ASYMPTOTIC THEORY OF STATISTICS
FROM UNIT ROOT TEST REGRESSIONS
WHEN THE ALTERNATIVE IS A
BREAKING-TREND-STATIONARY MODEL

Antonio Noriega-Muro[*]

Universidad de Guanajuato

Resumen: Se derivan modelos de regresión cuya estructura provee una conexión entre pruebas para una raíz unitaria y pruebas sobre la presencia de parámetros asociados con la función determinística de la tendencia lineal de la regresión. Estas regresiones de prueba son equivalentes a las propuestas por Perron (1989). Con estas ecuaciones de regresión, extendemos los resultados asintóticos de Perron al derivar distribuciones límite para los estimadores de los componentes determinísticos de estas regresiones. Las representaciones asintóticas de estas distribuciones muestran que no hay conflicto entre pruebas para raíces unitarias y pruebas de cambio estructural, modeladas por variables *dummy*.

Abstract: We derive test regressions whose structure provides a link between tests for a unit root and tests on the nullity of the parameters associated with the regression's trend function. These test regressions turn out to be equivalent to those proposed by Perron (1989). Using these regression equations, we extend Perron's (1989) asymptotic results by deriving limiting distributions of the deterministic components for all the models considered. The asymptotic representations of these distributions show that there is no conflict between testing for unit roots and for structural breaks: acceptance of a unit root rules out acceptance of a structural break, as modelled by a dummy variable.

1. Introduction

The seminal work of Nelson and Plosser (1982) drew the attention of both economists and econometricians to the problem associated with the practice of detrending macroeconomic time series by removing a linear trend prior to the analysis of short-run or 'business cycle' fluctuations. In particular they pointed out the potential existence of stochastic non-stationarity due to the presence of unit roots.

Using the statistical techniques developed in Fuller (1976) and Dickey and Fuller (1979), they concluded that the majority of US macroeconomic time series can be well characterized as integrated processes, i.e., processes with a unit root. These findings have serious implications in many fields of economics. In terms of economic policy, unit roots are associated with the concept of 'persistency' of innovations or 'shocks' to the economic system. Policy effects can vary dramatically depending on the nature of the non-stationarity associated with macroeconomic variables (see for example Dickey *et al.*, 1986). In terms of econometric modelling, the theory of co-integration is heavily dependent on the existence of unit roots. This theory not only has interesting connections with the concept of equilibrium in economics, but also helps us in understanding the way trending economic variables move together in the long-run (see *inter alia* Engle and Granger, 1987; Dolado and Jenkinson, 1987). The theory of co-integration has also produced interesting results in the area of forecasting (Granger, 1986; Engle and Yoo, 1987).

All these interactions between economic theory, economic policy, econometric modelling and statistical theory have led to a vast amount of research on the nature, existence and implications of unit roots in macroeconomic time series. Yet, there is still much debate as to whether the unit root hypothesis is supported empirically.

Empirical evidence for the unit root hypothesis is usually obtained by employing the 'standard' theory of unit root testing (Dickey and Fuller, 1979, 1981; Said and Dickey, 1984; Phillips, 1987; Phillips and Perron, 1988; among others). This assumes, on the basis of the traditional method of removing (deterministic) trends from the data, that the relevant alternative to the unit root hypothesis is that of stationary fluctuations around a linear time trend. This latter model is usually parameterized by an intercept and a linear trend, together with a stationary autoregressive component.

This 'standard' theory has undergone a reappraisal during the last lustrum, prompted by the recognition of the potential danger of misspecifying the alternative hypothesis against which the unit root null was tested. It has been argued that if the alternative hypothesis is *'underparameterized'*, then a unit root will hardly ever be rejected, even when the 'true' model is stochastically stationary.

Indeed, the work of Rappoport and Reichlin (1987) and Perron (1989) (Perron henceforth), opened an important avenue for research by pointing out the crucial role of the alternative hypothesis in unit root tests. The reappraisal led to the recognition that 'fixed' linear time trends are extremely simple parameterizations representing secular or long-run movements in economic time series. It was then proposed to broaden the alternative hypothesis by recognizing that most macro-economic time series have undergone structural breaks, due to economic crises, changes in institutional arrangements, wars, etc. Taking into account these considerations, test regressions for unit root tests were then modified by including a set of dummy variables to reflect the changes in the trend function due to the shocks mentioned above.

These 'new' tests, formally analyzed by Perron, consist of the (normalized) Ordinary Least Squares estimator of the autoregressive parameter and its t-statistic (for testing the equality of this parameter to unity) in an AR(1) model which includes, apart from an intercept and a trend, a set of dummy variables with which the break in the trend function is modelled. For these test statistics, asymptotic distributions were derived and, by simulating these distributions, critical values computed, under the null hypothesis that the Data Generating Process (DGP) is a driftless random walk.

In this paper, we focus mainly on the models and assumptions of Perron who, as mentioned above, provided the asymptotic theory of the tests for unit roots when structural breaks are taken into consideration.[1]

The paper is structured as follows. In Section 2, we offer a different interpretation of the hypotheses and models analyzed by Perron. Using the approach of Nelson and Plosser (1982)[2] we derive test regressions whose structure provides a link between tests of a unit root and tests on

---

[1] Perron's approach encompasses that of Rappoport and Reichlin (1987).

[2] Other authors using the same approach include, Bhargava (1986), Schmidt and Phillips (1989), and Haldrup (1990).

the nullity of the parameters associated with the regression's deter-
ministic components, i.e. the trend function. These test regressions turn
out to be equivalent to those proposed by Perron for testing a unit root.
This approach, however, allows us not only to rationalize the null and
alternative hypotheses used by Perron, but also to define, upon impos-
ing the restriction implied by the null hypothesis, the appropriate form
of the DGP.

   In Section 3, using these regression equations, we extend Perron's
asymptotic results by deriving limiting distributions of the deterministic
components for all the models considered. The asymptotic repre-
sentations of these distributions show interesting relationships between
the deterministic components and the autoregressive parameter. In par-
ticular, we show that there is no conflict between testing for unit roots
and for structural breaks: acceptance of a unit root rules out acceptance
of a structural break, as modelled by a dummy variable. As a by-product of
the derivation of the asymptotic distributions, rates of convergence are
also found not only for the AR parameter, but also for the parameters of
the trend function. It can be seen that the DGP used in the asymptotic
results is different from the one used by Perron.

2. Models and Procedures

The Trend-Stationary (TS) model (originally studied by Nelson and
Plosser (1982) as the relevant non-stationary deterministic alternative to
the stochastic, non-stationary, unit root process) is by no means the only
deterministic specification of a non-stationary process. Although the
trending nature of many macroeconomic time series has suggested the TS
model as the obvious alternative to the unit root process, it is also
apparent from visual inspection of the data that most of the series display
heterogeneous behaviour in their historical trends. This behaviour has
been described as "*variable trends*" by Stock and Watson (1988), "*break-
ing trends*" by Perron, and "*segmented trends*" by Rappoport and Reichlin
(1987, 1989). Whatever it is called, it implies the presence of *structural
breaks* in the data, giving rise to a more general class of TS processes, the
class of Breaking-Trend-Stationary (BTS) processes.

   It has been found that misspecification of the alternative hypothesis
can lead to the (spurious) acceptance of stochastic non-stationarity (i.e.

a unit root) when in fact the DGP is that of a (non-stationary) deterministically heterogeneous nature, as the BTS model briefly introduced above.

In this Section we apply the approach followed by Nelson and Plosser (1982) for testing the null hypothesis of a unit root. Here, however, we allow for the presence of structural breaks under the alternative hypothesis, using the models and assumptions in Perron.

Following Perron's practice, we also consider three types of breaks in the trend function: a 'step' change in the trend (or 'crash' model) [model (A)]; a change in the slope of the trend (or 'breaking-trend' model) [model (B)]; and a combination of models (A) and (B) [model (C)]. A fundamental assumption in Perron's analysis is that the dating of the break points (denoted by T_B) in any of the three models (A), (B) or (C) is known *a priori*, in the sense that the dates are chosen after viewing the data.

Since a very similar analysis can be applied to the three models (A)-(C), let us concentrate on model (A), leaving the study of models (B) and (C) to the next subsection.

2.1. Model (A)

Assume a situation in which interest centers on testing whether or not *deviations from a broken trend* are stationary. Consider a stochastic process $\{y_t\}_0^\infty$ generated according to the following DGP:

$$(y_t - \mu - \beta t - \theta^A DU_t) = \left(\frac{1}{1 - \alpha L}\right)\varepsilon_t \tag{1}$$

with $DU_t = 1$ if $t > T_B$ and 0 otherwise, where T_B denotes the date of the break. From (1), it is clear that if $\alpha = 1$, deviations from a trend function with a crash are non-stationary, being accumulations of persistent shocks; while if $\alpha < 1$, these deviations will be stationary since the effects of the shocks will tend to disappear as time passes.

