JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **16**: 289–326 (2001)
DOI: 10.1002/jae.616

BOUNDS TESTING APPROACHES TO THE ANALYSIS OF LEVEL RELATIONSHIPS

M. HASHEM PESARAN,[a]* YONGCHEOL SHIN[b] AND RICHARD J. SMITH[c]

[a] *Trinity College, Cambridge CB2 1TQ, UK*
[b] *Department of Economics, University of Edinburgh, 50 George Square, Edinburgh EH8 9JY, UK*
[c] *Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK*

SUMMARY

This paper develops a new approach to the problem of testing the existence of a 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 univariate equilibrium correction mechanism. The asymptotic distributions of these statistics are non-standard under the null hypothesis that there exists no level relationship, irrespective of whether the regressors are $I(0)$ or $I(1)$. Two sets of asymptotic critical values are provided: one when all regressors are purely $I(1)$ and the other if they are all purely $I(0)$. These two sets of critical values provide a band covering all possible classifications of the regressors into purely $I(0)$, purely $I(1)$ 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 is demonstrated by a re-examination of the earnings equation included in the UK Treasury macroeconometric model.

1. INTRODUCTION

Over the past decade considerable attention has been paid in empirical economics to testing for the existence of relationships in levels between variables. In the main, this analysis has been based on the use of cointegration techniques. Two principal approaches have been adopted: the two-step residual-based procedure for testing the null of no-cointegration (see Engle and Granger, 1987; Phillips and Ouliaris, 1990) and the system-based reduced rank regression approach due to Johansen (1991, 1995). In addition, 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) have been considered. All of these methods concentrate on cases in which the underlying variables are integrated of order one. This inevitably involves a certain degree of pre-testing, thus introducing a further degree of uncertainty into the analysis of levels relationships. (See, for example, Cavanagh, Elliott and Stock, 1995.)

This paper proposes a new approach to testing for the existence of a relationship between variables in levels which is applicable irrespective of whether the underlying regressors are purely

* Correspondence to: M. H. Pesaran, Faculty of Economics and Politics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD. E-mail: hashem.pesaran@econ.cam.ac.uk
Contract/grant sponsor: ESRC; Contract/grant numbers: R000233608; R000237334.
Contract/grant sponsor: Isaac Newton Trust of Trinity College, Cambridge.

Received 16 February 1999
Revised 13 February 2001

$I(0)$, purely $I(1)$ 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 a *conditional* unrestricted equilibrium correction model (ECM). It is shown that the asymptotic distributions of both statistics are non-standard under the null hypothesis that there exists no relationship in levels between the included variables, irrespective of whether the regressors are purely $I(0)$, purely $I(1)$ 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 set of regressors which are a mixture of $I(0)/I(1)$ variables.

Two sets of asymptotic critical values are provided for the two polar cases which assume that all the regressors are, on the one hand, purely $I(1)$ and, on the other, purely $I(0)$. Since these two sets of critical values provide *critical value bounds* for all classifications of the regressors into purely $I(1)$, purely $I(0)$ or mutually cointegrated, 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 the integration/cointegration status of the underlying regressors. However, if the Wald or F-statistic falls inside these bounds, inference is inconclusive and knowledge of the order of the integration of the underlying variables is required before conclusive inferences can be made. A bounds procedure is also provided for the related cointegration test proposed by Banerjee *et al.* (1998) which is based on earlier contributions by Banerjee *et al.* (1986) and Kremers *et al.* (1992). Their test is based on the t-statistic associated with the coefficient of the lagged dependent variable in an unrestricted conditional ECM. The asymptotic distribution of this statistic is obtained for cases in which all regressors are purely $I(1)$, which is the primary context considered by these authors, as well as when the regressors are purely $I(0)$ or mutually cointegrated. The relevant critical value bounds for this t-statistic are also detailed.

The empirical relevance of the proposed bounds procedure is demonstrated in a re-examination of the earnings equation included in the UK Treasury macroeconometric model. This is a particularly relevant application because there is considerable doubt concerning the order of integration of variables such as the degree of unionization of the workforce, 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. Sargan's work is particularly noteworthy as it is some of the first to articulate and apply an ECM to wage rate determination. Sargan, however, did not consider the problem of testing for the existence of a levels relationship between real wages and its determinants.

The relationship in *levels* underlying the UK 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 *et al.* (1991). In order to identify our model as corresponding to the bargaining theory of wage determination, we require that the level of the unemployment rate enters the wage equation, but not vice versa; see Manning (1993). This assumption, of course, does not preclude the rate of change of earnings from entering the unemployment equation, or there being other level relationships between the remaining four variables. Our approach accommodates both of these possibilities.

A number of conditional ECMs in these five variables were estimated and we found that, if a sufficiently high order is selected for the lag lengths of the included variables, the hypothesis that there exists no relationship in levels between these variables is rejected, irrespective of whether they are purely $I(0)$, purely $I(1)$ or mutually cointegrated. Given a level relationship between these variables, the autoregressive distributed lag (ARDL) modelling approach (Pesaran and Shin, 1999) is used to estimate our preferred ECM of average earnings.

The plan of the paper is as follows. The vector autoregressive (VAR) model which underpins the analysis of this and later sections is set out in Section 2. This section also addresses the issues involved in testing for the existence of relationships in levels between variables. Section 3 considers the Wald statistic (or the F-statistic) for testing the hypothesis that there exists no level relationship between the variables under consideration and derives the associated asymptotic theory together with that for the t-statistic of Banerjee *et al.* (1998). Section 4 discusses the power properties of these tests. Section 5 describes the empirical application. Section 6 provides some concluding remarks. The Appendices detail proofs of results given in Sections 3 and 4.

The following notation is 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 VAR model of order p (VAR(p)):

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

where L is the lag operator, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are unknown $(k+1)$-vectors of intercept and trend coefficients, the $(k+1, k+1)$ matrix lag polynomial $\mathbf{\Phi}(L) = \mathbf{I}_{k+1} - \sum_{i=1}^{p} \mathbf{\Phi}_i L^i$ with $\{\Phi_i\}_{i=1}^{p}$ $(k+1, k+1)$ matrices of unknown coefficients; see Harbo *et al.* (1998) and Pesaran, Shin and Smith (2000), henceforth HJNR and PSS respectively. 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 $\mathbf{Z}_0 \equiv (\mathbf{z}_{1-p}, \ldots, \mathbf{z}_0)$. We assume:

Assumption 1. *The roots of* $|\mathbf{I}_{k+1} - \sum_{i=1}^{p} \mathbf{\Phi}_i z^i| = 0$ *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* $IN(\mathbf{0}, \mathbf{\Omega})$, $\mathbf{\Omega}$ *positive definite.*

Assumption 1 permits the elements of $\mathbf{z}_t$ to be purely $I(1)$, purely $I(0)$ 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 homoscedastic process; see, for example, PSS, Assumption 4.1.

We may re-express the lag polynomial $\mathbf{\Phi}(L)$ in vector equilibrium correction model (ECM) form; i.e. $\mathbf{\Phi}(L) \equiv -\mathbf{\Pi} L + \mathbf{\Gamma}(L)(1 - L)$ in which the long-run multiplier matrix is defined by $\mathbf{\Pi} \equiv$

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

$-(\mathbf{I}_{k+1} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i)$, and the short-run response matrix lag polynomial $\boldsymbol{\Gamma}(L) \equiv \mathbf{I}_{k+1} - \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i L^i$, $\boldsymbol{\Gamma}_i = -\sum_{j=i+1}^{p} \Phi_j$, $i = 1, \ldots, p-1$. Hence, the VAR(p) model (1) may be rewritten in vector ECM form as

$$\Delta \mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \boldsymbol{\Pi} \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t \qquad t = 1, 2, \ldots \tag{2}$$

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

$$\mathbf{a}_0 \equiv -\boldsymbol{\Pi}\boldsymbol{\mu} + (\boldsymbol{\Gamma} + \boldsymbol{\Pi})\boldsymbol{\gamma}, \mathbf{a}_1 \equiv -\boldsymbol{\Pi}\boldsymbol{\gamma} \tag{3}$$

and the sum of the short-run coefficient matrices $\boldsymbol{\Gamma} \equiv \mathbf{I}_m - \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i = -\boldsymbol{\Pi} + \sum_{i=1}^{p} i\boldsymbol{\Phi}_i$. As detailed in PSS, Section 2, if $\boldsymbol{\gamma} \neq \mathbf{0}$, the resultant constraints (3) on the trend coefficients $\mathbf{a}_1$ in (2) ensure that the deterministic trending behaviour of the level process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is invariant to the (cointegrating) rank of $\boldsymbol{\Pi}$; a similar result holds for the intercept of $\{\mathbf{z}_t\}_{t=1}^{\infty}$ if $\boldsymbol{\mu} \neq \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$. Consequently, critical regions defined in terms of the Wald and F-statistics suggested below are asymptotically *similar*.[2]

The focus of this paper is on the conditional modelling of the *scalar* variable 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

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

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

$$\varepsilon_{yt} = \boldsymbol{w}_{yx} \boldsymbol{\Omega}_{xx}^{-1} \boldsymbol{\varepsilon}_{xt} + u_t \tag{4}$$

where $u_t \sim IN(0, \omega_{uu})$, $\omega_{uu} \equiv \omega_{yy} - \boldsymbol{w}_{yx}\boldsymbol{\Omega}_{xx}^{-1}\boldsymbol{w}_{xy}$ and u_t is independent of $\boldsymbol{\varepsilon}_{xt}$. Substitution of (4) into (2) together with a similar partitioning of $\mathbf{a}_0 = (a_{y0}, \mathbf{a}_{x0}')'$, $\mathbf{a}_1 = (a_{y1}, \mathbf{a}_{x1}')'$, $\boldsymbol{\Pi} = (\boldsymbol{\pi}_y', \boldsymbol{\Pi}_x')'$, $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_y', \boldsymbol{\Gamma}_x')'$, $\boldsymbol{\Gamma}_i = (\boldsymbol{\gamma}_{yi}', \boldsymbol{\Gamma}_{xi}')'$, $i = 1, \ldots, p-1$, provides a conditional model for Δy_t in terms of $\mathbf{z}_{t-1}$, $\Delta \mathbf{x}_t$, $\Delta \mathbf{z}_{t-1}, \ldots$; i.e. the *conditional* ECM

$$\Delta y_t = c_0 + c_1 t + \boldsymbol{\pi}_{y.x} \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{w}' \Delta \mathbf{x}_t + u_t \qquad t = 1, 2, \ldots \tag{5}$$

where $\boldsymbol{w} \equiv \Omega_{xx}^{-1} \boldsymbol{w}_{xy}$, $c_0 \equiv a_{y0} - \boldsymbol{w}'\mathbf{a}_{x0}$, $c_1 \equiv a_{y1} - \boldsymbol{w}'\mathbf{a}_{x1}$, $\boldsymbol{\psi}_i' \equiv \boldsymbol{\gamma}_{yi} - \boldsymbol{w}'\boldsymbol{\Gamma}_{xi}$, $i = 1, \ldots, p-1$, and $\boldsymbol{\pi}_{y.x} \equiv \boldsymbol{\pi}_y - \boldsymbol{w}'\Pi_x$. The deterministic relations (3) are modified to

$$c_0 = -\boldsymbol{\pi}_{y.x}\boldsymbol{\mu} + (\boldsymbol{\gamma}_{y.x} + \boldsymbol{\pi}_{y.x})\boldsymbol{\gamma} \qquad c_1 = -\boldsymbol{\pi}_{y.x}\boldsymbol{\gamma} \tag{6}$$

where $\boldsymbol{\gamma}_{y.x} \equiv \boldsymbol{\gamma}_y - \boldsymbol{w}'\boldsymbol{\Gamma}_x$.

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

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

[2] See also Nielsen and Rahbek (1998) for an analysis of similarity issues in cointegrated systems.

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

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

In the application of Section 6, Assumption 3 is an identifying assumption for the bargaining theory of wage determination. Under Assumption 3,

$$\Delta \mathbf{x}_t = \mathbf{a}_{x0} + \mathbf{a}_{x1}t + \boldsymbol{\Pi}_{xx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{xi}\Delta \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_{xt} \qquad t = 1, 2, \ldots. \tag{7}$$

Thus, 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 (7); see Granger and Lin (1995).[3] Assumption 3 restricts consideration to cases in which there exists *at most* one conditional level relationship between y_t and $\mathbf{x}_t$, irrespective of the level of integration of the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$; see (10) below.[4]

Under Assumption 3, the conditional ECM (5) now becomes

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{w}'\Delta \mathbf{x}_t + u_t \tag{8}$$

$t = 1, 2, \ldots$, where

$$c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\mu} + [\boldsymbol{\gamma}_{y.x} + (\pi_{yy}, \boldsymbol{\pi}_{yx.x})]\boldsymbol{\gamma},\ c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\gamma} \tag{9}$$

and $\boldsymbol{\pi}_{yx.x} \equiv \boldsymbol{\pi}_{yx} - \boldsymbol{w}'\boldsymbol{\Pi}_{xx}$.[5]

The next assumption together with Assumptions 5a and 5b below which constrain the maximal order of integration of the system (8) and (7) to be unity defines the cointegration properties of the system.

Assumption 4. *The matrix* $\boldsymbol{\Pi}_{xx}$ *has rank r,* $0 \le r \le k$.

Under Assumption 4, from (7), we may express $\boldsymbol{\Pi}_{xx}$ as $\boldsymbol{\Pi}_{xx} = \boldsymbol{\alpha}_{xx}\boldsymbol{\beta}_{xx}'$, 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 (8) and (7) is unity, under Assumptions 1, 3 and 4, the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is mutually cointegrated of order r, $0 \le r \le k$. However, in contradistinction to, for example, Banerjee, Dolado and Mestre (1998), BDM henceforth, who concentrate on the case $r = 0$, we do not wish to impose an *a priori* specification of r.[6] When $\boldsymbol{\pi}_{xy} = \mathbf{0}$ *and* $\boldsymbol{\Pi}_{xx} = \mathbf{0}$, then $\mathbf{x}_t$ is weakly exogenous for π_{yy} and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx}$ in (8); see, for example,

[3] 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*.

[4] Assumption 3 may be straightforwardly assessed via a test for the exclusion of the lagged level y_{t-1} in (7). The asymptotic properties of such a test are the subject of current research.

[5] PSS and HJNR consider a similar model but where $\mathbf{x}_t$ is purely $I(1)$; that is, under the additional assumption $\boldsymbol{\Pi}_{xx} = \mathbf{0}$. If current and lagged values of a weakly exogenous purely $I(0)$ vector $\mathbf{w}_t$ are included as additional explanatory variables in (8), the lagged level vector $\mathbf{x}_{t-1}$ should be augmented to include the cumulated sum $\sum_{s=1}^{t-1} \mathbf{w}_s$ in order to preserve the asymptotic similarity of the statistics discussed below. See PSS, sub-section 4.3, and Rahbek and Mosconi (1999).

[6] BDM, pp. 277–278, also briefly discuss the case when $0 < r \le k$. However, in this circumstance, as will become clear below, the validity of the limiting distributional results for their procedure requires the imposition of further implicit and untested assumptions.