We can write (1), upon multiplying both sides by $(1 - \alpha L)$, as:

$$y_t = \mu_*^A + \beta_*^A t + \theta^A DU_t + \theta_*^A DU_{t-1} + \alpha^A y_{t-1} + \varepsilon_t \tag{2}$$

where $\mu_*^A = [\mu(1 - \alpha) + \alpha\beta]$, $\beta_*^A = \beta(1 - \alpha)$, $\theta_*^A = -\alpha\theta^A$, and we have renamed α as α^A in (2) to avoid any confusion with models (B) and (C) introduced below.

However, as discussed in Noriega-Muro (1993a), the interaction of DU_t and DU_{t-1} in (2) create a situation of asymptotic collinearity. We can apply a simple reparameterization to Model (A) by adding and subtracting the term $\alpha\theta DU_t$ to (2), to obtain:

$$y_t = \mu_*^A + \beta_*^A t + \varphi^A DU_t + \varphi_*^A D(TB)_t + \alpha^A y_{t-1} + \varepsilon_t \qquad (2')$$

where now the parameter vector

$$\underline{\beta}^A = (\mu_*^A, \beta_*^A, \varphi^A, \varphi_*^A, \alpha^A)$$

$$\equiv ([\mu(1-\alpha) + \alpha\beta], \beta(1-\alpha), \theta^A(1-\alpha), \alpha\theta^A, \alpha)$$

has null values $(\beta, 0, 0, \theta_A, 1)$, and

$$DU_t - DU_{t-1} = D(TB)_t = 1$$

if $t = T_B + 1$ and 0 otherwise. As we can see, the only difference in model (2') relates to the dummy variables and their respective parameters. Note that, *in the empirical applications* for testing a unit root in Model (A), (2') has the same form as the regression model proposed by Perron (albeit without the extra lags of the first differences of the data) [equation (12) in his paper].

However, *in deriving limiting distributions*, the regression equation used by Perron is equation (2') *without the dummy variable* $D(TB)_t$ [his equation (A.6)], and the DGP assumed is a driftless random walk. This is because he only derives limiting distributions for the (bias of the) OLS estimator of the autoregressive parameter and its corresponding t-statistic. He argues that these statistics, computed from his regression (A.6), are invariant with respect to the parameters in the DGP.

Under the null hypothesis, $H_0 : \alpha = 1$ becomes:[3]

$$y_t = \beta + \theta^A D(TB)_t + y_{t-1} + \varepsilon_t , \qquad (3)$$

---

[3] Note that repeated substitution in (3) takes us back to (1) with $\alpha = 1$, i.e. $y_t = y_0 + \beta t$ $+ \theta^A DU_t + \sum_{i=1}^{t} \varepsilon_i$, with $y_0 = \mu$ since $[D(TB)_t + D(TB)_{t-1} + \ldots + D(TB)_1] = DU_t$.

which is precisely Perron's model (A) (the 'crash' model) under the null hypothesis of a unit root. Now, if we consider (2') under the assumption that $\alpha = 0$ we get:

$$y_t = \mu + \beta t + \theta^A DU_t + \varepsilon_t , \qquad (4)$$

which has the same form as Perron's model (A) under the alternative hypothesis. Note that in (4) there is no autoregressive component. Following Perron, we can construct a regression equation for testing the null of a unit root by nesting the corresponding models under the null (3) and under the (specific) alternative (4). This strategy produces precisely Perron's regression equation for his model (A) [equation (2')].

By using (2') as the regression equation to test $H_0 : \alpha = 1$, the parameters related to the deterministic variables have a clear and straightforward interpretation. That is, under the null hypothesis, $H_0 : \alpha = 1$ and this implies that $\mu^A_* = \beta$, $\beta^A_* = 0$, $\varphi^A = 0$ and $\varphi^A_* = \theta^A$, and we get equation (3), implying that deviations from a breaking trend are non-stationary, while under the alternative, (1) implies that these deviations are stationary.

Following the lines of Haldrup (1990) it should be possible to prove for model (2') that, as $T \to \infty$, $\hat{\mu}^A_* \xrightarrow{P} \beta$, $\hat{\beta}^A_* \xrightarrow{P} 0$, $\hat{\alpha}^A \xrightarrow{P} 1$, $\varphi^A \xrightarrow{P} 0$ and $\varphi^A_* \xrightarrow{P} \theta^A$, where a hat denotes OLSEs of the parameters in (2'). Asymptotic results related to this model and the models presented below, are the subject matter of the next Section.

Therefore, as stressed above, model (2') offers a different interpretation to Perron's regression equation for Model (A), which, on the one hand, helps to rationalize the forms of the null and alternative hypotheses used by Perron, and, on the other, provides a clearer interpretation of the deterministic components and their relationship with the autoregressive parameter.

2.2. Models (B) and (C)

Following the same arguments of model (A), consider again testing whether or not *deviations from a broken trend* are stationary, but now let us analyze the case of a shift not in the trend's level but in its slope, that is, a change in the rate of growth of the series [Model (B)]. Consider then the next equation:

$$(y_t - \mu - \beta t - \theta^B DT_t^*) = \left(\frac{1}{1 - \alpha L}\right)\varepsilon_t \qquad (5)$$

with $DT_t^* = t - T_B$ if $t > T_B$ and 0 otherwise. Again, $\alpha = 1$ implies that deviations from a broken trend are non-stationary, being accumulations of persistent shocks. Multiplying both sides of (5) by $(1 - \alpha L)$ gives:

$$y_t = \mu_*^B + \beta_*^B t + \theta^B DT_t^* + \theta_*^B DT_{t-1}^* + \alpha^B y_{t-1} + \varepsilon_t , \qquad (6)$$

where $\mu_*^B = [\mu(1 - \alpha) + \alpha\beta]$; $\beta_*^B = \beta(1 - \alpha)$, $\theta_*^B = -\alpha\theta^B$, and we have renamed α as α^B. As with model (A) above, under $H_0 : \alpha = 1$, we get Perron's model (B) under the null, while if we let $\alpha = 0$ we get Perron's model (B) under the alternative.

Adding and subtracting $\alpha\theta^B DT_t^*$ to (6) produces the following regression equation for model (B):

$$y_t = \mu_*^B + \beta_*^B t + \varphi^B DT_t^* + \varphi_*^B DU_t + \alpha^B y_{t-1} + \varepsilon_t \qquad (6')$$

where the parameter vector

$$\underline{\beta}_*^B = (\mu_*^B, \beta_*^B, \varphi^B, \varphi_*^B, \alpha^B)$$

$$\equiv ([\mu(1 - \alpha) + \alpha\beta], \beta(1 - \alpha), \theta^B(1 - \alpha), \alpha\theta^B, \alpha),$$

has null values $(\beta, 0, 0, \theta^B, 1)$.

Finally model (C) results from combining models (A) and (B), i.e.:[4]

$$(y_t - \mu - \beta t - \theta^A DU_t - \theta^B DT_t^*) = \left(\frac{1}{1 - \alpha L}\right)\varepsilon_t , \qquad (7)$$

with $\alpha = 1$ implying non-stationary deviations from a trend with both a shift in level and in slope. We can rewrite (7) as:

---

[4] Note that in presenting Model (C) as a combination of Models (A) and (B) we are using the dummy variable DT_t^*, instead of the one used by Perron, namely DT_t ($DT_t = t$ if $t > T_B$ and 0 otherwise). This does not alter the results, but will greatly simplify the algebra of the asymptotic theory in the next section. The same practice was also followed by Zivot and Andrews (1992).

$$y_t = \mu_*^C + \beta_*^C t + \theta^A DU_t + \theta_*^A DU_{t-1} + \theta^B DT_t^* + \theta_*^B DT_{t-1}^* + \alpha^C y_{t-1} + \varepsilon_t \quad (8)$$

where $\mu_*^C = [\mu(1 - \alpha) + \alpha\beta]$; $\beta_*^C = \beta(1 - \alpha)$; $\theta_*^A = -\alpha\theta^A$, $\theta_*^B = -\alpha\theta^B$, and $\alpha^C = \alpha$. However, note that in equation (8), there is perfect collinearity between the dummy variables; specifically

$$DT_t^* - DT_{t-1}^* = DU_t \quad (9)$$

Using (9) into (8) results in the following alternative regression equation

$$y_t = \mu_*^C + \beta_*^C t + (\theta^A + \alpha\theta^B)DU_t + (-\alpha\theta^A)DU_{t-1}$$

$$+ \theta^B(1 - \alpha)DT_t^* + \alpha^C y_{t-1} + \varepsilon_t$$

$$y_t = \mu_*^C + \beta_*^C t + \theta^C DU_t + \theta_*^C DU_{t-1} + \theta_{**}^C DT_t^* + \alpha^C y_{t-1} + \varepsilon_t , \quad (10)$$

where the notation is obvious. We can reparameterize (10) by adding and subtracting $\alpha\theta^A DU_t$ to get the following equation for testing a unit root

$$y_t = \mu_*^C + \beta_*^C t + \varphi^C DU_t + \varphi_*^C D(TB)_t + \varphi_{**}^C DT_t^* + \alpha^C y_{t-1} + \varepsilon_t \quad (10')$$

where the parameter vector

$$\underline{\beta}^C = (\mu_*^C, \beta_*^C, \varphi^C, \varphi_*^C, \varphi_{**}^C, \alpha^C)$$

$$\equiv ([\mu(1 - \alpha) + \alpha\beta], \beta(1 - \alpha), [\theta^A(1 - \alpha) + \alpha\theta^B], \alpha\theta^A, \theta^B(1 - \alpha), \alpha)$$

has null values $(\beta, 0, \theta^B, \theta^A, 0, 1)$. Note that, in the empirical applications for testing a unit root in Model (C), (10') has the same form as the regression model proposed by Perron[5] (albeit without the extra lags of the first differences of the data) [equation (14) in his paper].

Under the null $H_0 : \alpha = 1$, (10') reduces to

[5] But see footnote 4.

$$\Delta y_t = \beta + (\theta^A + \theta^B)DU_t - \theta^A DU_{t-1} + \varepsilon_t$$

$$= \beta + \theta^A D(TB) + \theta^B DU_t + \varepsilon_t \,,$$

which is Perron's null model (C). Setting $\alpha = 0$, (10') produces Perron's alternative hypothesis for model (C):

$$y_t = \mu + \beta t + \theta^A DU_t + \theta^B DT_i^* + \varepsilon_t \,.$$

As in the case of model (A), it should be possible to prove that, under H_0,

$$(\hat{\mu}_*^B, \hat{\beta}_*^B, \hat{\varphi}^B, \hat{\varphi}_*^B, \hat{\alpha}^B) \xrightarrow{P} (\beta, 0, 0, \theta^B, 1) \quad \text{for model (B)}$$

$$(\hat{\mu}_*^C, \hat{\beta}_*^C, \hat{\varphi}^C, \hat{\varphi}_*^C, \hat{\varphi}_{**}^C, \hat{\alpha}^C) \xrightarrow{P} (\beta, 0, 0^B, \theta^A, 0, 1) \quad \text{for (C)}$$

We will proceed to the asymptotic analysis of the parameters for each model in Section 3.

To summarize, we have offered a different interpretation to Perron's regression models for testing a unit root. This interpretation will prove to be very useful in the derivation of the asymptotic results presented in the next Section.

3. Asymptotic Behaviour of Statistics from Unit Root test Regressions when the Alternative is a Breaking-Trend-Stationary Model

Apart from checking the consistency of these OLS estimators, the purpose of this Section is to present the asymptotic theory related to the three models studied in Section 2, i.e., the derivation of limiting distributions, the study of the invariance of these distributions to the parameters in the DGP, and the equivalence between asymptotic distributions.

These results extend Perron's findings in two directions. First, by using a different interpretation to the problem of testing whether or not deviations from a broken trend are stationary (introduced in Section 2), we propose the use of a broader DGP for each model considered. This in turn allows us not to restrict the analysis solely to the autoregressive parameter and its t-statistic, (for which Perron showed invariance to the

parameters in the DGP) but also to extend it to the entire set of parameters in each model. This makes it possible to analyze the relationships between the asymptotic behaviour of the estimators. The small sample counterparts of these large sample results, are investigated in Noriega-Muro (1993c).

The innovations driving the models (denoted ε_t) are assumed to satisfy assumption 1 in Perron, reproduced below for convenience:

ASSUMPTION 1: *(a)* $E(\varepsilon_t) = 0 \supset t$; *(b)* $\sup_t E|\varepsilon_t|^{\beta + \xi} < \infty$ *for some* $\beta > 2$ *and* $\xi > 0$; *(c)* $\sigma^2 = \lim_{T \to \infty} T^{-1}E(S_T^2)$ *exists and* $\sigma^2 > 0$, *where* $S_T = \Sigma_1^T \varepsilon_t$; *(d)* $\{\varepsilon_t\}_1^\infty$ *is strong mixing with mixing numbers* α_m *that satisfy:* $\Sigma_1^\infty \alpha_m^{1 - 2/\beta} < \infty$.

In the Appendix at the end of paper, we express the OLS estimators of the various parameters as functions of normalized partial sums of the innovations ε_t, i.e., objects like $T^{-1/2}S_T$, $T^{-3/2}\sum_{t=1}^{T} S_t$, etc. The process $T^{-1/2}S_T$ is assumed to satisfy an invariance principle, i.e., $T^{-1/2}S_T$ converges weakly to $\sigma W(1)$ [denoted $T^{-1/2}S_T \overset{D}{\Rightarrow} \sigma W(1)$], where W is a limit process known as Wiener process or Brownian motion. Some other related weak convergence results are presented as Lemmas 1 and 2 in the Appendix. In the theorems bellow, the asymptotic distributions of the OLS estimators are, therefore, expressed as functions of Wiener processes.

The representation of these distributions is lengthy, so we make notational economies by writing the various stochastic processes without the argument. Furthermore, following Phillips and Perron (1988), (some) integrals are understood to be taken over the interval [0, 1], and with respect to Lebesgue measure, unless otherwise indicated. Thus, we shall use, for instance, W, $\int W$, and $\int rW$ in place of $W(r)\int_0^1 W(r)dr$, and $\int_0^1 rW(r)dr$.

The algebraic computations of all the OLS estimators and t-statistics to be analyzed in what follows are tedious and were carried out using the computerized algebra package *REDUCE* (see for example, Rayna, 1987 or Hearn, 1987). The proofs for the following theorems are presented in the Appendix at the end of this paper. The relevant *REDUCE* programmes are contained in Noriega-Muro (1992).

3.1. Model (A): "Crash" Model

For convenience, we present here a rewritten version of the DGP for Model (A) [see equation (1)]:

$$y_t = \mu + \beta t + \theta^A DU_t + \sum_{j=1}^{t} \varepsilon_j. \tag{3'}$$

We now present the main result of this Section.

THEOREM 1: *Let* $\{y_t\}_0^T$ *be a sample of size* $T + 1$ *generated from equation (3') (Model (A) under the null hypothesis) and* $\lambda = T_B/T$. *Let also Assumption 1 hold. Then, for the regression model (2'), as* $T \to \infty$[6]

a) $T(\hat{\alpha}^A - 1) \quad \overset{D}{\Rightarrow} \quad D_A^{-1}A_A$

a') $t_{\hat{\alpha}_*^A} \quad \overset{D}{\Rightarrow} \quad (\sigma/\sigma_\varepsilon)(\Delta_A D_A)^{-1/2}A_A$

b) $T\hat{\beta}_*^A \quad \overset{D}{\Rightarrow} \quad -\beta D_A^{-1}A_A$

b') $t_{\hat{\beta}_*^A} \quad \overset{D}{\Rightarrow} \quad -(\sigma/\sigma_\varepsilon)(\Delta_A D_A)^{-1/2}A_A$

c) $T^{1/2}(\hat{\mu}_*^A - \beta) \quad \overset{D}{\Rightarrow} \quad \sigma G D_A^{-1}$

c') $T^{-1/2}t_{\hat{\mu}_*^A} \quad \overset{D}{\Rightarrow} \quad (\beta/\sigma_\varepsilon)(D_A/\Gamma_\mu)^{1/2}$

d) $T^{1/2}\hat{\varphi}^A \quad \overset{D}{\Rightarrow} \quad \sigma H D_A^{-1}$

d') $t_{\hat{\varphi}^A} \quad \overset{D}{\Rightarrow} \quad (\sigma/\sigma_\varepsilon)(\Delta_\varphi D_A)^{-1/2}H$

e) $(\hat{\varphi}_*^A - \theta^A) \quad \overset{D}{\Rightarrow} \quad \varepsilon_{T_B+1}$

e') $t_{\hat{\varphi}_*^A} \quad \overset{D}{\Rightarrow} \quad (\theta^A + \varepsilon_{T_B+1})/\sigma_\varepsilon$

where '^' denotes an OLS estimator, and

---

[6] In what follows, the sign '*' as a subscript to the OLS estimator of α indicates that it is being expressed as a deviation from its true value 1.