Johansen (1995, Theorem 8.1, p. 122). In the more general case where $\mathbf{\Pi}_{xx}$ is non-zero, as π_{yy} and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx} - \boldsymbol{w}'\mathbf{\Pi}_{xx}$ are variation-free from the parameters in (7), $\mathbf{x}_t$ is also weakly exogenous for the parameters of (8).

Note that under Assumption 4 the maximal cointegrating rank of the long-run multiplier matrix $\mathbf{\Pi}$ for the system (8) and (7) is $r+1$ and the minimal cointegrating rank of $\mathbf{\Pi}$ is r. The next assumptions provide the conditions for the maximal order of integration of the system (8) and (7) to be unity. First, we consider the requisite conditions for the case in which $rank(\mathbf{\Pi}) = r$. In this case, under Assumptions 1, 3 and 4, $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx} - \boldsymbol{\phi}'\mathbf{\Pi}_{xx} = \mathbf{0}'$ 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, $\mathbf{\Pi}$ has rank r and is given by

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

Hence, we may express $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ where $\boldsymbol{\alpha} = (\boldsymbol{\alpha}'_{yx}, \boldsymbol{\alpha}'_{xx})'$ and $\boldsymbol{\beta} = (\mathbf{0}, \boldsymbol{\beta}'_{xx})'$ are $(k+1, r)$ matrices of full column rank; cf. HJNR, p. 390. 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 $rank(\mathbf{\Pi}) = r$, the matrix $(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\mathbf{\Gamma}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ is full rank $k-r+1$, $0 \le r \le k$.*

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

Second, if the long-run multiplier matrix $\mathbf{\Pi}$ has rank $r+1$, then under Assumptions 1, 3 and 4, $\pi_{yy} \neq 0$ and $\mathbf{\Pi}$ may be expressed as $\mathbf{\Pi} = \boldsymbol{\alpha}_y\boldsymbol{\beta}'_y + \boldsymbol{\alpha\beta}'$, 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 $rank(\mathbf{\Pi}) = r+1$, the matrix $\boldsymbol{\alpha}^{\perp\prime}\mathbf{\Gamma}\boldsymbol{\beta}^{\perp}$ is full rank $k-r$, $0 \le r \le k$.*

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

Using (A.1) in Appendix A, it is easily seen that $\boldsymbol{\pi}_{y.x}(\mathbf{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \boldsymbol{\pi}_{y.x}\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$, where $\{\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t\}$ is a mean zero stationary process. Therefore, under Assumptions 1, 3, 4 and 5b, that is, $\pi_{yy} \neq 0$, it immediately follows that there exists a *conditional level relationship* between y_t and $\mathbf{x}_t$ defined by

$$y_t = \theta_0 + \theta_1 t + \boldsymbol{\theta}\mathbf{x}_t + v_t,\ t = 1, 2, \ldots \tag{10}$$

where $\theta_0 \equiv \boldsymbol{\pi}_{y.x}\boldsymbol{\mu}/\pi_{yy}$, $\theta_1 \equiv \boldsymbol{\pi}_{y.x}\boldsymbol{\gamma}/\pi_{yy}$, $\boldsymbol{\theta} \equiv -\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ and $v_t = \boldsymbol{\pi}_{y.x}\mathbf{C}^*(L)\varepsilon_t/\pi_{yy}$, also a zero mean stationary process. If $\boldsymbol{\pi}_{yx.x} = \alpha_{yy}\boldsymbol{\beta}'_{yx} + (\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx} \neq \mathbf{0}'$, the level relationship between y_t and $\mathbf{x}_t$ is *non-degenerate*. Hence, from (10), $y_t \sim I(0)$ if $rank(\boldsymbol{\beta}_{yx}, \boldsymbol{\beta}_{xx}) = r$ and $y_t \sim I(1)$ if $rank(\boldsymbol{\beta}_{yx}, \boldsymbol{\beta}_{xx}) = r+1$. In the former case, $\boldsymbol{\theta}$ is the vector of conditional *long-run* multipliers and, in this sense, (10) may be interpreted as a conditional *long-run* level relationship between y_t and $\mathbf{x}_t$, whereas, in the latter, because the processes $\{y_t\}_{t=1}^{\infty}$ and $\{\mathbf{x}_t\}_{t=1}^{\infty}$ are *cointegrated*, (10) represents the conditional *long-run* level relationship between y_t and $\mathbf{x}_t$. Two *degenerate* cases arise. First,

if $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$, clearly, from (10), y_t is (trend) stationary or $y_t \sim I(0)$ whatever the value of r. Consequently, the differenced variable Δy_t depends only on its own lagged level y_{t-1} in the conditional ECM (8) and *not* on the lagged levels $\mathbf{x}_{t-1}$ of the forcing variables. Second, if $\pi_{yy} = 0$, that is, Assumption 5a holds, and $\boldsymbol{\pi}_{yx.x} = (\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx} \neq \mathbf{0}'$, as $rank(\boldsymbol{\Pi}) = r$, $\boldsymbol{\pi}_{yx.x} = (\boldsymbol{\phi} - \boldsymbol{w})'\boldsymbol{\alpha}_{xx}\boldsymbol{\beta}'_{xx}$ which, from the above, yields $\boldsymbol{\pi}_{yx.x}(\mathbf{x}_t - \boldsymbol{\mu}_x - \boldsymbol{\gamma}_x t) = \boldsymbol{\pi}_{y.x}\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t,\ t = 1, 2, \ldots,$ where $\boldsymbol{\mu} = (\mu_y, \boldsymbol{\mu}'_x)'$ and $\boldsymbol{\gamma} = (\gamma_y, \boldsymbol{\gamma}'_x)'$ are partitioned conformably with $\mathbf{z}_t = (y_t, \mathbf{x}'_t)'$. Thus, in (8), Δy_t depends only on the lagged level $\mathbf{x}_{t-1}$ through the linear combination $(\boldsymbol{\phi} - \boldsymbol{w})'\boldsymbol{\alpha}_{xx}$ of the lagged mutually cointegrating relations $\boldsymbol{\beta}'_{xx}\mathbf{x}_{t-1}$ for the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$. Consequently, $y_t \sim I(1)$ whatever the value of r. Finally, if *both* $\pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$, there are no level effects in the conditional ECM (8) with no possibility of any level relationship between y_t and $\mathbf{x}_t$, degenerate or otherwise, and, again, $y_t \sim I(1)$ whatever the value of r.

Therefore, in order to test for the absence of level effects in the conditional ECM (8) and, more crucially, the absence of a level relationship between y_t and $\mathbf{x}_t$, the emphasis in this paper is a test of the *joint* hypothesis $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ in (8).[7,8] In contradistinction, the approach of BDM may be described in terms of (8) using Assumption 5b:

$$\Delta y_t = c_0 + c_1 t + \alpha_{yy}(\beta_{yy} y_{t-1} + \boldsymbol{\beta}'_{yx}\mathbf{x}_{t-1}) + (\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{w}'\Delta\mathbf{x}_t + u_t \tag{11}$$

BDM test for the exclusion of y_{t-1} in (11) when $r = 0$, that is, $\boldsymbol{\beta}_{xx} = \mathbf{0}$ in (11) or $\boldsymbol{\Pi}_{xx} = \mathbf{0}$ in (7) and, thus, $\{\mathbf{x}_t\}$ is purely $I(1)$; cf. HJNR and PSS.[9] Therefore, BDM consider the hypothesis $\alpha_{yy} = 0$ (or $\pi_{yy} = 0$).[10] More generally, when $0 < r \leq k$, BDM require the imposition of the untested subsidiary hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$; that is, the limiting distribution of the BDM test is obtained under the *joint* hypothesis $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}$ in (8).

In the following sections of the paper, we focus on (8) and differentiate between five cases of interest delineated according to how the deterministic components are specified:

- **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 (8) becomes

$$\Delta y_t = \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{w}'\Delta\mathbf{x}_t + u_t \tag{12}$$

- **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

$$\Delta y_t = \pi_{yy}(y_{t-1} - \mu_y) + \boldsymbol{\pi}_{yx.x}(\mathbf{x}_{t-1} - \boldsymbol{\mu}_x) + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{w}'\Delta\mathbf{x}_t + u_t \tag{13}$$

[7] This joint hypothesis may be justified by the application of Roy's union-intersection principle to tests of $\pi_{yy} = 0$ in (8) 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 (8).

[8] A related approach to that of this paper is Hansen's (1995) test for a unit root in a univariate time series which, in our context, would require the imposition of the subsidiary hypothesis $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$.

[9] The BDM test is based on earlier contributions of Kremers *et al.* (1992), Banerjee *et al.* (1993), and Boswijk (1994).

[10] Partitioning $\boldsymbol{\Gamma}_{xi} = (\boldsymbol{\gamma}_{xy,i}, \boldsymbol{\Gamma}_{xx,i})$, $i = 1, \ldots, p-1$, conformably with $\mathbf{z}_t = (y_t, \mathbf{x}'_t)'$, BDM also set $\boldsymbol{\gamma}_{xy,i} = \mathbf{0}$, $i = 1, \ldots, p-1$, which implies $\boldsymbol{\gamma}_{xy} = \mathbf{0}$, where $\boldsymbol{\Gamma}_x = (\boldsymbol{\gamma}_{xy}, \boldsymbol{\Gamma}_{xx})$; that is, Δy_t does *not Granger cause* $\Delta\mathbf{x}_t$.

- **Case III** (unrestricted intercepts; no trends) $c_0 \neq 0$ and $c_1 = 0$. Again, $\boldsymbol{\gamma} = \mathbf{0}$. Now, the intercept restriction $c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\mu}$ is ignored and the ECM is

$$\Delta y_t = c_0 + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{w}'\Delta \mathbf{x}_t + u_t \tag{14}$$

- **Case IV** (unrestricted intercepts; restricted trends) $c_0 \neq 0$ and $c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\gamma}$.

$$\Delta y_t = c_0 + \pi_{yy}(y_{t-1} - \gamma_y t) + \boldsymbol{\pi}_{yx.x}(\mathbf{x}_{t-1} - \boldsymbol{\gamma}_x t) + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{w}'\Delta \mathbf{x}_t + u_t \tag{15}$$

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

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{w}'\Delta \mathbf{x}_t + u_t \tag{16}$$

It should be emphasized 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, 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 (9).

In the following exposition, we concentrate on Case IV, that is, (15), which may be specialized to yield the remainder.

3. BOUNDS TESTS FOR A LEVEL RELATIONSHIPS

In this section we develop bounds procedures for testing for the existence of a level relationship between y_t and $\mathbf{x}_t$ using (12)–(16); see (10). The main approach taken here, cf. Engle and Granger (1987) and BDM, is to test for the *absence* of *any* level relationship between y_t and $\mathbf{x}_t$ via the exclusion of the lagged level variables y_{t-1} and $\mathbf{x}_{t-1}$ in (12)–(16). Consequently, we define the constituent null hypotheses $H_0^{\pi_{yy}} : \pi_{yy} = 0$, $H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$, and alternative hypotheses $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$, $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$. Hence, the joint null hypothesis of interest in (12)–(16) is given by:

$$H_0 = H_0^{\pi_{yy}} \cap H_0^{\pi_{yx.x}} \tag{17}$$

and the alternative hypothesis is correspondingly stated as:

$$H_1 = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}} \tag{18}$$

However, as indicated in Section 2, not only does the alternative hypothesis H_1 of (17) cover the case of interest in which $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ but also permits $\pi_{yy} \neq 0$, $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ and $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$; cf. (8). That is, the possibility of *degenerate* level relationships between y_t and $\mathbf{x}_t$ is admitted under H_1 of (18). We comment further on these alternatives at the end of this section.

For ease of exposition, we consider Case IV and rewrite (15) in matrix notation as

$$\Delta\mathbf{y} = \iota_T c_0 + \mathbf{Z}^*_{-1}\boldsymbol{\pi}^*_{y.x} + \Delta\mathbf{Z}_-\boldsymbol{\psi} + \mathbf{u} \tag{19}$$

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

$$\hat{\boldsymbol{\pi}}^*_{y.x} \equiv (\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\tilde{\mathbf{Z}}^*_{-1})^{-1}\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} \tag{20}$$

where $\tilde{\mathbf{Z}}^*_{-1} \equiv \overline{\mathbf{P}}_\iota\mathbf{Z}^*_{-1}$, $\widetilde{\Delta\mathbf{Z}}_- \equiv \overline{\mathbf{P}}_\iota\Delta\mathbf{Z}_-$, $\widetilde{\Delta\mathbf{y}} \equiv \overline{\mathbf{P}}_\iota\Delta\mathbf{y}$, $\overline{\mathbf{P}}_\iota \equiv \mathbf{I}_T - \iota_T(\iota'_T\iota_T)^{-1}\iota'_T$ and $\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-} \equiv \mathbf{I}_T - \widetilde{\Delta\mathbf{Z}}_-(\widetilde{\Delta\mathbf{Z}}'_-\widetilde{\Delta\mathbf{Z}}_-)^{-1}\widetilde{\Delta\mathbf{Z}}'_-$. The Wald and the F-statistics for testing the null hypothesis H_0 of (17) against the alternative hypothesis H_1 of (18) are respectively:

$$W \equiv \hat{\boldsymbol{\pi}}^{*'}_{y.x}\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\tilde{\mathbf{Z}}^*_{-1}\hat{\boldsymbol{\pi}}^*_{y.x}/\hat{\omega}_{uu}, \quad F \equiv \frac{W}{k+2} \tag{21}$$

where $\hat{\omega}_{uu} \equiv (T-m)^{-1}\sum_{t=1}^T \tilde{u}_t^2$, $m \equiv (k+1)(p+1)+1$ is the number of estimated coefficients and $\tilde{u}_t$, $t = 1, 2, \ldots, T$, are the least squares (LS) residuals from (19).

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 $\tilde{\mathbf{W}}_{k-r+1}(a) \equiv \mathbf{W}_{k-r+1}(a) - \int_0^1 \mathbf{W}_{k-r+1}(a)da$, and de-meaned and de-trended $(k-r+1)$-vector standard Brownian motion $\hat{\mathbf{W}}_{k-r+1}(a) \equiv \tilde{\mathbf{W}}_{k-r+1}(a) - 12\left(a-\frac{1}{2}\right)\int_0^1\left(a-\frac{1}{2}\right)\tilde{\mathbf{W}}_{k-r+1}(a)da$, 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 (17), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (21) has the representation*

$$W \Rightarrow \mathbf{z}'_r\mathbf{z}_r + \int_0^1 dW_u(a)\mathbf{F}_{k-r+1}(a)'\left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'da\right)^{-1}\int_0^1 \mathbf{F}_{k-r+1}(a)dW_u(a) \tag{22}$$

where $\mathbf{z}_r \sim N(\mathbf{0}, \mathbf{I}_r)$ is distributed independently of the second term in (22) and

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

$r = 0, \ldots, k$, and Cases I–V are defined in (12)–(16), $a \in [0,1]$.

The asymptotic distribution of the Wald statistic W of (21) depends on the dimension and cointegration rank of the forcing variables $\{\mathbf{x}_t\}$, k and r respectively. In Case IV, referring to (11), the first component in (22), $\mathbf{z}_r'\mathbf{z}_r \sim \chi^2(r)$, corresponds to testing for the exclusion of the r-dimensional stationary vector $\boldsymbol{\beta}_{xx}'\mathbf{x}_{t-1}$, that is, the hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$, whereas the second term in (22), 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 specialize Theorem 3.1 to the two polar cases in which, first, the process for the forcing variables $\{\mathbf{x}_t\}$ is purely integrated of order zero, that is, $r = k$ and $\boldsymbol{\Pi}_{xx}$ is of full rank, and, second, the $\{\mathbf{x}_t\}$ process is not mutually cointegrated, $r = 0$, and, hence, the $\{\mathbf{x}_t\}$ process is purely 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 (17), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (21) has the representation*

$$W \Rightarrow \mathbf{z}_k'\mathbf{z}_k + \frac{(\int_0^1 F(a)dW_u(a))^2}{(\int_0^1 F(a)^2 da)} \tag{23}$$

where $\mathbf{z}_k \sim N(\mathbf{0}, I_k)$ is distributed independently of the second term in (23) and

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

$r = 0, \ldots, k$, where Cases I–V are defined in (12)–(16), $a \in [0, 1]$.

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 (17), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (21) has the representation*

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

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

In practice, however, it is unlikely that one would possess *a priori* knowledge of the rank r of $\boldsymbol{\Pi}_{xx}$; 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 (12)–(16) predicated on *a prior* determination of the cointegration rank r in (7) 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 $0 \le r \le k$.[11] 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. Asymptotic critical value bounds for the F-statistics covering Cases I–V are set out in Tables CI(i)–CI(v) for sizes 0.100, 0.050, 0.025 and 0.010; the lower bound values assume that the forcing variables $\{\mathbf{x}_t\}$ are purely $I(0)$, and the upper bound values assume that $\{\mathbf{x}_t\}$ are purely $I(1)$.[12]

Hence, we suggest a *bounds procedure* to test $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (17) within the conditional ECMs (12)–(16). 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 conditional ECMs (12)–(16), derived from the underlying VAR(p) model (2), may also be interpreted as an autoregressive distributed lag model of orders $(p, p, \ldots, p)$ (ARDL$(p, \ldots, p)$). However, one could also allow for differential lag lengths on the lagged variables y_{t-i} and $\mathbf{x}_{t-i}$ in (2) to arrive at, for example, an ARDL$(p, p_1, \ldots, p_k)$ model 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 (12)–(16) as well as allowing for short-run feedbacks from the lagged dependent variables, Δy_{t-i}, $i = 1, \ldots, p$, to $\Delta \mathbf{x}_t$ in (7). 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), HJNR, Johansen (1992, 1995), PSS, and Urbain (1992), where it is assumed in addition that $\boldsymbol{\Pi}_{xx} = \mathbf{0}$ or $\mathbf{x}_t$ is purely $I(1)$ in (7).

To conclude this section, we reconsider the approach of BDM. There are three scenarios for the deterministics given by (12), (14) and (16). Note that the restrictions on the deterministics' coefficients (9) are ignored in Cases II of (13) and IV of (15) and, thus, Cases II and IV are now subsumed by Cases III of (14) and V of (16) respectively. As noted below (11), BDM impose but do not test the implicit hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$; 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 (17). BDM 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. For example, in Case V, they consider the t-statistic

$$t_{\pi_{yy}} = \frac{\hat{\mathbf{y}}_{-1}' \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}, \hat{\mathbf{X}}_{-1}} \widehat{\Delta \mathbf{y}}}{\hat{\omega}_{uu}^{1/2} (\hat{\mathbf{y}}_{-1}' \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}, \hat{\mathbf{X}}_{-1}} \hat{\mathbf{y}}_{-1})^{1/2}} \tag{24}$$

where $\hat{\omega}_{uu}$ is defined in the line after (21), $\widehat{\Delta \mathbf{y}} \equiv \overline{\mathbf{P}}_{\iota_T, \tau_T} \Delta \mathbf{y}$, $\hat{\mathbf{y}}_{-1} \equiv \overline{\mathbf{P}}_{\iota_T, \tau_T} \mathbf{y}_{-1}$, $\mathbf{y}_{-1} \equiv (y_0, \ldots, y_{T-1})'$, $\hat{\mathbf{X}}_{-1} \equiv \overline{\mathbf{P}}_{\iota_T, \tau_T} \mathbf{X}_{-1}$, $\mathbf{X}_{-1} \equiv (\mathbf{x}_0, \ldots, \mathbf{x}_{T-1})'$, $\widehat{\Delta \mathbf{Z}}_{-} \equiv \overline{\mathbf{P}}_{\iota_T, \tau_T} \Delta \mathbf{Z}_{-}$, $\overline{\mathbf{P}}_{\iota_T, \tau_T} \equiv \overline{\mathbf{P}}_{\iota_T} - \overline{\mathbf{P}}_{\iota_T} \boldsymbol{\tau}_T (\boldsymbol{\tau}_T' \overline{\mathbf{P}}_{\iota_T} \boldsymbol{\tau}_T)^{-1} \boldsymbol{\tau}_T' \overline{\mathbf{P}}_{\iota_T}$, $\overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}, \hat{\mathbf{X}}_{-1}} = \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}} - \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}} \hat{\mathbf{X}}_{-1} (\hat{\mathbf{X}}_{-1}' \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}} \hat{\mathbf{X}}_{-1})^{-1} \hat{\mathbf{X}}_{-1}' \overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}}$ and $\overline{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_{-}} \equiv \mathbf{I}_T - \widehat{\Delta \mathbf{Z}}_{-} (\widehat{\Delta \mathbf{Z}}_{-}' \widehat{\Delta \mathbf{Z}}_{-})^{-1} \widehat{\Delta \mathbf{Z}}_{-}'$.

[11] 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 $0 \le r \le k$.

[12] 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.

Table CI. Asymptotic critical value bounds for the F-statistic. Testing for the existence of a levels relationship[a]

Table CI(i) Case I: No intercept and no trend

	0.100		0.050		0.025		0.010		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 CI(ii) Case II: Restricted intercept and no trend

	0.100		0.050		0.025		0.010		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	3.80	3.80	4.60	4.60	5.39	5.39	6.44	6.44	2.03	2.03	1.77	1.77
1	3.02	3.51	3.62	4.16	4.18	4.79	4.94	5.58	1.69	2.02	1.01	1.25
2	2.63	3.35	3.10	3.87	3.55	4.38	4.13	5.00	1.52	2.02	0.69	0.96
3	2.37	3.20	2.79	3.67	3.15	4.08	3.65	4.66	1.41	2.02	0.52	0.78
4	2.20	3.09	2.56	3.49	2.88	3.87	3.29	4.37	1.34	2.01	0.42	0.65
5	2.08	3.00	2.39	3.38	2.70	3.73	3.06	4.15	1.29	2.00	0.35	0.56
6	1.99	2.94	2.27	3.28	2.55	3.61	2.88	3.99	1.26	2.00	0.30	0.49
7	1.92	2.89	2.17	3.21	2.43	3.51	2.73	3.90	1.23	2.01	0.26	0.44
8	1.85	2.85	2.11	3.15	2.33	3.42	2.62	3.77	1.21	2.01	0.23	0.40
9	1.80	2.80	2.04	3.08	2.24	3.35	2.50	3.68	1.19	2.01	0.21	0.36
10	1.76	2.77	1.98	3.04	2.18	3.28	2.41	3.61	1.17	2.00	0.19	0.33

Table CI(iii) Case III: Unrestricted intercept and no trend

	0.100		0.050		0.025		0.010		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

(*Continued overleaf*)

Table CI. (*Continued*)

Table CI(iv) Case IV: Unrestricted intercept and restricted trend

	0.100		0.050		0.025		0.010		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	5.37	5.37	6.29	6.29	7.14	7.14	8.26	8.26	3.17	3.17	2.68	2.68
1	4.05	4.49	4.68	5.15	5.30	5.83	6.10	6.73	2.45	2.77	1.41	1.65
2	3.38	4.02	3.88	4.61	4.37	5.16	4.99	5.85	2.09	2.57	0.92	1.20
3	2.97	3.74	3.38	4.23	3.80	4.68	4.30	5.23	1.87	2.45	0.67	0.93
4	2.68	3.53	3.05	3.97	3.40	4.36	3.81	4.92	1.72	2.37	0.51	0.76
5	2.49	3.38	2.81	3.76	3.11	4.13	3.50	4.63	1.62	2.31	0.42	0.64
6	2.33	3.25	2.63	3.62	2.90	3.94	3.27	4.39	1.54	2.27	0.35	0.55
7	2.22	3.17	2.50	3.50	2.76	3.81	3.07	4.23	1.48	2.24	0.31	0.49
8	2.13	3.09	2.38	3.41	2.62	3.70	2.93	4.06	1.44	2.22	0.27	0.44
9	2.05	3.02	2.30	3.33	2.52	3.60	2.79	3.93	1.40	2.20	0.24	0.40
10	1.98	2.97	2.21	3.25	2.42	3.52	2.68	3.84	1.36	2.18	0.22	0.36

Table CI(v) Case V: Unrestricted intercept and unrestricted trend

	0.100		0.050		0.025		0.010		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

[a] The critical values are computed via stochastic simulations using $T = 1000$ and 40,000 replications for the F-statistic for testing $\boldsymbol{\phi} = \mathbf{0}$ in the regression: $\Delta y_t = \boldsymbol{\phi}'\mathbf{z}_{t-1} + \mathbf{a}'\mathbf{w}_t + \xi_t$, $t = 1, \ldots, T$, where $\mathbf{x}_t = (x_{1t}, \ldots, x_{kt}')$ and

$$\left\{\begin{array}{ll} \mathbf{z}_{t-1} = (y_{t-1}, \mathbf{x}'_{t-1})', \mathbf{w}_t = 0 & \textbf{Case I} \\ \mathbf{z}_{t-1} = (y_{t-1}, \mathbf{x}'_{t-1}, 1)', \mathbf{w}_t = 0 & \textbf{Case II} \\ \mathbf{z}_{t-1} = (y_{t-1}, \mathbf{x}'_{t-1})', \mathbf{w}_t = 1 & \textbf{Case III} \\ \mathbf{z}_{t-1} = (y_{t-1}, \mathbf{x}'_{t-1}, t)', \mathbf{w}_t = 1 & \textbf{Case IV} \\ \mathbf{z}_{t-1} = (y_{t-1}, \mathbf{x}'_{t-1})', \mathbf{w}_t = (1, t)' & \textbf{Case V} \end{array}\right\}$$

The variables y_t and $\mathbf{x}_t$ are generated from $y_t = y_{t-1} + \varepsilon_{1t}$ and $\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_{2t}$, $t = 1, \ldots, T$, where $y_0 = 0$, $\mathbf{x}_0 = 0$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \boldsymbol{\varepsilon}'_{2t})'$ is drawn as $(k+1)$ independent standard normal variables. If $\mathbf{x}_t$ is purely $I(1)$, $\mathbf{P} = \mathbf{I}_k$ whereas $\mathbf{P} = \mathbf{0}$ if $\mathbf{x}_t$ is purely $I(0)$. The critical values for $k = 0$ correspond to the squares of the critical values of Dickey and Fuller's (1979) unit root t-statistics for Cases I, III and V, while they match those for Dickey and Fuller's (1981) 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 purely $I(0)$, while the columns headed '$I(1)$' refer to the upper bound obtained when $\mathbf{x}_t$ is purely $I(1)$.

Theorem 3.2 *(Limiting distribution of $t_{\pi_{yy}}$). If Assumptions 1-4 and 5a hold and $\boldsymbol{\gamma}_{xy} = \mathbf{0}$, where $\boldsymbol{\Gamma}_x = (\boldsymbol{\gamma}_{xy}, \boldsymbol{\Gamma}_{xx})$, then under $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (17), as $T \to \infty$, the asymptotic distribution of the t-statistic $t_{\pi_{yy}}$ of (24) has the representation*

$$\int_0^1 dW_u(a) F_{k-r}(a) \left(\int_0^1 F_{k-r}(a)^2 \, da \right)^{-1/2} \tag{25}$$

where

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

$r = 0, \ldots, k$, and Cases I, III and V are defined in (12), (14) and (16), $a \in [0, 1]$.

The form of the asymptotic representation (25) 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).

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

Corollary 3.3 *(Limiting distribution of $t_{\pi_{yy}}$ 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 (17), as $T \to \infty$, the asymptotic distribution of the t-statistic $t_{\pi_{yy}}$ of (24) has the representation*

$$\int_0^1 dW_u(a) F(a) \left(\int_0^1 F(a)^2 \, da \right)^{-1/2}$$

where

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

and Cases I, III and V are defined in (12), (14) and (16), $a \in [0, 1]$.

Corollary 3.4 *(Limiting distribution of $t_{\pi_{yy}}$ if $\{\mathbf{x}_t\} \sim I(1)$). If Assumptions 1-4 and 5a hold, $\boldsymbol{\gamma}_{xy} = \mathbf{0}$, where $\boldsymbol{\Gamma}_x = (\boldsymbol{\gamma}_{xy}, \boldsymbol{\Gamma}_{xx})$, and $r = 0$, that is, $\{\mathbf{x}_t\} \sim I(1)$, then under $H_0^{\pi_{yy}} : \pi_{yy} = 0$, as $T \to \infty$, the asymptotic distribution of the t-statistic $t_{\pi_{yy}}$ of (24) has the representation*

$$\int_0^1 dW_u(a) F_k(a) \left(\int_0^1 F_k(a)^2 \, da \right)^{-1/2}$$

where $F_k(a)$ is defined in Theorem 3.2 for Cases I, III and V, $a \in [0, 1]$.

As above, it may be shown by simulation that the asymptotic critical values obtained from Corollaries 3.3 ($r = k$ and $\{\mathbf{x}_t\}$ is purely $I(0)$) and 3.4 ($r = 0$ and $\{\mathbf{x}_t\}$ is purely $I(1)$) 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 $t_{\pi_{yy}}$ for the exclusion of y_{t-1} in the conditional ECMs (12), (14) and (16) without prior knowledge of the cointegrating rank r.[13] These asymptotic critical value bounds are given in Tables CII(i), CII(iii) and CII(v) for Cases I, III and V for sizes 0.100, 0.050, 0.025 and 0.010.