$$D_A = \{(4 - 3\lambda^3)\,(\textstyle\int W\,)^2 + \textstyle\int W[2(3\lambda - 4)\textstyle\int_0^\lambda W + 12(\lambda^2 - 1)\textstyle\int rW|$$

$$+ \lambda^{-1}(\textstyle\int_0^\lambda W\,)^2 + 12(1 - \lambda)\textstyle\int rW\,[\textstyle\int_0^\lambda W + \textstyle\int rW\,]$$

$$+ (3\lambda^2 - 3\lambda + 1)(\lambda - 1)\textstyle\int W^2\,\}.$$

$$A_A = \{[(-e\lambda^3 + 6\lambda^2 - 2)\textstyle\int W + (2 - 3\lambda)\textstyle\int_0^\lambda W + 6(\lambda - 1)^2\textstyle\int rW\,]W(1)$$

$$+ [(3\lambda - 4)\textstyle\int W + \lambda^{-1}\textstyle\int_0^\lambda W + 6(1 - \lambda)\textstyle\int rW\,]W(\lambda)$$

$$+ 6\textstyle\int W\,[(1 - \lambda^2)\textstyle\int W + (\lambda - 1)(\textstyle\int_0^\lambda W + 2\textstyle\int rW\,)]$$

$$+ (1/2)(3\lambda^2 - 3\lambda + 1)(\lambda - 1)\,[W(1)^2 - \sigma_\varepsilon^2/\sigma^2]\}.$$

$$G = \{[3\lambda(1 - \lambda^2)\textstyle\int W + 6\lambda(\lambda - 1)\textstyle\int rW + (3\lambda - 4 + \lambda^{-1})\textstyle\int_0^\lambda W\,]$$

$$\times (1/2)(W(1)^2 - \sigma_\varepsilon^2/\sigma^2) + W(1)[6\lambda\textstyle\int W(\textstyle\int rW - \textstyle\int W\,) + 2\lambda^{-1}\textstyle\int_0^\lambda W$$

$$\times ((3\lambda^2 + 1)\textstyle\int W - \textstyle\int_0^\lambda W - 3(2\lambda^2 - 2\lambda + 1)\textstyle\int rW\,) + 3\lambda(\lambda - 1)^2\textstyle\int W^2]$$

$$+ \lambda^{-1}[4\textstyle\int W\,(\textstyle\int W - \textstyle\int_0^\lambda W\,) + 6\textstyle\int rW((\lambda^2 - 2)\textstyle\int W + \textstyle\int_0^\lambda W + 2(1 - \lambda)\textstyle\int rW)$$

$$- (3\lambda^2 - 4\lambda + 1)\textstyle\int W^2]W(\lambda) + 6\textstyle\int W\,[\lambda\textstyle\int W\,(\textstyle\int W - \textstyle\int_0^\lambda W\,)$$

$$+ \lambda(\lambda - 1)\textstyle\int W^2 - \lambda^{-1}\textstyle\int_0^\lambda W(2(\lambda - 1)\textstyle\int rW - \textstyle\int_0^\lambda W + \textstyle\int W)\}.$$

$$\Gamma_\mu = \{\lambda^{-1}[4\textstyle\int W\,(\textstyle\int W - 2\textstyle\int_0^\lambda W\,) + 4(\textstyle\int_0^\lambda W\,)^2 + 12\textstyle\int rW\,((\lambda^2 - 1)$$

$$\times (\textstyle\int W - \textstyle\int_0^\lambda W\,) + (1 - \lambda)\textstyle\int_{\lambda} rW\,) + (3\lambda^4 - 6\lambda^2 + 4\lambda - 1)\textstyle\int W^2]\,\}.$$

$$H = \{[(2 - 3\lambda\textstyle\int W^2 - (6\lambda\textstyle\int \dot{W} - 12\textstyle\int rW\,)\,(\textstyle\int rW - \textstyle\int W\,)$$

$$+ 2\lambda^{-1}\textstyle\int_0^\lambda W(3\textstyle\int rW - \textstyle\int W\,)]W(1) + \lambda^{-1}[12\textstyle\int rW(\textstyle\int W - \textstyle\int rW\,)$$

$$- 4(\textstyle\int W\,)^2 + \textstyle\int W^2]W(\lambda) + 6\textstyle\int W[\textstyle\int W(2\textstyle\int rW - \lambda\textstyle\int W\,)$$

$$+ (\lambda - 1)\int W^2 - \lambda^{-1}\int_0^\lambda W(2\int rW - \int W)\} - (1/2)[6(\lambda - 1)\int rW$$

$$+ (4 - 3\lambda)\int W - \int_0^\lambda W](W(1)^2 - \sigma_\varepsilon^2/\sigma^2)\},$$

$$\sigma_\varepsilon^2 = \lim_{T \to \infty} E(T^{-1}\sum_{t=1}^T \varepsilon_t^2), \quad \Delta_A = (3\lambda^2 - 3\lambda + 1)(\lambda - 1),$$

$$\Delta_\varphi = \lambda^{-1}[4(\int W)^2 + 12\int rW(\int rW - \int W) - \int W^2].$$

PROOF. See Appendix.

Simple algebra shows that parts $a)$ and $a')$ of Theorem 1 are a restatement of the results derived by Perron. As we can see from $a)$ and $a')$, both the normalized bias $T(\hat{\alpha}_*^A)$ and the t-statistic $t_{\hat{\alpha}_*^A}$ are not affected by either the drift β, or the parameter θ^A, i.e., to the presence of the dummy variable DU_t in the DGP.

Parts $a)$ and $b)$ show the super-consistency property of the OLS estimators for α and β.

Parts $a')$-$b')$ of Theorem 1 show us that the t-statistics on $\hat{\alpha}_*^A$ and $\hat{\beta}_*^A$ do not diverge and, moreover, one is the negative of the other. This result implies that, asymptotically, whenever we, say, accept (reject) a unit root, we will accept (reject) $\beta_*^A \equiv \beta(1 - \alpha) = 0$. Hence, this asymptotic result rules out the possibility of the simultaneous acceptance of deterministic and stochastic trends.

Parts $c)$ and $d)$ of the theorem show the consistency of the estimators of the constant term and the shift-in-mean dummy variable, respectively, at the usual rate of $T^{-1/2}$. Part $c')$ shows that the t-statistic will diverge asymptotically. Thus, as $T \to \infty$, the null hypothesis $H_0 : \mu_*^A = 0$, will always be rejected when the true value of μ is β, rather than zero. Therefore, valid asymptotic inference can be performed about $\hat{\mu}_*^A$.

Part $e)$ shows that the parameter related to the dummy variable $D(TB)_t$ is inconsistent, and that the limiting behaviour of the bias $(\hat{\varphi}_*^A - \theta^A)$ is, in contrast to the rest of the estimators, standard; that is, it is not a functional of Brownian motion processes. Instead, it is simply the (zero mean) innovation at the time of the break, ε_{T_B+1}. Result $e)$ is not exclusive to the unit root asymptotics, it also arises in a standard

asymptotic framework. In order to isolate result $e)$ of Theorem 1 from the unit root asymptotics, consider the following regression

$$y_t = \gamma x_t + \theta DU_t + \vartheta D(TB)_t + \varepsilon_t \tag{11}$$

Under standard moments' assumptions, Noriega-Muro (1993a) found that

$$(\hat{\vartheta} - \vartheta) \overset{D}{\Rightarrow} \varepsilon_{T_B + 1}, \text{ as } T \to \infty \tag{12}$$

The inconsistency arises because the dummy variable $D(TB)_t$ is an asymptotically 'ill-behaved' variable, in the sense that, in contrast to any other 'well-behaved' variable, say z_t, for which $\lim_{T \to \infty} (T^{-1} \sum_{t=1}^{T} z_t^2)$ is finite and different from zero, for $D(TB)_t$, $\lim_{T \to \infty} (T^{-\delta} \sum_{t=1}^{T} D(TB)_t^2) = 0$, for $\delta > 0$.

The other dummy variables which we have used so far are, in the sense defined above, asymptotically well behaved. For instance, $\lim_{T \to \infty} (T^{-1} \sum_{t=1}^{T} DU_t^2) = (1 - \lambda)$, which is nonzero and finite since we assumed $\lambda = T_B / T$ to be a constant $\in (0, 1)$. Standard (root T) asymptotics apply to $T^{1/2}(\hat{\theta} - \theta)$ and $T^{1/2}(\hat{\gamma} - \gamma)$.