As is emphasized in the Proof of Theorem 3.2 given in Appendix A, if the asymptotic analysis for the t-statistic $t_{\pi_{yy}}$ of (24) is conducted under $H_0^{\pi_{yy}}:\pi_{yy}=0$ only, the resultant limit distribution for $t_{\pi_{yy}}$ depends on the nuisance parameter $\boldsymbol{w}-\boldsymbol{\phi}$ in addition to the cointegrating rank r, 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 $i=1,\ldots,p-1$, then the limit distribution also is dependent on the nuisance parameter $\boldsymbol{\gamma}_{xy}/(\gamma_{yy}-\boldsymbol{\phi}'\boldsymbol{\gamma}_{xy})$; see Appendix A. Consequently, in general, where $\boldsymbol{w}\neq\boldsymbol{\phi}$ or $\boldsymbol{\gamma}_{xy}\neq\mathbf{0}$,

Table CII. Asymptotic critical value bounds of the t-statistic. Testing for the existence of a levels relationship[a]

Table CII(i): Case I: No intercept and no trend

	0.100		0.050		0.025		0.010		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	−2.24	−3.64	−2.58	−3.97	−0.42	−1.71	0.98	1.09
4	−1.62	−3.26	−1.95	−3.60	−2.24	−3.89	−2.58	−4.23	−0.42	−1.98	0.98	1.07
5	−1.62	−3.49	−1.95	−3.83	−2.24	−4.12	−2.58	−4.44	−0.42	−2.22	0.98	1.05
6	−1.62	−3.70	−1.95	−4.04	−2.24	−4.34	−2.58	−4.67	−0.42	−2.43	0.98	1.04
7	−1.62	−3.90	−1.95	−4.23	−2.24	−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	−2.58	−5.07	−0.42	−2.81	0.98	1.04
9	−1.62	−4.26	−1.95	−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 CII(iii) Case III: Unrestricted intercept and no trend

	0.100		0.050		0.025		0.010		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	0.72	0.71
1	−2.57	−2.91	−2.86	−3.22	−3.13	−3.50	−3.43	−3.82	−1.53	−1.80	0.72	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	−2.86	−4.19	−3.13	−4.46	−3.43	−4.79	−1.53	−2.65	0.72	0.92
6	−2.57	−4.04	−2.86	−4.38	−3.13	−4.66	−3.43	−4.99	−1.53	−2.83	0.72	0.93
7	−2.57	−4.23	−2.86	−4.57	−3.13	−4.85	−3.43	−5.19	−1.53	−3.00	0.72	0.94
8	−2.57	−4.40	−2.86	−4.72	−3.13	−5.02	−3.43	−5.37	−1.53	−3.16	0.72	0.96
9	−2.57	−4.56	−2.86	−4.88	−3.13	−5.18	−3.42	−5.54	−1.53	−3.31	0.72	0.96
10	−2.57	−4.69	−2.86	−5.03	−3.13	−5.34	−3.43	−5.68	−1.53	−3.46	0.72	0.96

(*Continued overleaf*)

[13] Although Corollary 3.3 does not require $\boldsymbol{\gamma}_{xy}=\mathbf{0}$ and $H_0^{\pi_{yx.x}}:\boldsymbol{\pi}_{yx.x}=\mathbf{0}'$ is automatically satisfied under the conditions of Corollary 3.4, the simulation critical value bounds result requires $\boldsymbol{\gamma}_{xy}=\mathbf{0}$ and $H_0^{\pi_{yx.x}}:\boldsymbol{\pi}_{yx.x}=\mathbf{0}'$ for $0<r<k$.

Table CII. (*Continued*)

Table CII(v) Case V: Unrestricted intercept and unrestricted trend

	0.100		0.050		0.025		0.010		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	−3.65	−3.66	−3.96	−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	−3.65	−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	−2.18	−3.49	0.57	0.90
9	−3.13	−4.82	−3.41	−5.15	−3.65	−5.44	−3.96	−5.79	−2.18	−3.62	0.57	0.91
10	−3.13	−4.96	−3.41	−5.29	−3.65	−5.59	−3.96	−5.94	−2.18	−3.75	0.57	0.92

[a] The critical values are computed via stochastic simulations using $T = 1000$ and 40 000 replications for the t-statistic for testing $\phi = 0$ in the regression: $\Delta y_t = \phi y_{t-1} + \boldsymbol{\delta}'\mathbf{x}_{t-1} + \mathbf{a}'\mathbf{w}_t + \xi_t$, $t = 1, \ldots, T$, where $\mathbf{x}_t = (x_{1t}, \ldots, x_{kt})'$ and

$$\left\{ \begin{array}{ll} \mathbf{w}_t = 0 & \textbf{Case I} \\ \mathbf{w}_t = 1 & \textbf{Case III} \\ \mathbf{w}_t = (1, t)' & \textbf{Case V} \end{array} \right\}$$

The variables y_t and $\mathbf{x}_t$ are generated from $y_t = y_{t-1} + \varepsilon_{1t}$ and $\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_{2t}$, $t = 1, \ldots, T$, where $y_0 = 0$, $\mathbf{x}_0 = 0$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \boldsymbol{\varepsilon}'_{2t})'$ is drawn as $(k+1)$ independent standard normal variables. If $\mathbf{x}_t$ is purely $I(1)$, $\mathbf{P} = \mathbf{I}_k$ whereas $\mathbf{P} = \mathbf{0}$ if $\mathbf{x}_t$ is purely $I(0)$. The critical values for $k = 0$ correspond to those of Dickey and Fuller's (1979) unit root t-statistics. The columns headed '$I(0)$' refer to the lower critical values bound obtained when $\mathbf{x}_t$ is purely $I(0)$, while the columns headed '$I(1)$' refer to the upper bound obtained when $\mathbf{x}_t$ is purely $I(1)$.

although the t-statistic $t_{\pi_{yy}}$ has a well-defined limiting distribution under $H_0^{\pi_{yy}} : \pi_{yy} = 0$, the above bounds testing procedure for $H_0^{\pi_{yy}} : \pi_{yy} = 0$ based on $t_{\pi_{yy}}$ is not asymptotically *similar*.[14]

Consequently, in the light of the consistency results for the above statistics discussed in Section 4, see Theorems 4.1, 4.2 and 4.4, we suggest the following procedure for ascertaining the existence of a level relationship between y_t and $\mathbf{x}_t$: test H_0 of (17) using the bounds procedure based on the Wald or F-statistic of (21) from Corollaries 3.1 and 3.2: (a) if H_0 is not rejected, proceed no further; (b) if H_0 is rejected, test $H_0^{\pi_{yy}} : \pi_{yy} = 0$ using the bounds procedure based on the t-statistic $t_{\pi_{yy}}$ of (24) from Corollaries 3.3 and 3.4. If $H_0^{\pi_{yy}} : \pi_{yy} = 0$ is false, a large value of $t_{\pi_{yy}}$ should result, at least asymptotically, confirming the existence of a level relationship between y_t and $\mathbf{x}_t$, which, however, may be degenerate (if $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$).

4. THE ASYMPTOTIC POWER OF THE BOUNDS PROCEDURE

This section first demonstrates that the proposed bounds testing procedure based on the Wald statistic of (21) described in Section 3 is consistent. Second, it derives the asymptotic distribution

[14] In principle, the asymptotic distribution of $t_{\pi_{yy}}$ under $H_0^{\pi_{yy}} : \pi_{yy} = 0$ may be simulated from the limiting representation given in the Proof of Theorem 3.2 of Appendix A after substitution of consistent estimators for $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}_{xy}^{\phi} \equiv \boldsymbol{\gamma}_{xy}/\gamma_{yy.x}^{\phi}$ under $H_0^{\pi_{yy}} : \pi_{yy} = 0$, where $\gamma_{yy.x}^{\phi} \equiv \gamma_{yy} - \boldsymbol{\phi}'\boldsymbol{\gamma}_{xy}$. Although such estimators may be obtained straightforwardly, unfortunately, they necessitate the use of parameter estimators from the marginal ECM (7) for $\{\mathbf{x}_t\}_{t=1}^{\infty}$.

of the Wald statistic of (21) under a sequence of local alternatives. Finally, we show that the bounds procedure based on the t-statistic of (24) is consistent.

In the discussion of the consistency of the bounds test procedure based on the Wald statistic of (21), because the rank of the long-run multiplier matrix $\mathbf{\Pi}$ may be either r or $r+1$ under the alternative hypothesis $H_1 = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}}$ of (18) 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. First, under $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$, the rank of $\mathbf{\Pi}$ is $r+1$ so Assumption 5b applies; in particular, $\alpha_{yy} \neq 0$. Second, under $H_0^{\pi_{yy}} : \pi_{yy} = 0$, the rank of $\mathbf{\Pi}$ 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, $\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx} \neq \mathbf{0}'$.

Theorem 4.1 *(Consistency of the Wald statistic 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 (18) the Wald statistic W (21) is consistent against $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ in Cases I–V defined in (12)–(16).*

Theorem 4.2 *(Consistency of the Wald statistic 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 (18) and $H_0^{\pi_{yy}} : \pi_{yy} = 0$ of (17) the Wald statistic W (21) is consistent against $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ in Cases I–V defined in (12)–(16).*

Hence, combining Theorems 4.1 and 4.2, the bounds procedure of Section 3 based on the Wald statistic W (21) defines a consistent test of $H_0 = H_0^{\pi_{yy}} \cap H_0^{\pi_{yx.x}}$ of (17) against $H_1 = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}}$ of (18). This result holds irrespective of whether the forcing variables $\{\mathbf{x}_t\}$ are purely $I(0)$, purely $I(1)$ or mutually cointegrated.

We now turn to consider the asymptotic distribution of the Wald statistic (21) under a suitably specified sequence of local alternatives. Recall that under Assumption 5b, $\boldsymbol{\pi}_{y.x}[= (\pi_{yy}, \boldsymbol{\pi}_{yx.x})] = (\alpha_{yy}\beta_{yy}, \alpha_{yy}\boldsymbol{\beta}'_{xy} + (\boldsymbol{\alpha}_{yx} - \boldsymbol{w}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx})$. Consequently, we define the sequence of local alternatives

$$H_{1T} : \boldsymbol{\pi}_{y.xT}[= (\pi_{yyT}, \boldsymbol{\pi}_{yx.xT})] = (T^{-1}\alpha_{yy}\beta_{yy}, T^{-1}\alpha_{yy}\boldsymbol{\beta}'_{xy} + T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}'_{xx}) \tag{26}$$

Hence, under Assumption 3, defining

$$\mathbf{\Pi}_T \equiv \begin{pmatrix} \pi_{yyT} & \boldsymbol{\pi}_{yxT} \\ \mathbf{0} & \mathbf{\Pi}_{xxT} \end{pmatrix}$$

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

$$\mathbf{\Pi}_T - \mathbf{\Pi} = T^{-1}\boldsymbol{\alpha}_y\boldsymbol{\beta}'_y + T^{-1/2}\begin{pmatrix} \boldsymbol{\delta}_{yx} \\ \boldsymbol{\delta}_{xx} \end{pmatrix}\boldsymbol{\beta}' \tag{27}$$

In order to detail the limit distribution of the Wald statistic under the sequence of local alternatives H_{1T} of (26), 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, $\mathbf{J}^*_{k-r+1}(a) = \mathbf{W}_{k-r+1}(a) + \mathbf{ab}'\int_0^a \mathbf{J}^*_{k-r+1}(r)\,dr$ and $d\mathbf{J}^*_{k-r+1}(a) = d\mathbf{W}_{k-r+1}(a) + \mathbf{ab}'\mathbf{J}^*_{k-r+1}(a)\,da$, where $\mathbf{W}_{k-r+1}(a)$ is a $(k-r+1)$-dimensional standard Brownian motion, $\mathbf{a} = [(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\mathbf{\Omega}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{-1/2}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\boldsymbol{\alpha}_y$, $\mathbf{b} = [(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\mathbf{\Omega}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{1/2}[(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{\Gamma}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{-1}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\boldsymbol{\beta}_y$, together with the de-meaned and de-meaned and de-trended counterparts $\tilde{\mathbf{J}}^*_{k-r+1}(a) = (\tilde{J}^*_u(a), \tilde{\mathbf{J}}^*_{k-r}(a)')'$ 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 $H_{1T}: \pi_{y.x} = T^{-1}\alpha_{yy}\boldsymbol{\beta}_y' + T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}'$ of (26), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (21) has the representation*

$$W \Rightarrow \mathbf{z}_r'\mathbf{z}_r + \int_0^1 dJ_u^*(a)\mathbf{F}_{k-r+1}(a)' \left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'\,da\right)^{-1} \int_0^1 \mathbf{F}_{k-r+1}(a)\,dJ_u^*(a) \quad (28)$$

where $\mathbf{z}_r \sim N(\mathbf{Q}^{1/2}\boldsymbol{\eta}, \mathbf{I}_r)$, $\mathbf{Q}[= \mathbf{Q}^{1/2\prime}\mathbf{Q}^{1/2}] = p\lim_{T\to\infty}(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*)$, $\boldsymbol{\eta} \equiv (\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx})'$, is distributed independently of the second term in (28) and

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

$r = 0, \ldots, k$, and Cases I–V are defined in (12)–(16), $a \in [0, 1]$.

The first component of (28) $\mathbf{z}_r'\mathbf{z}_r$ 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 $H_{1T}^{\pi_{yx.x}}$: $\boldsymbol{\pi}_{yx.xT} = T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}_{xx}'$ under $H_0^{\pi_{yy}} : \pi_{yy} = 0$. The second term in (28) 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 $\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx} = \mathbf{0}'$. Note that under H_0 of (17), that is, $\alpha_{yy} = 0$ and $\boldsymbol{\delta}_{yx} - \boldsymbol{w}'\boldsymbol{\delta}_{xx} = \mathbf{0}'$, the limiting representation (28) reduces to (22) as should be expected.

The proof for the consistency of the bounds test procedure based on the t-statistic of (24) requires that the rank of the long-run multiplier matrix $\boldsymbol{\Pi}$ is $r + 1$ under the alternative hypothesis $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$. Hence, Assumption 5b applies; in particular, $\alpha_{yy} \neq 0$.

Theorem 4.4 *(Consistency of the t-statistic 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 (18) the t-statistic $t_{\pi_{yy}}$ (24) is consistent against $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ in Cases I, III and V defined in (12), (14) and (16).*

As noted at the end of Section 3, Theorem 4.4 suggests the possibility of using $t_{\pi_{yy}}$ to discriminate between $H_0^{\pi_{yy}} : \pi_{yy} = 0$ and $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$, although, if $H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ is false, the bounds procedure given *via* Corollaries 3.3 and 3.4 is not asymptotically similar.

AN APPLICATION: UK EARNINGS EQUATION

Following the modelling approach described earlier, this section provides a re-examination of the earnings equation included in the UK Treasury macroeconometric model described in Chan, Savage and Whittaker (1995), CSW hereafter. The theoretical basis of the Treasury's earnings equation is the bargaining model advanced in Nickell and Andrews (1983) and reviewed, for example, in Layard *et al.* (1991, Chapter 2). Its theoretical derivation is based on a Nash bargaining framework where firms and unions set wages to maximize a weighted average of firms' profits and unions'

utility. Following Darby and Wren-Lewis (1993), the theoretical real wage equation underlying the Treasury's earnings equation is given by

$$w_t = \frac{Prod_t}{1 + f(UR_t)(1 - RR_t)/Union_t} \tag{29}$$

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 (29) after allowing for a wedge effect that takes account of the difference between the 'real product wage' which is the focus of the firms' decision, and the 'real consumption wage' which concerns the union.[15] 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 different theoretical rationales for the inclusion of ΔUR_t in the wage equation: the differential moderating effects of long- 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 equilibrium correction formulation used here automatically allow for such lagged effects.

We begin our empirical analysis from the maintained assumption that the time series properties of the key variables in the Treasury's earnings equation can be well approximated by a log-linear $VAR(p)$ model, augmented with appropriate deterministics such as intercepts and 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'.[16] Also, as in CSW, we include two dummy variables to account for the effects of incomes policies on average earnings. These dummy variables are defined by

$$D7475_t = 1, \text{ over the period } 1974q1 - 1975q4, 0 \text{ elsewhere}$$

$$D7579_t = 1, \text{ over the period } 1975q1 - 1979q4, 0 \text{ elsewhere}$$

The asymptotic theory developed in the paper is not affected by the inclusion of such 'one-off' dummy variables.[17] 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 ECM of interest can be written as

$$\Delta w_t = c_0 + c_1 t + c_2 D7475_t + c_3 D7579_t + \pi_{ww} w_{t-1} + \boldsymbol{\pi}_{wx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta\mathbf{z}_{t-i} + \boldsymbol{\delta}'\Delta\mathbf{x}_t + u_t \tag{30}$$

[15] 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.