Notice that the result in (12) is nothing but the asymptotic counterpart of Salkever's (1976) result on the use of dummy variables to compute predictions. For model (11), we have that

$$\hat{\vartheta} = y_{T_B+1} - \hat{\gamma} x_{T_B+1} - \hat{\theta} DU_{T_B+1} \equiv y_{T_B+1} - \hat{\gamma} x_{T_B+1} - \hat{\theta} = \hat{\varepsilon}_{T_B+1} \tag{13}$$

By using (11) for $t = T_B+1$ into (13), we get

$$\hat{\vartheta} = \gamma x_{T_B+1} + \theta + \vartheta + \varepsilon_{T_B+1} - \hat{\gamma} x_{T_B+1} - \hat{\theta}$$

$$(\hat{\vartheta} - \vartheta) = \varepsilon_{T_B+1} - [(\hat{\theta} - \theta) + (\hat{\gamma} - \gamma) x_{T_B+1}]$$

The result in (12) follows immediately since $\hat{\theta}$ and $\hat{\gamma}$ are, as argued above, consistent estimators of θ and γ, respectively.

3.2. Model (B): "Broken Trend" Model

From Section 2.2, the DGP for model (B) can be expressed as:

$$y_t = \mu + \beta t + \theta^B DT_t^* + \sum_{j=1}^{t} \varepsilon_j \tag{5'}$$

For the same reasons as for model (A), *in deriving limiting distributions*, the regression equation used by Perron is equation (6') without the dummy variable DU_t [his equation (A.7)], and the DGP assumed is a driftless random walk. For this model we have the following asymptotic results.

THEOREM 2: *Let* $\{y_t\}_0^T$ *be a sample of size* $T + 1$ *generated from equation (5') [Model (B) under the null hypothesis], and* $\lambda = T_B / T$. *Let also Assumption 1 hold. Then, for the regression model (6'), as* $T \to \infty$

a) $T(\hat{\alpha}^B - 1) \quad \overset{D}{\Rightarrow} \quad D_B^{-1} A_B$

a') $t_{\hat{\alpha}_*^B} \quad \overset{D}{\Rightarrow} \quad (\sigma / \sigma_\varepsilon)(\Delta_B D_B)^{-1/2} A_B$

b) $T\hat{\beta}_*^B \quad \overset{D}{\Rightarrow} \quad -\beta D_B^{-1} A_B$

b') $t_{\hat{\beta}_*^B} \quad \overset{D}{\Rightarrow} \quad -(\sigma / \sigma_\varepsilon)[\Delta_B D_B]^{-1/2} A_B$

c) $T^{1/2}(\hat{\mu}_*^B - \beta) \quad \overset{D}{\Rightarrow} \quad \sigma J_B D_B^{-1}$

c') $T^{-1/2} t_{\hat{\mu}_*^B} \quad \overset{D}{\Rightarrow} \quad (\beta / \sigma_\varepsilon)[\lambda(\lambda - 1)D_B / W]^{1/2}$

d) $T\hat{\varphi}^B \quad \overset{D}{\Rightarrow} \quad -\theta^B A_B D_B^{-1}$

d') $t_{\hat{\varphi}^B} \quad \overset{D}{\Rightarrow} \quad -(\sigma / \sigma_\varepsilon)[\Delta_B D_B]^{-1/2} A_B$

e) $(\hat{\varphi}_*^B - \theta^B) \quad \overset{D}{\Rightarrow} \quad 0$

e') $T^{-1/2} t_{\hat{\varphi}_*^B} \quad \overset{D}{\Rightarrow} \quad (\theta^B / \sigma_\varepsilon)[\lambda(\lambda - 1)D_B / Z]^{1/2}$

where '^' denotes an OLS estimator, and

$$D_B = \{\lambda^3(\lambda-1)^3\int W^2 + 4(\lambda^2+\lambda+1)\lambda^3\int W\,(\int W - 2\int_0^\lambda W) + 12\lambda^3(\lambda+1)$$

$$\times[\int W\int_0^\lambda rW - \int rW\,(\int W - \int_0^\lambda W)] + 4\lambda^2(4\lambda^2-2\lambda+1)(\int_0^\lambda W)^2$$

$$+ 12\lambda(3\lambda-4\lambda^2-1)\int_0^\lambda W\int_0^\lambda rW + 12\lambda^3[\int rW(\int rW - 2\int_0^\lambda rW)]$$

$$+ 12(3\lambda^2-3\lambda+1)(\int_0^\lambda rW)^2\,\}.$$

$$A_B = \{2\lambda^3(\lambda-1)[(2\lambda+1)(\int W - \int_0^\lambda W) - 3(\int rW - \int_0^\lambda rW)]W(1)$$

$$+ 2(\lambda-1)\,[\lambda(8\lambda^2-8\lambda+3)\int_0^\lambda W - \lambda^3(2\lambda+1)\int W + 3\lambda^3\int rW$$

$$- 3(4\lambda^2-5\lambda+2)\int_0^\lambda rW]W(\lambda) + 6\lambda^3(\lambda+1)\int W(\int W - 2\int_0^\lambda W)$$

$$- 12\lambda^3[\int W(\int rW - \int_0^\lambda rW) - \int_0^\lambda W\int rW] + 6\lambda(4\lambda^2-3\lambda+1)(\int_0^\lambda W)^2$$

$$+ 6\int_0^\lambda W[\lambda(4\lambda^2-3\lambda+1)\int_0^\lambda W - 2(3\lambda^2-3\lambda+1)\int_0^\lambda rW\,]$$

$$+ (1/2)\lambda^3(\lambda-1)^3(W(1)^2 - \sigma_\varepsilon^2/\sigma^2)\,\}.$$

$$J_B = -[\{4\lambda(2\lambda^2-\lambda-1)[2\lambda\int_0^\lambda W(\int W - \int_0^\lambda W) - 3\int W\int_0^\lambda rW] + 12\lambda$$

$$\times(\lambda-1)[\int_0^\lambda rW((4\lambda+1)\int_0^\lambda W + 3\int rW - 3\int_0^\lambda rW) - 4\int_0^\lambda W\int rW\,]\,\}W(1)$$

$$+\{\ 8\lambda(2\lambda^3-\lambda^2-\lambda)\int W(\int_0^\lambda W - \int W) + 24\lambda\int W(\int W - 2\int_0^\lambda W)$$

$$+ 12\lambda(2\lambda^2-\lambda)\int W(2\int rW - \int_0^\lambda rW) + 12\lambda\int W(5\int_0^\lambda rW - 6\int rW)$$

$$+ 24\lambda\int_0^\lambda W(\int_0^\lambda W - (\lambda^2-3)\int rW) + 12\lambda^{-1}\int_0^\lambda rW((3\lambda-1)\int_0^\lambda rW$$

$$- (2\lambda^2+3\lambda-1)\int_0^\lambda W) + 12\lambda\int rW((5\lambda-9)\int_0^\lambda rW - 2(2\lambda-3)\int rW)$$

$$+ 2\lambda(3-2\lambda)(\lambda-1)^3\int W^2\ \}\ W(\lambda) + \{4\lambda\int W[6\int_0^\lambda rW(\int rW - \int_0^\lambda rW)$$

$$- \int W(2\int_0^\lambda W + 3(\lambda+1)\int_0^\lambda rW)] + 24\lambda\int_0^\lambda W[\int W(2\int_0^\lambda W + (\lambda+3)\int rW$$

$$+ 2\lambda\int_0^\lambda rW) - \int_0^\lambda W(1 + (\lambda+3)\int rW - (1/2\lambda^2)(3\lambda-1)\int_0^\lambda rW)$$

$$-3\int rW(\int rW - \int_0^\lambda rW) + (1/4)(\lambda - 1)^3\int W^2] \} + 2(\lambda - 1)^3\lambda(2\lambda\int_0^\lambda W$$

$$-3\int_0^\lambda rW(W(1)^2 - \sigma_\varepsilon^2/\sigma^2)].$$

$$W = 4\lambda^3(\lambda - 1) \{4(\lambda^2 + \lambda - 1)[\int W(\int W - 2\int_0^\lambda W) + (\int_0^\lambda W)^2] - 12(\lambda + 1)$$

$$\times (\int W - \int_0^\lambda W)(\int rW - \int_0^\lambda rW) + 12\int rW(\int rW - 2\int_0^\lambda rW) + 3\lambda^{-3}(3\lambda^3 + 3\lambda^2$$

$$- 3\lambda + 1)(\int_0^\lambda rW)^2 + (\lambda - 1)^3\int rW^2\},$$

$$Z = 4 \{\lambda^3(\lambda^3 - 4)(\int W)^2 + 6\lambda^2(\lambda - 2)\int W(\int_0^\lambda rW - \lambda\int_0^\lambda W)$$

$$- 6\lambda^3(\lambda^2 - 2)\int W\int rW - 3\lambda\int_0^\lambda W(\lambda\int_0^\lambda W - 2\int_0^\lambda rW)$$

$$+ 3\lambda^3(3\lambda - 4)(\int rW)^2 + 6\lambda^2(2\lambda - 3)\int rW(\lambda\int_0^\lambda W - \int_0^\lambda rW)$$

$$- 3(\int_0^\lambda rW)^2 - \lambda^3(\lambda - 1)^3\int W^2\}.$$

$$\sigma_\varepsilon^2 = \lim_{T \to \infty} E(T^{-1}\sum_{t=1}^T \varepsilon_t^2), \quad \Delta_B = [\lambda(\lambda - 1)]^3.$$

PROOF. See Appendix.

As in Model (A), parts $a)$ and $a')$ of Theorem 2 show the invariance of the normalized bias $T(\hat{\alpha}^B - 1)$ and the t-statistic on $\hat{\alpha}_*^B$ with respect to the value of both the drift β, and the 'breaking trend' parameter, θ^B. Parts $a)$ and $b)$ show the super-consistency properties of the OLS estimators for α^B and β^B.

Some other results from Theorem 2 mimic those of Theorem 1, namely, the asymptotic equality between $t_{\hat{\alpha}_*^B}$ and $-t_{\hat{\beta}_*^B}$, the consistency of the constant term at the usual rate of $\sqrt{T}$ and the divergent nature of its t-statistic, which, once again, will allow valid asymptotic inference on $\hat{\mu}_*^B$.

Part $d)$ shows that super-consistency is also a feature of $\hat{\varphi}^B$, which will differ from the asymptotic distribution of $T\hat{\alpha}_*^B$ only by (minus) the value of θ^B. $d')$ is an interesting result: the t-statistic of $\hat{\varphi}^B$ is asymptotically equivalent to the negative of the t-statistic on the autoregressive

parameter. This implies that, whenever we get a large negative value for $t_{\hat{\alpha}_*^B}$ (and therefore we reject the null $H_0 : \alpha = 1$), we will get a large positive value for $t_{\hat{\varphi}^B}$ of approximately the same numerical value, so that we will reject the hypothesis that $\varphi^B = 0$. Hence, there is no conflict between these two hypotheses asymptotically: whenever we, say, reject a unit root, we will not be able to reject the presence of a (deterministic) shift in trend.

Finally, part *e)* shows us that $\hat{\varphi}_*^B$ converges to its true value of θ^B, while its *t*-statistic will diverge, implying that, as T grows, it will always reject $H_0 : \dot{\varphi}_* = 0$. This is a nice result, since the (true) null value of φ_*^B is θ^B, not zero.

3.3. Model (C): "Crash + Breaking Trend" Model

In this sub-section we present the asymptotic results for model (C) which, as it was established in Section 2.2, considers whether or not deviations from a trend with both a jump and a change in slope are stationary. For convenience, we rewrite below the DGP for this model as follows:

$$y_t = \mu + \beta t + \theta^A DU_t + \theta^B DT_t^* + \sum_{j=1}^{t} \varepsilon_j \qquad (7')$$

For the same reasons as for model (A), *in deriving limiting distributions*, the regression equation used by Perron is equation (10') without the dummy variable $D(TB)_t$ [his equation (A.8)].

For this model we have the following asymptotic results.

THEOREM 3. *Let* $\{y_t\}_0^T$ *be a sample of size* $T + 1$ *generated from equation (7')* [Model (C) under the null hypothesis] *and* $\lambda = T_B / T$. *Let Assumption 1 hold. Then, for regression model (10')1, as* $T \rightarrow \infty$

a) $T(\hat{\alpha}^C - 1) \quad \overset{D}{\Rightarrow} \quad D_B^{-1} A_B$

a') $t_{\hat{\alpha}_*^C} \quad \overset{D}{\Rightarrow} \quad (\sigma / \sigma_\varepsilon) [\Delta_B D_B]^{-1/2} A_B$

b) $T\hat{\beta}_*^C \quad \overset{D}{\Rightarrow} \quad -\beta D_B^{-1} A_B$

b') $t_{\hat{\beta}_*^C} \quad \overset{D}{\Rightarrow} \quad -(\sigma / \sigma_\varepsilon) [\Delta_B D_B]^{-1/2} A_B$

$c)\ T^{1/2}(\hat{\mu}_*^C - \beta) \quad \overset{\text{D}}{\Rightarrow} \quad \sigma J_B D_B^{-1}$

$c')\ T^{-1/2} t_{\hat{\mu}_*^C} \quad \overset{\text{D}}{\Rightarrow} \quad (\beta/\sigma_\varepsilon)[\lambda(\lambda-1)D_B/W]^{1/2}$

$d)\ T^{1/2}(\hat{\varphi}^C - \theta^B) \quad \overset{\text{D}}{\Rightarrow} \quad \sigma C_C[\lambda(\lambda-1)D_B]^{-1}$

$d')\ T^{-1/2} t_{\hat{\varphi}^C} \quad \overset{\text{D}}{\Rightarrow} \quad (\theta^B/\sigma_\varepsilon)\,[\lambda(\lambda-1)D_B/Z]^{1/2}$

$e)\ (\hat{\varphi}_*^C - \theta^A) \quad \overset{\text{D}}{\Rightarrow} \quad \varepsilon_{T_B+1}$

$e')\ t_{\hat{\varphi}_*^C} \quad \overset{\text{D}}{\Rightarrow} \quad (\theta^A + \varepsilon_{T_B+1})/\sigma_\varepsilon$

$f)\ T\hat{\varphi}_{**}^C \quad \overset{\text{D}}{\Rightarrow} \quad -\theta^B A_B D_B^{-1}$

$f')\ t_{\hat{\varphi}_{**}^C} \quad \overset{\text{D}}{\Rightarrow} \quad -(\sigma/\sigma_\varepsilon)[\Delta_B D_B]^{-1/2} A_B$

where '^' denotes an OLS estimator,

$$
\begin{aligned}
C_C = {}& \lambda W(1)\,\{4\lambda\textstyle\int W[3\lambda^2(\lambda+1)\int rW - 3\lambda^3\int W - \lambda(2\lambda^3 - 9\lambda^2 + 1)\int_0^\lambda W \\
&+ 3(\lambda^3 - 4\lambda^2 + 1)\int_0^\lambda rW\,] + 12\lambda\int_0^\lambda W[(5\lambda - 3)\int_0^\lambda rW - \lambda(2\lambda - 1)\int_0^\lambda W \\
&- \lambda(3\lambda - 1)\int rW\,] - 12\lambda^3(\int rW)^2 - 12\int_0^\lambda rW[\lambda(\lambda^2 - 6\lambda + 3)\int rW \\
&+ (3\lambda - 2)\int_0^\lambda rW\,] + 2\lambda^3(\lambda - 1)^3\int W^2\,\} + W(\lambda)\,\{4\lambda\int W \\
&\times [\lambda(2\lambda^4 - 3\lambda^3 - 2\lambda + 6)\int W - 6\lambda(\lambda^2 - 3\lambda + 3)\int_0^\lambda W - 3\lambda \\
&\times (3\lambda^2 - 8\lambda + 6)(\lambda + 1)\int rW + 3(\lambda^2 - 3\lambda + 4)\int_0^\lambda rW\,] + 12\int_0^\lambda W \\
&\times [\lambda\int_0^\lambda W + \lambda^2(4\lambda^2 - 11\lambda + 9)\int rW - (\lambda + 1)\int_0^\lambda rW] + 12\lambda\int rW \\
&\times [\lambda(3\lambda^2 - 8\lambda + 6)\int rW - 2(\lambda^2 - 3\lambda + 3)\int_0^\lambda rW] + 12(\int_0^\lambda rW)^2 \\
&- 2\lambda^2(4\lambda - 3)(\lambda - 1)^3\int W^2\,\} + 6\,\{2\lambda^2(\int W)^2[\lambda^3\int W - (2\lambda^3 - \lambda + 2) \\
&\times \int_0^\lambda W - \lambda^2\int rW + (4\lambda^2 - 3)\int_0^\lambda rW\,] + 2(\int_0^\lambda W)^2[\lambda^2(4\lambda^2 - 7\lambda + 6)\int W
\end{aligned}
$$

$$+ \lambda^2(8\lambda - 9)\int rW + \int_0^\lambda rW] - 4\lambda\int w\int_0^\lambda W[\lambda(\lambda^2 + \lambda - 3)\int rW$$

$$+ (5\lambda^2 - 6\lambda + 2)\int_0^\lambda rW] + 12\lambda(\lambda - 1)[\int W\int_0^\lambda rW(\int_0^\lambda rW - \lambda\int rW)$$

$$+ \int_0^\lambda W\int rW(\lambda\int rW - \int_0^\lambda rW)] - 2\lambda(\int_0^\lambda W)^3 + \lambda^2(\lambda - 1)^2\int W^2[\lambda^2\int W$$

$$- (2\lambda - 1)\int_0^\lambda W] \} + \lambda^2(\lambda - 1)^2[2\lambda(4\lambda - 1)\int_0^\lambda W - 2\lambda^2(\lambda + 2)\int W$$

$$+ 6\lambda^2\int rW - 6(2\lambda - 1)\int_0^\lambda rW]\lambda^2(\lambda - 1)^2(W(1)^2 - \sigma_\varepsilon^2 / \sigma^2)$$

and the rest of the expressions have been defined in Theorem 2.