[16] It is important, however, that, at a future date, a fresh investigation of the possible effects of the replacement ratio on real wages should be undertaken.

[17] However, both the asymptotic theory and associated critical values must be modified if the fraction of periods in which the dummy variables are non-zero does not tend to zero with the sample size T. In the present application, both dummy variables included in the earning equation are zero after 1979, and the fractions of observations where $D7475_t$ and $D7579_t$ are non-zero are only 7.6% and 19.2% respectively.

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 LS.[18] 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 these equations is an identification requirement for the bargaining theory of wages which permits it to be distinguished from other alternatives, such as the efficiency wage theory which postulates that labour productivity is partly determined by the level of real wages.[19] It is clear that, in our framework, the bargaining theory and the efficiency wage theory cannot be entertained simultaneously, at least not in the long run.

The above specification is also based on the assumption that the disturbances u_t are serially uncorrelated. It is therefore important that the lag order p of the underlying *VAR* is selected appropriately. There is a delicate balance between choosing p sufficiently large to mitigate the residual serial correlation problem and, at the same time, sufficiently small so that the conditional ECM (30) is not unduly over-parameterized, particularly in view of the limited time series data which are available.

Finally, a decision must be made concerning the time trend in (30) and whether its coefficient should be restricted.[20] This issue can only be settled in light of the particular sample period under consideration. The time series data used are quarterly, cover the period 1970q1-1997q4, and are seasonally adjusted (where relevant).[21] To ensure comparability of results for different choices of p, all estimations use the same sample period, 1972q1–1997q4 ($T = 104$), with the first eight observations reserved for the construction of lagged variables.

The five variables in the earnings equation were constructed from primary sources in the following manner: $w_t = \ln(ERPR_t/PYNONG_t)$, $Wedge_t = \ln(1 + TE_t) + \ln(1 - TD_t) - \ln(RPIX_t/PYNONG_t)$, $UR_t = \ln(100 \times ILOU_t/(ILOU_t + WFEMP_t))$, $Prod_t = \ln((YPROM_t + 278.29 \times YMF_t)/(EMF_t + ENMF_t))$, and $Union_t = \ln(UDEN_t)$, 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 the 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 (thousands), 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') measured by union membership as a percentage of employment.[22] The time series plots of the five variables included in the VAR model are given in Figures 1–3.

[18] See Assumption 3 and the following discussion. By construction, the contemporaneous effects Δx_t are uncorrelated with the disturbance term u_t and instrumental variable estimation which has been particularly popular in the empirical wage equation literature is not necessary. Indeed, given the unrestricted nature of the lag distribution of the conditional ECM (30), it is difficult to find suitable instruments: namely, variables that are not already included in the model, which are uncorrelated with u_t *and* also have a reasonable degree of correlation with the included variables in (30).

[19] For a discussion of the issues that surround the identification of wage equations, see Manning (1993).

[20] See, for example, PSS and the discussion in Section 2.

[21] 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 CSW, pp. 46–51 and p. 11 of the Annex.

[22] The data series for *UDEN* assumes a constant rate of unionization from 1980q4 onwards.

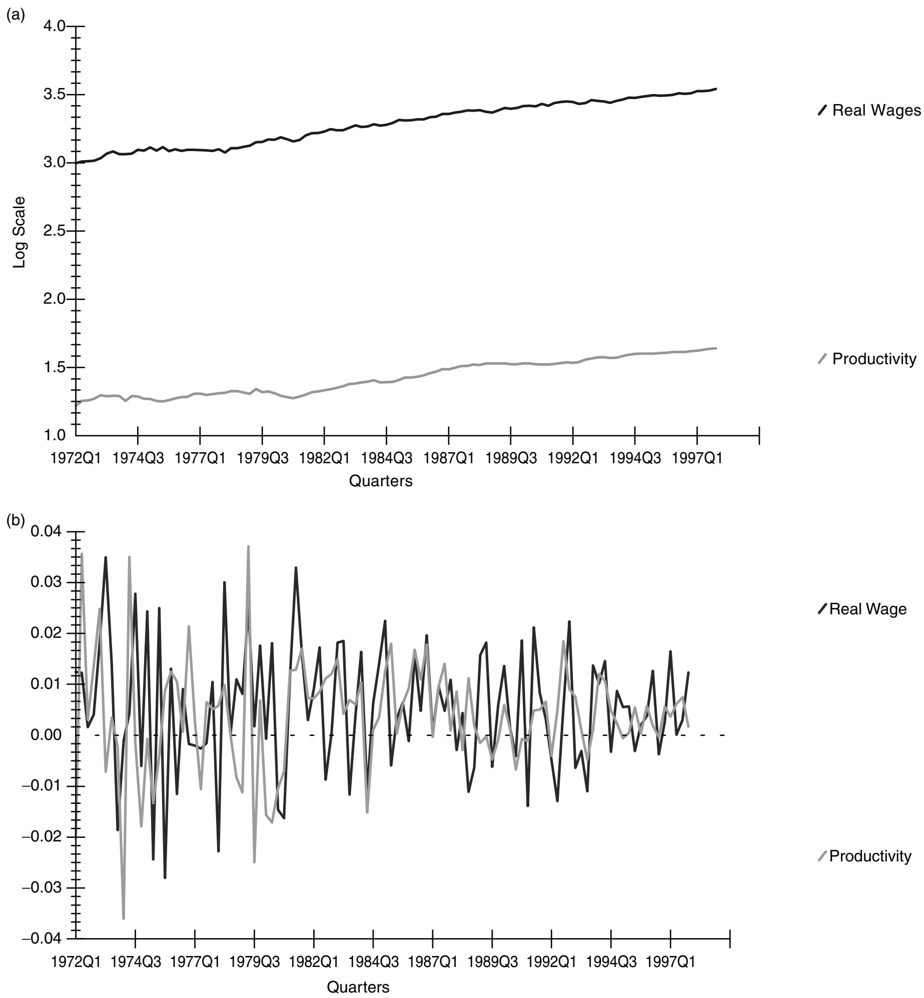

Figure 1. (a) Real wages and labour productivity. (b) Rate of change of real wages and labour productivity

It is clear from Figure 1 that real wages (average earnings) and productivity show steadily rising trends with real wages growing at a faster rate than productivity.[23] This suggests, at least initially, that a linear trend should be included in the real wage equation (30). Also the application of 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 cases of real wages and productivity. This does not necessarily preclude the other three variables (UR, $Wedge$, and $Union$) having levels impact on real wages. Following the methodology developed in this paper, it is possible to test for the existence of a real wage equation involving the levels of these five variables irrespective of whether they are purely $I(0)$, purely $I(1)$, or mutually cointegrated.

[23] Over the period 1972q1–97q4, real wages grew by 2.14% per annum as compared to labour productivity that increased by an annual average rate of 1.54% over the same period.

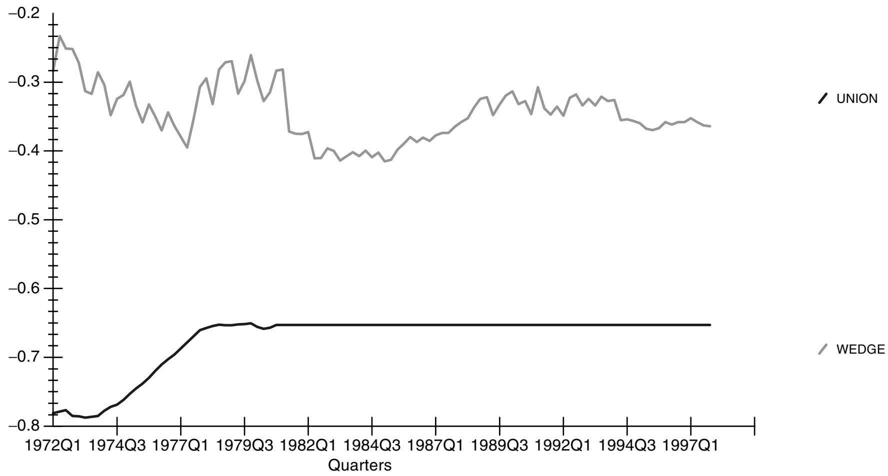

Figure 2. The wedge and the unionization variables

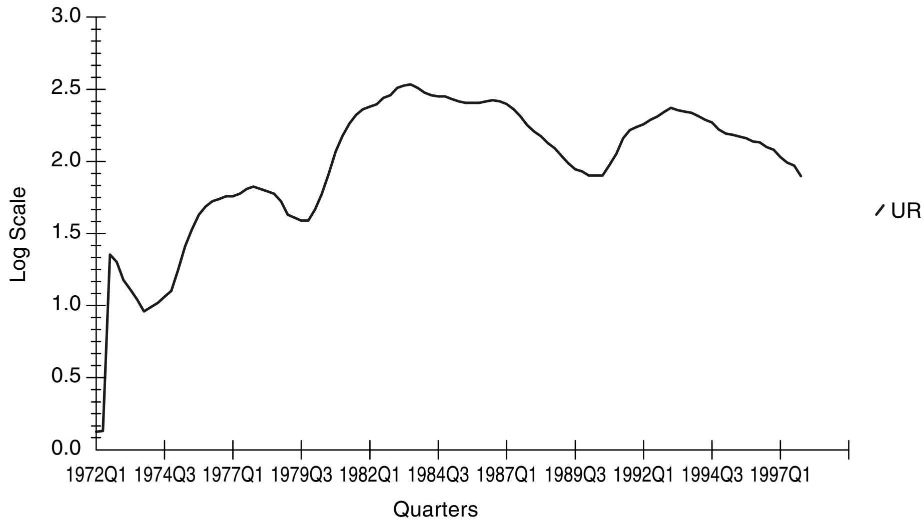

Figure 3. The unemployment rate

To determine the appropriate lag length p and whether a deterministic linear trend is required in addition to the productivity variable, we estimated the conditional model (30) by LS, with and without a linear time trend, for $p = 1, 2, \ldots, 7$. As pointed out earlier, all regressions were computed over the same period 1972q1–1997q4. We found that lagged changes of the productivity variable, $\Delta Prod_{t-1}$, $\Delta Prod_{t-2}, \ldots$, were insignificant (either singly or jointly) in all 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 lagged changes of all other variables. Table I gives Akaike's and Schwarz's Bayesian Information Criteria, denoted

respectively by AIC and SBC, and Lagrange multiplier (LM) statistics for testing the hypothesis of no residual serial correlation against orders 1 and 4 denoted by $\chi^2_{SC}(1)$ and $\chi^2_{SC}(4)$ respectively.

As might be expected, the lag order selected by AIC, $\widehat{p}_{aic} = 6$, irrespective of whether a deterministic trend term is included or not, is much larger than that selected by SBC. This latter criterion gives estimates $\widehat{p}_{sbc} = 1$ if a trend is included and $\widehat{p}_{sbc} = 4$ if 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 tests, it seems prudent to select p to be either 5 or 6.[24] Nevertheless, for completeness, in what follows we report test results for $p = 4$ and 5, as well as for our preferred choice, namely $p = 6$. The results in Table I also indicate that there is little to choose between the conditional ECM with or without a linear deterministic trend.

Table II gives the values of the F- and t-statistics for testing the existence of a level earnings equation under three different scenarios for the deterministics, Cases III, IV and V of (14), (15) and (16) respectively; see Sections 2 and 3 for detailed discussions.

The various statistics in Table II should be compared with the critical value bounds provided in Tables CI and CII. First, consider the bounds F-statistic. As argued in PSS, the statistic F_{IV} which sets the trend coefficient to zero under the null hypothesis of no level relationship, Case IV of (15), is more appropriate than F_V, Case V of (16), which ignores this constraint. Note that, if the trend coefficient c_1 is not subject to this restriction, (30) implies a quadratic trend in the level of real wages under the null hypothesis of $\pi_{ww} = 0$ and $\boldsymbol{\pi}_{wx.x} = \mathbf{0}'$, which is empirically implausible. The critical value bounds for the statistics F_{IV} and F_V are given in Tables CI(iv) and CI(v). Since $k = 4$, the 0.05 critical value bounds are (3.05, 3.97) and (3.47, 4.57) for F_{IV} and F_V, respectively.[25] The test outcome depends on the choice of the lag order p. For $p = 4$, the

Table I. 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 underlying VAR model for the conditional ECM (30), with zero restrictions on the coefficients of lagged changes in the productivity variable. $AIC_p \equiv LL_p - s_p$ and $SBC_p \equiv LL_p - (s_p/2)\ln T$ denote Akaike's and Schwarz's Bayesian Information Criteria for a given lag order p, where LL_p is the maximized log-likelihood value of the model, 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 LM statistics for testing no residual serial correlation against orders 1 and 4. The symbols *, **, and *** denote significance at 0.01, 0.05 and 0.10 levels, respectively.

[24] 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.

[25] Following a suggestion from one of the referees we also computed critical value bounds for our sample size, namely $T = 104$. For $k = 4$, the 5% critical value bounds associated with F_{IV} and F_V statistics turned out to be (3.19,4.16) and (3.61,4.76), respectively, which are only marginally different from the asymptotic critical value bounds.

Table II. F- and t-statistics for testing the existence of a levels 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[c]	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: See the notes to Table I. F_{IV} is the F-statistic for testing $\pi_{ww} = 0$, $\boldsymbol{\pi}_{wx.x} = \mathbf{0}'$ and $c_1 = 0$ in (30). F_V is the F-statistic for testing $\pi_{ww} = 0$ and $\boldsymbol{\pi}_{wx.x} = \mathbf{0}'$ in (30). F_{III} is the F-statistic for testing $\pi_{ww} = 0$ and $\boldsymbol{\pi}_{wx.x} = \mathbf{0}'$ in (30) with c_1 set equal to 0. t_V and t_{III} are the t-ratios for testing $\pi_{ww} = 0$ in (30) with and without a deterministic linear trend. [a] indicates that the statistic lies below the 0.05 lower bound, [b] that it falls within the 0.05 bounds, and [c] that it lies above the 0.05 upper bound.

hypothesis that there exists no level earnings equation is not rejected at the 0.05 level, irrespective of whether the regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated. For $p = 5$, the bounds test is inconclusive. For $p = 6$ (selected by AIC), the statistic F_V is still inconclusive, but $F_{IV} = 4.78$ lies outside the 0.05 critical value bounds and rejects the null hypothesis that there exists no level earnings equation, irrespective of whether the regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated.[26] This finding is even more conclusive when the bounds F-test is applied to the earnings equations without a linear trend. The relevant test statistic is F_{III} and the associated 0.05 critical value bounds are (2.86, 4.01).[27] For $p = 4$, $F_{III} = 3.63$, and the test result is inconclusive. However, for $p = 5$ and 6, the values of F_{III} are 5.23 and 5.42 respectively and the hypothesis of no levels earnings equation is conclusively rejected.

The results from the application of the bounds t-test to the earnings equations are less clear-cut and do not allow the imposition of the trend restrictions discussed above. The 0.05 critical value bounds for t_{III} and t_V, when $k = 4$, are (−2.86, −3.99) and (−3.41, −4.36).[28] Therefore, if a linear trend is included, the bounds t-test does not reject the null even if $p = 5$ or 6. However, when the trend term is excluded, the null is rejected for $p = 5$. Overall, these test results support the existence of a levels earnings equation when a sufficiently high lag order is selected and when the statistically insignificant deterministic trend term is excluded from the conditional ECM (30). Such a specification is in accord with the evidence on the performance of the alternative conditional ECMs set out in Table I.

In testing the null hypothesis that there are no level effects in (30), namely ($\pi_{ww} = 0$, $\boldsymbol{\pi}_{wx.x} = 0$) it is important that the coefficients of lagged changes remain unrestricted, otherwise these tests could be subject to a pre-testing problem. However, for the subsequent estimation of levels effects and short-run dynamics of real wage adjustments, the use of a more parsimonious specification seems advisable. To this end we adopt the ARDL approach to the estimation of the level relations

[26] The same conclusion is also reached for $p = 7$.
[27] See Table CI(iii).
[28] See Tables CII(iii) and CII(v).

discussed in Pesaran and Shin (1999).[29] First, the (estimated) orders of an $ARDL(p, p_1, p_2, p_3, p_4)$ model in the five variables $(w_t, Prod_t, UR_t, Wedge_t, Union_t)$ were selected by searching across the $7^5 = 16,807$ ARDL models, spanned by $p = 0, 1, \ldots, 6$, and $p_i = 0, 1, \ldots, 6$, $i = 1, \ldots, 4$, using the AIC criterion.[30] This resulted in the choice of an $ARDL(6, 0, 5, 4, 5)$ specification with estimates of the levels relationship given by

$$\underset{}{w_t =} \underset{(0.050)}{1.063}\, Prod_t \underset{(0.034)}{-0.105}\, UR_t \underset{(0.265)}{-0.943}\, Wedge_t \underset{(0.311)}{+1.481}\, Union_t \underset{(0.242)}{+2.701} + \hat{v}_t \tag{31}$$

where $\hat{v}_t$ is the equilibrium correction term, and the standard errors are given in parenthesis. All levels estimates are highly significant and have the expected signs. The coefficients of the productivity and the wedge variables are insignificantly different from unity. In the Treasury's earnings equation, the levels coefficient of the productivity variable is imposed as unity and the above estimates can be viewed as providing empirical support for this *a priori* restriction. Our levels estimates of the effects of the unemployment rate and the union variable on real wages, namely -0.105 and 1.481, are also in line with the Treasury estimates of -0.09 and 1.31.[31] The main difference between the two sets of estimates concerns the levels coefficient of the wedge variable. We obtain a much larger estimate, almost twice that obtained by the Treasury. Setting the levels coefficients of the $Prod_t$ and $Wedge_t$ variables to unity provides the alternative interpretation that the share of wages (net of taxes and computed using *RPIX* rather than the implicit GDP deflator) has varied negatively with the rate of unemployment and positively with union strength.[32]

The conditional ECM regression associated with the above level relationship is given in Table III.[33] These estimates provide further direct evidence on the complicated dynamics that seem to exist between real wage movements and their main determinants.[34] All five lagged changes in real wages are statistically significant, further justifying the choice of $p = 6$. The equilibrium correction coefficient is estimated as -0.229 (0.0586) which is reasonably large and highly significant.[35] The auxiliary equation of the autoregressive part of the estimated conditional ECM has real roots 0.9231 and -0.9095 and two pairs of complex roots with moduli 0.7589 and 0.6381, which suggests an initially cyclical real wage process that slowly converges towards the equilibrium described by (31).[36] The regression fits reasonably well and passes the diagnostic tests against non-normal errors and heteroscedasticity. However, it fails the functional form misspecification test at

[29] Note that the ARDL approach advanced in Pesaran and Shin (1999) is applicable irrespective of whether the regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated.
[30] For further details, see Section 18.19 and Lesson 16.5 in Pesaran and Pesaran (1997).
[31] CSW do not report standard errors for the levels estimates of the Treasury earnings equation.
[32] We are grateful to a referee for drawing our attention to this point.
[33] Clearly, it is possible to simplify the model further, but this would go beyond the remit of this section which is first to test for the existence of a level relationship using an unrestricted ARDL specification and, second, if we are satisfied that such a levels relationship exists, to select a parsimonious specification.
[34] The standard errors of the estimates reported in Table III allow for the uncertainty associated with the estimation of the levels coefficients. This is important in the present application where it is not known with certainty whether the regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated. It is only in the case when it is known for certain that all regressors are $I(1)$ that it would be reasonable in large samples to treat these estimates as known because of their super-consistency.
[35] The equilibrium correction coefficient in the Treasury's earnings equation is estimated to be -0.1848 (0.0528), which is smaller than our estimate; see p. 11 in Annex of CSW. This seems to be because of the shorter lag lengths used in the Treasury's specification rather than the shorter time period 1971q1–1994q3. Note also that the t-ratio reported for this coefficient does not have the standard t-distribution; see Theorem 3.2.
[36] The complex roots are $0.34293 \pm 0.67703i$ and $-0.17307 \pm 0.61386i$, where $i = \sqrt{-1}$.

the 0.05 level which may be linked to the presence of some non-linear effects or asymmetries in the adjustment of the real wage process that our linear specification is incapable of taking into account.[37] Recursive estimation of the conditional ECM and the associated cumulative sum and cumulative sum of squares plots also suggest that the regression coefficients are generally stable over the sample period. However, these tests are known to have low power and, thus, may have missed important breaks. Overall, the conditional ECM earnings equation presented in Table III has a number of desirable features and provides a sound basis for further research.

Table III. Equilibrium correction form of the *ARDL*(6, 0, 5, 4, 5) earnings equation

Regressor	Coefficient	Standard error	p-value
$\hat{v}_{t-1}$	−0.229	0.0586	N/A
Δw_{t-1}	−0.418	0.0974	0.000
Δw_{t-2}	−0.328	0.1089	0.004
Δw_{t-3}	−0.523	0.1043	0.000
Δw_{t-4}	−0.133	0.0892	0.140
Δw_{t-5}	−0.197	0.0807	0.017
$\Delta Prod_t$	0.315	0.0954	0.001
ΔUR_t	0.003	0.0083	0.683
ΔUR_{t-1}	0.016	0.0119	0.196
ΔUR_{t-2}	0.003	0.0118	0.797
ΔUR_{t-3}	0.028	0.0113	0.014
ΔUR_{t-4}	0.027	0.0122	0.031
$\Delta Wedge_t$	−0.297	0.0534	0.000
$\Delta Wedge_{t-1}$	−0.048	0.0592	0.417
$\Delta Wedge_{t-2}$	−0.093	0.0569	0.105
$\Delta Wedge_{t-3}$	−0.188	0.0560	0.001
$\Delta Union_t$	−0.969	0.8169	0.239
$\Delta Union_{t-1}$	−2.915	0.8395	0.001
$\Delta Union_{t-2}$	−0.021	0.9023	0.981
$\Delta Union_{t-3}$	−0.101	0.7805	0.897
$\Delta Union_{t-4}$	−1.995	0.7135	0.007
$Intercept$	0.619	0.1554	0.000
$D7475_t$	0.029	0.0063	0.000
$D7579_t$	0.017	0.0063	0.009

$\overline{R}^2 = 0.5589$, $\hat{\sigma} = 0.0083$, $AIC = 339.57$, $SBC = 302.55$,
$\chi^2_{SC}(4) = 8.74[0.068]$, $\chi^2_{FF}(1) = 4.86[0.027]$
$\chi^2_N(2) = 0.01[0.993]$, $\chi^2_H(1) = 0.66[0.415]$.

Notes: The regression is based on the conditional ECM given by (30) using an *ARDL*(6, 0, 5, 4, 5) specification with dependent variable, Δw_t estimated over 1972q1–1997q4, and the equilibrium correction term $\hat{v}_{t-1}$ is given in (31). $\overline{R}^2$ is the adjusted squared multiple correlation coefficient, $\hat{\sigma}$ is the standard error of the regression, AIC and SBC are Akaike's and Schwarz's Bayesian Information Criteria, $\chi^2_{SC}(4)$, $\chi^2_{FF}(1)$, $\chi^2_N(2)$, and $\chi^2_H(1)$ denote chi-squared statistics to test for no residual serial correlation, no functional form mis-specification, normal errors and homoscedasticity respectively with p-values given in $[\cdot]$. For details of these diagnostic tests see Pesaran and Pesaran (1997, Ch. 18).

[37] The conditional ECM regression in Table III also passes the test against residual serial correlation but, as the model was specified to deal with this problem, it should not therefore be given any extra credit!

6. CONCLUSIONS

Empirical analysis of level relationships has been an integral part of time series econometrics and pre-dates the recent literature on unit roots and cointegration.[38] However, the emphasis of this earlier literature was on the estimation of level relationships rather than testing for their presence (or otherwise). Cointegration analysis attempts to fill this vacuum, but, typically, under the relatively restrictive assumption that the regressors, $\mathbf{x}_t$, entering the determination of the dependent variable of interest, y_t, are all integrated of order 1 or more. This paper demonstrates that the problem of testing for the existence of a level relationship between y_t and $\mathbf{x}_t$ is non-standard even if *all* the regressors under consideration are $I(0)$ because, under the null hypothesis of no level relationship between y_t and $\mathbf{x}_t$, the process describing the y_t process is $I(1)$, irrespective of whether the regressors $\mathbf{x}_t$ are purely $I(0)$, purely $I(1)$ or mutually cointegrated. The asymptotic theory developed in this paper provides a simple univariate framework for testing the existence of a single level relationship between y_t and $\mathbf{x}_t$ when it is not known with certainty whether the regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated.[39] Moreover, it is unnecessary that the order of integration of the underlying regressors be ascertained prior to testing the existence of a level relationship between y_t and $\mathbf{x}_t$. Therefore, unlike typical applications of cointegration analysis, this method is not subject to this particular kind of pre-testing problem. The application of the proposed bounds testing procedure to the UK earnings equation highlights this point, where one need not take an *a priori* position as to whether, for example, the rate of unemployment or the union density variable are $I(1)$ or $I(0)$.

The analysis of this paper is based on a single-equation approach. Consequently, it is inappropriate in situations where there may be more than one *level* relationship involving y_t. An extension of this paper and those of HJNR and PSS to deal with such cases is part of our current research, but the consequent theoretical developments will require the computation of further tables of critical values.

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,

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} t + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t \tag{A1}$$

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

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

[38] For an excellent review of this early literature, see Hendry *et al.* (1984).

[39] Of course, 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.

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 $\boldsymbol{\Omega}$ 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 that $\boldsymbol{\beta}_{*'}\mathbf{z}^*_t = \boldsymbol{\beta}'\boldsymbol{\mu} + \boldsymbol{\beta}'\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t = O_P(1)$. Hence, from Phillips and Solo (1992, Theorem 3.16, p. 983), defining $\tilde{\mathbf{Z}}^*_{-1} \equiv \overline{\mathbf{P}}_{\iota}\mathbf{Z}^*_{-1}$ and $\widetilde{\Delta\mathbf{Z}}_{-} \equiv \overline{\mathbf{P}}_{\iota}\Delta\mathbf{Z}_{-}$, it follows that

$$T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* = O_P(1), T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1), T^{-1}\widetilde{\Delta\mathbf{Z}}'_{-}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1)$$
$$T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* = O_P(1), T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1) \tag{A2}$$

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

$$T^{-1/2}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{u}} = O_P(1), T^{-1/2}\widetilde{\Delta\mathbf{Z}}'_{-}\tilde{\mathbf{u}} = O_P(1) \tag{A3}$$

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* $\mathbf{G}(a) = (\mathbf{G}_1(a)', G_2(a))'$, *where* $\mathbf{G}_1(a) \equiv (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{C}\tilde{\mathbf{B}}_{k+1}(a)$, $\tilde{\mathbf{B}}_{k+1}(a)[= (\tilde{B}_1(a)', \tilde{\mathbf{B}}_k(a)')'] = \mathbf{B}_{k+1}(a) - \int_0^1 \mathbf{B}_{k+1}(a)da$, *and* $G_2(a) \equiv a - \frac{1}{2}$, $a \in [0,1]$. *Then*

$$T^{-2}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{Z}}^*_{-1}\mathbf{B}_T \Rightarrow \int_0^1 \mathbf{G}(a)\mathbf{G}(a)'da, \; T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{u}} \Rightarrow \int_0^1 \mathbf{G}(a)d\tilde{B}^*_u(a)$$

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

Proof of Theorem 3.1 *Under* H_0 *of (17), the Wald statistic* W *of (21) can be written as*

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

where $\mathbf{A}_T \equiv T^{-1/2}\left(\boldsymbol{\beta}_*, T^{-1/2}\mathbf{B}_T\right)$. *Consider the matrix* $\mathbf{A}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\tilde{\mathbf{Z}}^*_{-1}\mathbf{A}_T$. *It follows from (A2) and Lemma A.1 that*

$$\mathbf{A}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\tilde{\mathbf{Z}}^*_{-1}\mathbf{A}_T = \begin{pmatrix} T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* & \mathbf{0}' \\ \mathbf{0} & T^{-2}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\tilde{\mathbf{Z}}^*_{-1}\mathbf{B}_T \end{pmatrix} + o_P(1) \tag{A4}$$

Next, consider $\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}}$*. From (A3) and Lemma A.1,*

$$\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}} = \begin{pmatrix} T^{-1/2}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}} \\ T^{-1}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\tilde{\mathbf{u}} \end{pmatrix} + o_P(1) \tag{A5}$$

Finally, the estimator for the error variance ω_{uu} *(defined in the line after (21)),*

$$\begin{aligned}\hat{\omega}_{uu} &= (T-m)^{-1}\left[\tilde{\mathbf{u}}'\tilde{\mathbf{u}} - \tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{A}_T(\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{A}_T)^{-1}\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}}\right] \\ &= (T-m)^{-1}\tilde{\mathbf{u}}'\tilde{\mathbf{u}} + o_P(1) = \omega_{uu} + o_P(1) \end{aligned} \tag{A6}$$

From (A4)–(A6) and Lemma A.1,

$$\begin{aligned} W &= T^{-1}\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}}/\omega_{uu} \\ &\quad + T^{-2}\tilde{\mathbf{u}}'\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\tilde{\mathbf{u}}/\omega_{uu} + o_P(1) \end{aligned} \tag{A7}$$

We consider each of the terms in the representation (A7) in turn. A central limit theorem allows us to state

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

Hence, the first term in (A7) converges in distribution to $\mathbf{z}_r'\mathbf{z}_r$*, a chi-square random variable with r degrees of freedom; that is,*

$$T^{-1}\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta \mathbf{Z}_-}}\tilde{\mathbf{u}}/\omega_{uu} \Rightarrow \mathbf{z}_r'\mathbf{z}_r \sim \chi^2(r) \tag{A8}$$