PROOF. See Appendix.

As we can see from the above results, the asymptotic distributions of models (B) and (C) are the same, with the exception of $\hat{\phi}^C$ and $\hat{\phi}_*^C$, this last not being part of model (B), but showing the same type of inconsistency already discussed for model (A). Therefore, the remarks made for Model (B) above apply for Model (C) as well.

Hence, models (B) and (C) are essentially equivalent, as was indeed suggested by Perron, who asserted that the asymptotic distribution of $t_{\hat{\alpha}^B}$ is identical to the asymptotic distribution of $t_{\hat{\alpha}^C}$. This fact led him to argue that, "it is not possible to test for a unit root under the maintained hypothesis that the trend function has a change in slope with the segments joined at the time of the change", Perron (p.1381). To see why this is so, consider detrending the raw series in order to remove the possible non-stationarities due to deterministic factors, including the breaks. Denote $\overset{+}{y}_t^i$ as the residuals from a regression of y_t on: (1) $i = A$: a constant, a time trend, and DU_t; (2) $i = B$: a constant, a time trend, and DT_t^*; (3) $i = C$: a constant, a time trend, DU_t, and DT_t. Also denote $\tilde{\alpha}^i$ the OLS estimator of α^i in the following regression equations:

$$\overset{+}{y}_t^i = \tilde{\alpha}^i \overset{+}{y}_{t-1}^i + \tilde{\varepsilon}_t, \quad i = A, B, C. \tag{14}$$

Perron argues that, as opposed to models (A) and (C), for which the asymptotic distributions of $t_{\hat{\alpha}^A}$ and $t_{\hat{\alpha}^C}$ in (2') and (10') are the same, respectively, as the asymptotic distributions of $t_{\tilde{\alpha}^A}$ and $t_{\tilde{\alpha}^C}$ in (14), for Model (B) such a correspondence does not hold for $t_{\hat{\alpha}^B}$ (6'). This is

because, apart from the dummy variable $D(TB)_t$, regressions (6') and (10') are equivalent.

3.4. Model (B1): The "Broken Trend" Model Revisited

Because of the asymptotic equivalence between models (B) and (C), in order to test the null hypothesis against the broken trend alternative under Model (B), Perron proposes running regression (6') excluding the dummy variable DU_t i.e.:

$$y_t = \hat{\mu}_*^{B1} + \hat{\beta}_*^{B1}t + \hat{\varphi}^{B1}DT_t^* + \hat{\alpha}^{B1}y_{t-1} + \hat{\varepsilon}_t \tag{15}$$

This is because the asymptotic distribution of $t_{\hat{\alpha}^{B1}}$ in this last regression will be the same as that of $t_{\tilde{\alpha}^\beta}$ in (14). The asymptotic distributions reported by Perron for Model (B) correspond to regression equation (15), when the process $\{y_t\}$ follows a driftless random walk.

The next Theorem presents the asymptotic theory corresponding to (15) including, as before, the deterministic parameters. For this reason, we make use of equation (5') as the DGP, instead of a driftless random walk.

THEOREM 4. *Let $\{y_t\}_0^T$ be a sample of size $T+1$ generated from equation (5'). Let Assumption 1 hold and $\lambda = T_B/T$. Then, for regression model (15) [Model (B1)], as $T \to \infty$*

a) $T(\hat{\alpha}^{B1} - 1) \quad \overset{D}{\Rightarrow} \quad D_{B1}^{-1}A_{B1}$

a') $t_{\hat{\alpha}_*^{B1}} \quad \overset{D}{\Rightarrow} \quad (\sigma/\sigma_\varepsilon)(\Delta_B D_{B1})^{-1/2}A_{B1}$

b) $T\hat{\beta}_*^{B1} \quad \overset{D}{\Rightarrow} \quad -\beta D_{B1}^{-1}A_{B1}$

b') $t_{\hat{\beta}_*^{B1}} \quad \overset{D}{\Rightarrow} \quad -(\sigma/\sigma_\varepsilon)(\Delta_B D_{B1})^{-1/2}A_{B1}$

c) $T^{1/2}(\hat{\mu}_*^{B1} - \beta) \quad \overset{D}{\Rightarrow} \quad \sigma D_{B1}^{-1}J_{B1}$

c') $T^{-1/2}t_{\hat{\mu}_*^{B1}} \quad \overset{D}{\Rightarrow} \quad (\beta/\sigma_\varepsilon)[D_{B1}/M]^{1/2}$

d) $T\hat{\varphi}^{B1} \quad \overset{D}{\Rightarrow} \quad -\theta^B D_{B1}^{-1}A_{B1}$

$$d')\ t_{\hat{\varphi}_{B1}} \quad \overset{\mathrm{D}}{\Rightarrow} \ -(\sigma/\sigma_\varepsilon)\,[\Delta_B D_{B1}]^{-1/2} A_{B1},$$

where '^' denotes an OLS estimator, and

$$D_{B1} = \lambda^3(4 - \lambda^3)(\int W)^2 + 6\lambda^2(2 - \lambda)\int W(\int_0^\lambda rW - \lambda\int_0^\lambda W) - 3\lambda^3\int rW$$

$$\times\,((3\lambda - 4)\int rW - 2(\lambda^2 - 2)\int W) - 3\lambda\int_0^\lambda W(2\int_0^\lambda rW - \lambda\int_0^\lambda W)$$

$$+\,6\lambda^2(3 - 2\lambda)\int rW(\lambda\int_0^\lambda W - \int_0^\lambda rW) + 3(\int_0^\lambda rW)^2 + \lambda^3(\lambda - 1)^3\int W^2,$$

$$A_{B1} = \lambda^2 W(1)[3(\lambda - 1)(\int_0^\lambda rW - \lambda\int_0^\lambda W) + 3\lambda(\lambda^2 - 3\lambda + 2)\int rW$$

$$-\,\lambda(\lambda^3 - 3\lambda^2 + 2)\int W] - 3(\lambda - 1)W(\lambda)((2\lambda - 3)\lambda^2\int Wr + (2 - \lambda)\lambda^2$$

$$\times \int W - \lambda\int_0^\lambda W + \int_0^\lambda rW) + \lambda^3(\lambda - 1)^3(1/2)[W(1)^2 - (\sigma_\varepsilon^2/\sigma^2)]$$

$$+\,3[\lambda^2\int W(2(\lambda^2 - \lambda - 1)\int_0^\lambda W - \lambda(\lambda^2 - 2)\int W - \lambda(3\lambda - 4)\int rW)$$

$$-\,\lambda^2(2\lambda - 3)(\int W\int_0^\lambda rW + \int_0^\lambda W\int rW) + \int_0^\lambda W(\lambda\int_0^\lambda W - \int_0^\lambda rW)].$$

$$J_{B1} = \{6\lambda\int W[(\lambda^2 + 1)(\lambda\int_0^\lambda W - \int_0^\lambda rW) - \lambda^3\int W] - 6\lambda^2(\int_0^\lambda W)^2$$

$$+\,6\lambda^3\int rW((\lambda + 1)\int W - \int rW) + 6\lambda(2\lambda^2 - 3\lambda + 3)\int rW(\int_0^\lambda rW - \lambda\int_0^\lambda W)$$

$$+\,6\int_0^\lambda rW(2\lambda\int_0^\lambda W - \int_0^\lambda rW) + \lambda^3(\lambda - 1)^3\int W^2\}\,W(1) + W(\lambda)\{(\lambda - 1)^2$$

$$\times\,[12\int W(\lambda\int W - \lambda\int_0^\lambda rW + \int_0^\lambda rW) + 18\int rW(\lambda\int_0^\lambda W - \int_0^\lambda rW)]$$

$$+\,6\lambda\int rW[(\lambda^3 - \lambda^2 - 6\lambda + 6)\int W - 3(\lambda^2 - 3\lambda + 2)\int rW] - 3\lambda(\lambda - 1)^2$$

$$\times\,\int W^2\}+(1/2)[W(1)^2 - (\sigma_\varepsilon^2/\sigma^2)]\,\{(\lambda - 1)^2\lambda^2(2 - \lambda)\int W$$