From Lemma A.1, the second term in (A7) weakly converges to

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

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

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

Now, noting that under H_0 *of (17) we may express* $\boldsymbol{\alpha}_y^{\perp} = (1, -\boldsymbol{w}')'$ *and* $\boldsymbol{\alpha}^{\perp} = (\mathbf{0}, \boldsymbol{\alpha}_{xx}^{\perp}{}')'$ *where* $\boldsymbol{\alpha}_{xx}^{\perp}{}'\boldsymbol{\alpha}_{xx} = \mathbf{0}$*, we define the* $(k-r+1)$*-vector of independent de-meaned standard Brownian motions,*

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

where $\tilde{B}_u^*(a) = \tilde{B}_1(a) - \boldsymbol{w}'\tilde{\mathbf{B}}_k(a)$ is independent of $\tilde{\mathbf{B}}_k(a)$ and $\tilde{\mathbf{B}}_{k+1}(a) \equiv (\tilde{B}_1(a), \tilde{\mathbf{B}}_k(a)')'$ is partitioned according to $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$, $a \in [0, 1]$. Hence, the second term in (A7) has the following asymptotic representation:

$$\int_0^1 d\tilde{W}_u(a)\begin{pmatrix}\tilde{\mathbf{W}}_{k-r+1}(a)\\ a-\frac{1}{2}\end{pmatrix}'\left(\int_0^1\begin{pmatrix}\tilde{\mathbf{W}}_{k-r+1}(a)\\ a-\frac{1}{2}\end{pmatrix}\begin{pmatrix}\tilde{\mathbf{W}}_{k-r+1}(a)\\ a-\frac{1}{2}\end{pmatrix}'da\right)^{-1} \times \int_0^1\begin{pmatrix}\tilde{\mathbf{W}}_{k-r+1}(a)\\ a-\frac{1}{2}\end{pmatrix}d\tilde{W}_u(a) \tag{A9}$$

Note that $d\tilde{W}_u(a)$ in (A9) may be replaced by $dW_u(a)$, $a \in [0, 1]$. Combining (A8) and (A9) 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 $\mathbf{Z}_{-1}^* = (\iota_T, \mathbf{Z}_{-1}')'$, 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,

$$\boldsymbol{\delta} = \begin{pmatrix}-\boldsymbol{\mu}'\\ \mathbf{I}_{k+1}\end{pmatrix}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp}) \text{ and } \mathbf{B}_T = (\boldsymbol{\delta}, \boldsymbol{\xi}).$$

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 emphasize 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 $H_0^{\boldsymbol{\pi}_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$; 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*

$$\widehat{\Delta\mathbf{y}}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\hat{\mathbf{Z}}_{-1}\mathbf{A}_T\left(\mathbf{A}_T'\hat{\mathbf{Z}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{Z}}_{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\hat{\mathbf{Z}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}}/\hat{\omega}_{uu} \tag{A10}$$

where $\mathbf{A}_T \equiv T^{-1/2}(\boldsymbol{\beta}, T^{-1/2}\mathbf{B}_T)$ and $\mathbf{B}_T = (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$. Note that only the diagonal element of the inverse in (A10) corresponding to $\boldsymbol{\beta}_y^{\perp}$ is relevant, which implies that we only need to consider the blocks $T^{-2}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{Z}}_{-1}\mathbf{B}_T$ and $T^{-1}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}}$ in (A10). Therefore, using (A2) and (A3), (A10) is asymptotically equivalent to

$$T^{-1}\hat{\mathbf{u}}'\overline{\mathbf{P}}_{\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{Z}}_{-1}\mathbf{B}_T\left(T^{-2}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\hat{\mathbf{Z}}_{-1}\mathbf{B}_T\right)^{-1}T^{-1}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\overline{\mathbf{P}}_{\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{u}}/\omega_{uu} \tag{A11}$$

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

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

where, for convenience, but without loss of generality, we have set $\boldsymbol{\beta}_y^{\perp} = (\beta_{yy}^{\perp}, \mathbf{0}')'$ *and* $\boldsymbol{\beta}^{\perp} = (\mathbf{0}, \boldsymbol{\beta}_{xx}^{\perp}{}')'$, $\boldsymbol{\lambda}_{xy}^{\phi} \equiv \boldsymbol{\gamma}_{xy}/\gamma_{yy.x}^{\phi}$, $\gamma_{yy.x}^{\phi} \equiv \gamma_{yy} - \boldsymbol{\phi}'\boldsymbol{\gamma}_{xy}$, $\boldsymbol{\gamma}_{yx.x}^{\phi} \equiv \boldsymbol{\gamma}_{yx} - \boldsymbol{\phi}'\boldsymbol{\Gamma}_{xx}$ *and* $\hat{\mathbf{B}}_k^{\phi}(a) \equiv \hat{\mathbf{B}}_k(a) - \boldsymbol{\lambda}_{xy}^{\phi}\hat{B}_u^{\phi}(a)$, $\hat{B}_u^{\phi}(a) \equiv \hat{B}_1(a) - \phi'\hat{\mathbf{B}}_k(a)$, $a \in [0, 1]$. *Hence, (A11) weakly converges to*

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

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

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* $H_1^{\pi_{yy}}: \pi_{yy} \neq 0$, *Assumption 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,* $\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma}t + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$, *where now* $\mathbf{C} \equiv \boldsymbol{\beta}^{\perp}[\boldsymbol{\alpha}^{\perp}{}'\boldsymbol{\Gamma}\boldsymbol{\beta}^{\perp}]^{-1}\boldsymbol{\alpha}^{\perp}{}'$. *We redefine* $\boldsymbol{\beta}_*$ *and* $\boldsymbol{\delta}$ *as the* $(k+2, r+1)$ *and* $(k+2, k-r)$ *matrices,*

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

and

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

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*

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

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 (21) multiplied by $\hat{\omega}_{uu}$ *may be written as*

$$\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^*\mathbf{A}_T\left(\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^*\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}} + 2\boldsymbol{\lambda}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}} + \boldsymbol{\lambda}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^*\boldsymbol{\lambda}_*, \quad (B1)$$

where $\boldsymbol{\lambda}_* \equiv \boldsymbol{\beta}_*(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})'(1, -\boldsymbol{w}')'$, $\mathbf{A}_T \equiv T^{-1/2}(\boldsymbol{\beta}_*, T^{-1/2}\mathbf{B}_T)$ *and* $\mathbf{B}_T \equiv (\boldsymbol{\delta}, T^{-1/2}\boldsymbol{\xi})$*. Note that (A6) 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 (B1) divided by* ω_{uu} *has the limiting representation*

$$\mathbf{z}_{r+1}'\mathbf{z}_{r+1} + \int_0^1 dW_u(a)\mathbf{F}_{k-r}(a)' \left(\int_0^1 \mathbf{F}_{k-r}(a)\mathbf{F}_{k-r}(a)'da\right)^{-1} \int_0^1 \mathbf{F}_{k-r}(a)dW_u(a) \qquad (B2)$$

where $\mathbf{z}_{r+1} \sim N(\mathbf{0}, \mathbf{I}_{r+1})$, $\mathbf{F}_{k-r}(a) = (\tilde{\mathbf{W}}_{k-r}(a)', a - \frac{1}{2})'$ *and* $\tilde{\mathbf{W}}_{k-r}(a) \equiv (\boldsymbol{\alpha}_{xx}^{\perp\prime}\boldsymbol{\Omega}_{xx}\boldsymbol{\alpha}_{xx}^{\perp})^{-1/2}\boldsymbol{\alpha}_{xx}^{\perp\prime}\tilde{\mathbf{B}}_k(a)$ *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. (22). 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 (B2) is unconditionally distributed as a* $\chi^2(k-r)$ *random variable and is independent of the first term; cf. (A4). Hence, the first term in (B1) divided by* ω_{uu} *has a limiting* $\chi^2(k+1)$ *distribution.*

The second term in (B1) may be written as

$$2(1, -\boldsymbol{w}')(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}} = 2T^{1/2}(1, -\boldsymbol{w}')(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})\left(T^{-1/2}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}}\right) = O_P(T^{1/2}), \quad (B3)$$

and the third term as

$$\begin{aligned}&(1, -\boldsymbol{w}')(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})'(1, -\boldsymbol{w}')' \\ &= T(1, -\boldsymbol{w}')(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})'(1, -\boldsymbol{w}')' = O_P(T) \qquad (B4)\end{aligned}$$

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

Proof of Theorem 4.2 *A similar decomposition to (B1) for the Wald statistic (21) holds under* $H_1^{\pi_{yx.x}} \cap H_0^{\pi_{yy}}$ *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* $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$*. Therefore, as in Theorem 3.2, note that we may write* $\boldsymbol{\alpha}_y^{\perp} = (1, -\boldsymbol{\phi}')'$ *for some k-vector* $\boldsymbol{\phi} \neq \boldsymbol{w}$*. Consequently, the first term divided by* ω_{uu} *may be written as*

$$\begin{aligned}&T^{-1}\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*\prime}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}}/\omega_{uu} \\ &\quad + T^{-2}\tilde{\mathbf{u}}'\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*\prime}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*\prime}\tilde{\mathbf{u}}/\omega_{uu} + o_P(1) \qquad (B5)\end{aligned}$$

cf. (A7). As in the Proof of Theorem 3.1, the first term of (B5) has the limiting representation $\mathbf{z}_r'\mathbf{z}_r$ *where* $\mathbf{z}_r \sim N(\mathbf{0}, \mathbf{I}_r)$*; cf. (22). The second term of (B5) has the limiting representation*

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

where $\tilde{B}_u^{\phi}(a) \equiv \tilde{B}_1(a) - \boldsymbol{\phi}'\tilde{\mathbf{B}}_k(a),\ a \in [0,1]$*; cf. Proof of Theorem 3.2. The second term of (B1) becomes*

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

and the third term

$$(1, -\boldsymbol{w}')\boldsymbol{\alpha}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\boldsymbol{\alpha}'(1, -\boldsymbol{w}')' = T(1, -\boldsymbol{w}')\boldsymbol{\alpha}$$
$$\times\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)\boldsymbol{\alpha}'(1, -\boldsymbol{w}')' = O_P(T)$$

The Theorem follows as $(1, -\boldsymbol{w}')\boldsymbol{\alpha} \neq \mathbf{0}'$ *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 (26). Hence,* $\boldsymbol{\Phi}(L)(\mathbf{z}_{tT} - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \boldsymbol{\xi}_{tT}$*, where* $\boldsymbol{\xi}_{tT} \equiv (\boldsymbol{\Pi}_T - \boldsymbol{\Pi})[\mathbf{z}_{(t-1)T} - \boldsymbol{\mu} - \boldsymbol{\gamma}(t-1)] + \boldsymbol{\varepsilon}_t$ *and* $\boldsymbol{\Pi}_T - \boldsymbol{\Pi}$ *is given in (27). Therefore,* $\Delta(\mathbf{z}_{tT} - \mu - \boldsymbol{\gamma}t) = \mathbf{C}\boldsymbol{\xi}_{tT} + \mathbf{C}^*(L)\Delta\boldsymbol{\xi}_{tT}$*,* $\mathbf{C}(z) = \mathbf{C} + (1-z)\mathbf{C}^*(z)$ *and* $\mathbf{C} = (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})[(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\boldsymbol{\Gamma}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})]^{-1}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'$*, and thus,*

$$[\mathbf{I}_{k+1} - (\mathbf{I}_{k+1} + T^{-1}\mathbf{C}\boldsymbol{\alpha}_y\boldsymbol{\beta}_y')L](\mathbf{z}_{tT} - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{C}\boldsymbol{\varepsilon}_{tT} + \mathbf{C}^*(L)\Delta\boldsymbol{\xi}_{tT} \qquad (B6)$$

where

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

Inverting (B6) yields

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

Note that $\Delta\boldsymbol{\xi}_{tT} = (\boldsymbol{\Pi}_T - \boldsymbol{\Pi})\Delta[\mathbf{z}_{(t-1)T} - \boldsymbol{\mu} - \boldsymbol{\gamma}(t-1)] + \Delta\boldsymbol{\varepsilon}_t$*. 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* $\boldsymbol{\Omega}$*,* $a \in [0,1]$*; cf. Johansen (1995, Theorem 14.1, p. 202).*

Similarly to (A4),

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

Therefore, expression (B1) for the Wald statistic (21) multiplied by $\hat{\omega}_{uu}$ *is revised to*

$$\hat{\omega}_{uu}W = T^{-1}\widetilde{\Delta}'\mathbf{y}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\widetilde{\Delta\mathbf{y}}$$
$$+ T^{-2}\widetilde{\Delta}'\mathbf{y}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\tilde{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\tilde{\mathbf{Z}}_{-1}^{*'}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\widetilde{\Delta\mathbf{y}} + o_P(1) \qquad (B7)$$

The first term in (B7) may be written as

$$T^{-1}\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}}$$
$$+2T^{-1}\tilde{\mathbf{u}}'\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT}$$
$$+T^{-1}\boldsymbol{\pi}^*_{yT}\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT} \quad (B8)$$

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

$$T^{-1/2}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT}=T^{-1/2}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}(\boldsymbol{\beta}_{y*}\alpha_{yy}T^{-1}+\boldsymbol{\beta}_*\boldsymbol{\eta}T^{-1/2})$$
$$=T^{-1}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\boldsymbol{\eta}+o_P(1) \quad (B9)$$

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, (B8) divided by* ω_{uu} *may be re-expressed as*

$$\left[\left(T^{-1/2}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}}\right)+\mathbf{Q}\boldsymbol{\eta}\right]'\mathbf{Q}^{-1}\left[\left(T^{-1/2}\boldsymbol{\beta}'_*\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{u}}\right)+\mathbf{Q}\boldsymbol{\eta}\right]/\omega_{uu}+o_P(1)=\mathbf{z}'_r\mathbf{z}_r+o_P(1) \quad (B9)$$

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

As $\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\widetilde{\Delta\mathbf{y}}=\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}(\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT}+\tilde{\mathbf{u}})$, $T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\widetilde{\Delta\mathbf{y}}=T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}(\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT}+\tilde{\mathbf{u}})$. *Consider the second term in (B7), in particular,* $T^{-1}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*'}_{yT}$ *which after substitution for* $\boldsymbol{\pi}^*_{yT}$ *becomes*

$$T^{-2}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_{y*}\alpha_{yy}+T^{-3/2}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\boldsymbol{\eta}=T^{-2}\mathbf{B}'_T\tilde{\mathbf{Z}}^{*'}_{-1}\overline{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}_-}}\tilde{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_{y*}\alpha_{yy}+o_P(1)$$
$$\Rightarrow\int_0^1\begin{pmatrix}(\boldsymbol{\beta}^{\perp}_y,\boldsymbol{\beta}^{\perp})'\mathbf{C}\tilde{\mathbf{J}}_{k+1}(a)\\ a-\frac{1}{2}\end{pmatrix}\tilde{\mathbf{J}}_{k+1}(a)'\mathbf{C}'\boldsymbol{\beta}_y\alpha_{yy}da$$

Therefore,

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

Consider

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

where $\tilde{J}_u(a)=\tilde{J}_1(a)-\boldsymbol{w}'\tilde{\mathbf{J}}_k(a)$ *is independent of* $\tilde{\mathbf{J}}_k(a)$ *and* $\tilde{\mathbf{J}}_{k+1}(a)\equiv(\tilde{J}_1(a),\tilde{\mathbf{J}}_k(a)')'$, $a\in[0,1]$. *Now,* $\tilde{\mathbf{J}}^*_{k-r+1}(a)$ *satisfies the stochastic integral and differential equations,* $\tilde{\mathbf{J}}^*_{k-r+1}(a)=\tilde{\mathbf{W}}_{k-r+1}(a)+\mathbf{ab}'\int_0^a\tilde{\mathbf{J}}^*_{k-r+1}(r)dr$ *and* $d\tilde{\mathbf{J}}^*_{k-r+1}(a)=d\tilde{\mathbf{W}}_{k-r+1}(a)+\mathbf{ab}'\tilde{\mathbf{J}}^*_{k-r+1}(a)da$, *where* $\mathbf{a}=[(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})'\Omega(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})]^{-1/2}(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})'\boldsymbol{\alpha}_y$ *and* $\mathbf{b}=[(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})'\boldsymbol{\Omega}(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})]^{1/2}\times[(\boldsymbol{\beta}^{\perp}_y,\boldsymbol{\beta}^{\perp})'\boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\perp}_y,\boldsymbol{\alpha}^{\perp})]^{-1}(\boldsymbol{\beta}^{\perp}_y,\boldsymbol{\beta}^{\perp})'\boldsymbol{\beta}_y$; *cf. Johansen (1995, Theorem 14.4, p. 207). Note that the first element of* $\tilde{\mathbf{J}}^*_{k-r+1}(a)$ *satisfies* $\tilde{J}^*_u(a)=\tilde{W}_u(a)+\omega^{-1/2}_{uu}\alpha_{yy}\mathbf{b}'\int_0^a\tilde{\mathbf{J}}^*_{k-r+1}(r)dr$ *and* $d\tilde{J}^*_u(a)=d\tilde{W}_u(a)+\omega^{-1/2}_{uu}\alpha_{yy}\mathbf{b}'\tilde{J}^*_{k-r+1}(a)da$.

Therefore,

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

Hence, the second term in (B7) weakly converges to

$$\omega_{uu}\int_0^1 d\tilde{J}_u^*(a)\mathbf{F}_{k-r+1}(a)'\left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'da\right)^{-1}\int_0^1 \mathbf{F}_{k-r+1}(a)\, d\tilde{J}_u^*(a) \qquad (B10)$$

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

Combining (B9) and (B10) gives the result stated in Theorem 4.3 as $\hat{\omega}_{uu} - \omega_{uu} = O_P(1)$ *under* H_{1T} *of (26) and noting* $d\tilde{J}_u^*(a)$ *may be replaced by* $dJ_u^*(a)$.■

Proof of Theorem 4.4 *We consider Case V; the remaining Cases I and III may be dealt with similarly. Under* $H_1^{\pi_{yy}}: \pi_{yy} \neq 0$, *from (10),* $\hat{\mathbf{y}}_{-1} = \hat{\mathbf{X}}_{-1}\boldsymbol{\theta} + \hat{\mathbf{v}}_{-1}$, *where* $\hat{\mathbf{v}}_{-1} \equiv \overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\mathbf{v}_{-1}$ *and* $\mathbf{v}_{-1} = (0, v_1, \ldots, v_{T-1})'$. *Therefore,* $\hat{\mathbf{y}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}} = \hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{Y}}$ *and* $\hat{\mathbf{y}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{y}}_{-1} = \hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{v}}_{-1}$.

As in Appendix A,

$$\begin{aligned} T^{-1/2}\boldsymbol{\beta}_{xx}^{\perp\prime}\mathbf{x}_{[Ta]} &= T^{-1/2}\boldsymbol{\beta}_{xx}^{\perp\prime}\boldsymbol{\mu}_x + T^{-1/2}\boldsymbol{\beta}_{xx}^{\perp\prime}\boldsymbol{\gamma}_x t + T^{-1/2}(\boldsymbol{\beta}_{xx}^{\perp\prime}\boldsymbol{\beta}_{xx}^{\perp})(\boldsymbol{\alpha}^{\perp\prime}\boldsymbol{\Gamma}\boldsymbol{\beta}^{\perp})^{-1}\boldsymbol{\alpha}^{\perp\prime}\mathbf{s}_{[Ta]} \\ &\quad + (\mathbf{0}, \boldsymbol{\beta}_{xx}^{\perp\prime})T^{-1/2}\mathbf{C}^*(L)\boldsymbol{\varepsilon}_{[Ta]} \end{aligned}$$

and noting that $\boldsymbol{\beta}_{xx}'\boldsymbol{\beta}_{xx}^{\perp} = \mathbf{0}$, $\boldsymbol{\beta}_{xx}'\mathbf{x}_t = T^{-1/2}\boldsymbol{\beta}_{xx}'\boldsymbol{\mu}_x + T^{-1/2}\boldsymbol{\beta}_{xx}'\boldsymbol{\gamma}_x t + (\mathbf{0}, \boldsymbol{\beta}_{xx}')\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$. *Consequently,*

$$\mathbf{A}_{xT}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\mathbf{A}_{xT} = \begin{pmatrix} T^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx} & \mathbf{0}' \\ \mathbf{0} & T^{-2}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp} \end{pmatrix} + o_P(1)$$

where $\mathbf{A}_{xT} \equiv T^{-1/2}(\boldsymbol{\beta}_{xx}, T^{-1/2}\boldsymbol{\beta}_{xx}^{\perp})$.

Now, because $T^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\hat{\mathbf{v}}_{-1} = O_P(1)$, $T^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\widehat{\Delta\mathbf{Z}_-} = O_P(1)$, $T^{-1}\widehat{\Delta\mathbf{Z}_-}'\widehat{\Delta\mathbf{Z}_-} = O_P(1)$ *and* $T^{-1}\widehat{\Delta\mathbf{Z}_-}'\hat{\mathbf{v}}_{-1} = O_P(1)$, *hence* $T^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{v}}_{-1} = O_P(1)$. *Also because* $T^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\hat{\mathbf{v}}_{-1} = O_P(1)$ *and* $T^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\widehat{\Delta\mathbf{Z}_-} = O_P(1)$, *hence* $T^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{v}}_{-1} = O_P(1)$; *cf. (A3). Hence, noting that* $T^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx} = O_P(1)$ *and* $T^{-2}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp} = O_P(1)$,

$$\begin{aligned} T^{-1}\hat{\mathbf{y}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{y}}_{-1} &= T^{-1}\hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{v}}_{-1} - T^{-1}\hat{\mathbf{v}}_{-1}'\mathbf{P}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{v}}_{-1} + o_P(1) \\ &= T^{-1}\hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{v}}_{-1} + o_P(1) \end{aligned}$$

where $\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}} \equiv \overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}} - \overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}(\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx})^{-1}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}$ *and* $\mathbf{P}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}} \equiv \overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}(\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp})^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-}}$. *Therefore, as* $T^{-1}\hat{\mathbf{v}}_{-1}'\hat{\mathbf{v}}_{-1} = O_P(1)$,

$$T^{-1}\hat{\mathbf{y}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{y}}_{-1} = O_P(1) \qquad (B11)$$

The numerator of $t_{\pi_{yy}}$ *of (24) may be written as* $\hat{\mathbf{y}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}} = \hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{u}} + \hat{\mathbf{v}}_{-1}'\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}_-},\hat{\mathbf{X}}_{-1}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda}$, *where* $\boldsymbol{\lambda} \equiv (\boldsymbol{\beta}_y, \boldsymbol{\beta})(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})'(1, -\mathbf{w}')'$. *Because* $T^{-1/2}\boldsymbol{\beta}_{xx}'\hat{\mathbf{X}}_{-1}'\hat{\mathbf{u}} = O_P(1)$ *and* $T^{-1/2}\widehat{\Delta\mathbf{Z}_-}'\hat{\mathbf{u}} =$

$O_P(1)$, $T^{-1/2}\boldsymbol{\beta}'_{xx}\hat{\mathbf{X}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{u}} = O_P(1)$, and, as $T^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}'_{-1}\hat{\mathbf{u}} = O_P(1)$, $T^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{u}} = O_P(1)$. Therefore,

$$T^{-1/2}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\hat{\mathbf{u}} = T^{-1/2}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{u}} - T^{-1/2}\hat{\mathbf{v}}'_{-1}\mathbf{P}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{u}} + o_P(1)$$
$$= T^{-1/2}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{u}} + o_P(1) = O_P(1)$$

noting $T^{-1/2}\hat{\mathbf{v}}'_{-1}\hat{\mathbf{u}} = O_P(1)$. Similarly, as $(1, -\boldsymbol{w}')(\boldsymbol{\alpha}_y, \boldsymbol{\alpha}) \neq \mathbf{0}'$, $T^{-1}\boldsymbol{\lambda}'\hat{\mathbf{Z}}'_{-1}\widehat{\Delta\mathbf{Z}}_- = O_P(1)$, $T^{-1}\boldsymbol{\lambda}'\hat{\mathbf{Z}}'_{-1}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx} = O_P(1)$ and $T^{-1}\boldsymbol{\lambda}'\hat{\mathbf{Z}}'_{-1}\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp} = O_P(1)$. Therefore,

$$T^{-1}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} = T^{-1}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} - T^{-1}\hat{\mathbf{v}}'_{-1}\mathbf{P}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} + o_P(1)$$
$$= T^{-1}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} + o_P(1) = O_P(1)$$

noting $T^{-1}\hat{\mathbf{v}}'_{-1}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} = O_P(1)$. Thus,

$$T^{-1/2}\hat{\mathbf{v}}'_{-1}\overline{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\hat{\mathbf{Z}}_{-1}\boldsymbol{\lambda} = O_P(T^{1/2}). \tag{B12}$$

Because $\hat{\omega}_{uu} - \omega_{uu} = o_P(1)$, combining (B11) and (B12) yields the desired result.■

ACKNOWLEDGEMENTS

We are grateful to the Editor (David Hendry) and three anonymous referees for their helpful comments on an earlier version of this paper. Our thanks are also owed to Michael Binder, Peter Burridge, Clive Granger, Brian Henry, Joon-Yong Park, Ron Smith, Rod Whittaker and seminar participants at the University of Birmingham. Partial financial support from the ESRC (grant Nos R000233608 and R000237334) and the Isaac Newton Trust of Trinity College, Cambridge, is gratefully acknowledged. Previous versions of this paper appeared as DAE Working Paper Series, Nos. 9622 and 9907, University of Cambridge.

REFERENCES

Banerjee A, Dolado J, Galbraith JW, Hendry DF. 1993. *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press: Oxford.

Banerjee A, Dolado J, Mestre R. 1998. Error-correction mechanism tests for cointegration in single-equation framework. *Journal of Time Series Analysis* **19**: 267–283.

Banerjee A, Galbraith JW, Hendry DF, Smith GW. 1986. Exploring equilibrium relationships in econometrics through static models: some Monte Carlo Evidence. *Oxford Bulletin of Economics and Statistics* **48**: 253–277.

Blanchard OJ, Summers L. 1986. Hysteresis and the European Unemployment Problem. In *NBER Macroeconomics Annual* 15–78.

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

Boswijk HP. 1994. Testing for an unstable root in conditional and structural error correction models. *Journal of Econometrics* **63**: 37–70.

Boswijk HP. 1995. Efficient inference on cointegration parameters in structural error correction models. *Journal of Econometrics* **69**: 133–158.

Cavanagh CL, Elliott G, Stock JH. 1995. Inference in models with nearly integrated regressors. *Econometric Theory* **11**: 1131–1147.

Chan A, Savage D, Whittaker R. 1995. The new treasury model. Government Economic Series Working Paper No. 128, (Treasury Working Paper No. 70).

Darby J, Wren-Lewis S. 1993. Is there a cointegrating vector for UK wages? *Journal of Economic Studies* **20**: 87–115.

Dickey DA, Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**: 427–431.

Dickey DA, Fuller WA. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**: 1057–1072.

Engle RF, Granger CWJ. 1987. Cointegration and error correction representation: estimation and testing. *Econometrica* **55**: 251–276.

Granger CWJ, Lin J-L. 1995. Causality in the long run. *Econometric Theory* **11**: 530–536.

Hansen BE. 1995. Rethinking the univariate approach to unit root testing: using covariates to increase power. *Econometric Theory* **11**: 1148–1171.

Harbo I, Johansen S, Nielsen B, Rahbek A. 1998. Asymptotic inference on cointegrating rank in partial systems. *Journal of Business Economics and Statistics* **16**: 388–399.

Hendry DF, Pagan AR, Sargan JD. 1984. Dynamic specification. In *Handbook of Econometrics* (Vol. II) Griliches Z, Intriligator MD (des). Elsevier: Amsterdam.

Johansen S. 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica* **59**: 1551–1580.

Johansen S. 1992. Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics* **52**: 389–402.

Johansen S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press: Oxford.

Kremers JJM, Ericsson NR, Dolado JJ. 1992. The power of cointegration tests. *Oxford Bulletin of Economics and Statistics* **54**: 325–348.

Layard R, Nickell S, Jackman R. 1991. *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford University Press: Oxford.

Lindbeck A, Snower D. 1989. *The Insider Outsider Theory of Employment and Unemployment*, MIT Press: Cambridge, MA.

Manning A. 1993. Wage bargaining and the Phillips curve: the identification and specification of aggregate wage equations. *Economic Journal* **103**: 98–118.

Nickell S, Andrews M. 1983. Real wages and employment in Britain. *Oxford Economic Papers* **35**: 183–206.

Nielsen B, Rahbek A. 1998. *Similarity issues in cointegration analysis*. Preprint No. 7, Department of Theoretical Statistics, University of Copenhagen.

Park JY. 1990. Testing for unit roots by variable addition. In *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, Fomby TB, Rhodes RF (eds). JAI Press: Greenwich, CT.

Pesaran MH, Pesaran B. 1997. *Working with Microfit 4.0: Interactive Econometric Analysis*, Oxford University Press: Oxford.

Pesaran MH, Shin Y. 1999. An autoregressive distributed lag modelling approach to cointegration analysis. Chapter 11 in *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Strom S (ed.). Cambridge University Press: Cambridge.

Pesaran MH, Shin Y, Smith RJ. 2000. Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics* **97**: 293–343.

Phillips AW. 1958. The relationship between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica* **25**: 283–299.

Phillips PCB, Durlauf S. 1986. Multiple time series with integrated variables. *Review of Economic Studies* **53**: 473–496.

Phillips PCB, Ouliaris S. 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica* **58**: 165–193.

Phillips PCB, Solo V. 1992. Asymptotics for linear processes. *Annals of Statistics* **20**: 971–1001.

Rahbek A, Mosconi R. 1999. Cointegration rank inference with stationary regressors in VAR models. *The Econometrics Journal* **2**: 76–91.

Sargan JD. 1964. Real wages and prices in the U.K. *Econometric Analysis of National Economic Planning*, Hart PE Mills G, Whittaker JK (eds). Macmillan: New York. Reprinted in Hendry DF, Wallis KF (eds.) *Econometrics and Quantitative Economics*. Basil Blackwell: Oxford; 275–314.

Shin Y. 1994. A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* **10**: 91–115.

Stock J, Watson MW. 1988. Testing for common trends. *Journal of the American Statistical Association* **83**: 1097–1107.

Urbain JP. 1992. On weak exogeneity in error correction models. *Oxford Bulletin of Economics and Statistics* **52**: 187–202.