$$+\,3(\int_0^\lambda rW - \lambda^2\int rW - \lambda\int_0^\lambda W)\,\} + 6\lambda(\int W)^2[\lambda^3\int W - (\lambda^3 + 3\lambda - 2)$$

$$\times\,\int_0^\lambda W - \lambda^2\int rW + (\lambda^2 + 3)\int_0^\lambda rW] + 6\int W\int_0^\lambda W[(\lambda(3\lambda - 2)\int_0^\lambda W$$

$$-2(\lambda^2 + 3\lambda + 3)\lambda\int rW - 2(3\lambda - 1)\int_0^\lambda rW\,] + 18(\int W\int_0^\lambda rW)(\int_0^\lambda rW$$

$$+ \lambda(\lambda - 2)\int rW) + 3(\lambda - 1)^2\lambda\int rW^2(\lambda^2\int W - \int_0^\lambda W) + 18\int_0^\lambda W\int rW$$

$$\times [\lambda\int_0^\lambda W - \lambda(\lambda - 2)\int rW - \int_0^\lambda rW\,].$$

$$M = 12\lambda\{(\lambda^2 - 3)\int rW[\lambda(\int W - \int_0^\lambda W) + \int_0^\lambda rW\,] + \int W[\lambda(\int W - 2\int_0^\lambda W)$$

$$+ 2\int_0^\lambda rW\,] + \int_0^\lambda W[\lambda(\int_0^\lambda W - 2\int_0^\lambda rW)] + \lambda(3 - 2\lambda)(\int rW)^2\ \}$$

$$+ 12(\int_0^\lambda rW)^2 + \lambda^2(\lambda^4 - 6\lambda^2 + 8\lambda - 3)\int W^2.$$

PROOF. See Appendix.

As in the case of model (B), parts *a)* and *a')* of Theorem 4 show the invariance of the normalized bias $T(\hat{\alpha}^{B1} - 1)$ and the *t*-statistic $t_{\hat{\alpha}^{B1}_*}$ with respect to the value of both the drift β, and the 'breaking-trend' parameter, θ^B.

As with the other models, parts *a)* and *b)* of Theorem 4 show the super-consistency properties of the OLS estimators for α^{B1} and β^{B1}_*. Part *d)* shows that the same property holds for the estimator of φ^{B1}.

Some other results from Theorem 4 mimic those of previous theorems, namely, the asymptotic equality between $t_{\hat{\alpha}^{B1}_*}, - t_{\hat{\beta}^{B1}_*}$ and $- t_{\hat{\varphi}^{B1}}$, providing the link between tests for a unit root and deterministic components. Also, appearing again are the consistency of the constant term at the usual rate of $T^{1/2}$ and the divergent nature of its *t*-statistic, which will allow valid asymptotic inference on $\hat{\mu}^{B1}_*$.

4. Conclusions

This paper has presented the asymptotic theory related to the models proposed by Perron for testing a unit root in time series with (deterministic) structural breaks. Apart from the limiting distributions of the (bias of the) autoregressive parameter and its *t*-statistic for all three models, which had already been derived by Perron, the rest of the asymptotic results presented here have not been reported before in the literature.

These results allow us to visualize the relationship between unit root testing and tests for the presence of structural breaks, as modelled by dummy variables.

The theorems show that, based on the formulation we followed for the testing strategy of models (B1) and (C), there is no conflict between testing hypotheses relating to the unit root and the presence of (deterministic) trends and/or structural breaks. That is, asymptotically, acceptance of a unit root rules out acceptance of a deterministic trend, with or without a structural break in it. For Model (A), however, this limiting 'mirror image' property does not extend to structural breaks: it only holds between the testing of a unit root and a deterministic trend.

The theorems also show both the dependence of some of the estimators and related t-statistics on the parameters in the DGP, and the rates of convergence of the various statistic analyzed.

Appendix

Most of the results needed in the proofs were carried out by the computerized algebra package REDUCE (see for example, Rayna, 1987). Based on the information provided below in the form of the matrices $(X'X)$ and $(X'\varepsilon)$, REDUCE computes the biases of the OLS estimators $\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$ for each model. The expression for each of the biases is then factored by REDUCE so that only the asymptotically relevant terms are analyzed. t-statistics are computed by REDUCE in a similar fashion (note that all t-statistics analyzed are specified with a zero coefficient null).

We only present the proof of Theorem 1, as the proofs of the rest of the theorems follow nearly identical steps. The REDUCE codes and details of the rest of the proofs are available from the author upon request. We will make use of the continuous mapping theorem (see Phillips, 1987) throughout the Appendix.

The following Lemmas about weak convergence of partial sums of the innovations of the type $S_T = \sum_{t=1}^{T} \varepsilon_t$ to functionals of Wiener processes, $W(r)$ will be required.

LEMMA 1: *Let ε_t satisfy Assumption 1 and $T_B = \lambda T$, $\supset T$. Then, as* $T \to \infty$

$$i)\ u \equiv T^{-1/2}\sum_{t=1}^{T}\dot{\varepsilon}_t \qquad \overset{D}{\Rightarrow}\ \sigma W(1)$$

$$ii)\ tu \equiv T^{-3/2}\sum_{t=1}^{T}t\varepsilon_t \qquad \overset{D}{\Rightarrow}\ \sigma[W(1) - \int_0^1 W(r)dr]$$

$$iii)\ u1 \equiv T^{-1/2}\sum_{t=T_{B}+1}^{T}\varepsilon_t \qquad \overset{D}{\Rightarrow}\ \sigma[W(1) - W(\lambda)]$$

$$iv)\ u2 \equiv T^{-1/2}\sum_{t=T_{B}+2}^{T}\varepsilon_t \qquad \overset{D}{\Rightarrow}\ \sigma[W(1) - W(\lambda)]$$

where $\sigma^2 = \lim_{T\to\infty}E(T^{-1}S_T^2)$.

PROOF: See Phillips and Perron (1988) for *i)-ii)*. Parts *iii)* and *iv)* are simple extensions of *i)*, since we can write, say *iii)*, as

$$u1 = T^{-1/2}(\sum_{t=1}^{T} - \sum_{t=1}^{T_B}\varepsilon_t) = T^{-1/2}(S_T - S_{[\lambda T]})$$

LEMMA 2. *Under the same conditions as Lemma* 1, *and* $S_0 = 0$:

$$i)\ T^{-3/2}\sum_{t=1}^{T}S_{t-1} \qquad \overset{D}{\Rightarrow}\ \sigma\int_0^1 W(r)dr$$

$$ii)\ T^{-2}\sum_{t=1}^{T}S_{t-1}^2 \qquad \overset{D}{\Rightarrow}\ \sigma^2\int_0^1 W(r)^2dr$$

$$iii)\ T^{-5/2}\sum_{t=1}^{T}tS_t \qquad \overset{D}{\Rightarrow}\ \sigma\int_0^1 rW(r)dr$$

$$iv)\ T^{-1}\sum_{t=1}^{T}\varepsilon_t S_{t-1} \qquad \overset{D}{\Rightarrow}\ (1/2)[\sigma^2 W(1)^2 - \sigma_\varepsilon^2]$$

$$v)\ T^{-3/2}\sum_{t=1}^{T_B}S_t \qquad \overset{D}{\Rightarrow}\ \sigma\int_0^\lambda W(r)dr$$

$$vi)\ T^{-2}\sum_{t=1}^{T_B}S_t^2 \qquad \overset{D}{\Rightarrow}\ \sigma\int_0^\lambda W(r)^2dr$$

vii) $T^{-5/2} \sum_{t=1}^{T_B} t S_t \qquad \overset{D}{\Rightarrow} \quad \sigma \int_0^{\lambda} r W(r) dr$

PROOF: See Phillips and Perron (1988) for *i)-v)* and Perron (1989) for *vi)-vii)*.

Proof of Theorem 1.
For equation (2') we have that:

$$
(X'X) = \begin{bmatrix}
T & \sum_{t=1}^{T} t & (T - T_B) & 1 & \sum_{t=1}^{T} y_{t-1} \\[2ex]
& \sum_{t=1}^{T} t^2 & \sum_{t=T_B+1}^{T} t & (T_B + 1) & \sum_{t=1}^{T} t y_{t-1} \\[2ex]
& & (T - T_B) & 1 & \sum_{t=T_B+1}^{T} y_{t-1} \\[2ex]
& & & 1 & y_{T_B} \\[2ex]
& & & & \sum_{t=1}^{T} y_{t-1}^2
\end{bmatrix}
$$

where

$$\sum_{t=1}^{T} t = (1/2)(T^2 + T),$$

$$\sum_{t=1}^{T} t^2 = (1/6)(2T^3 + 3T^2 + T),$$

$$\sum_{t=T_B+1}^{T} t = (1/2)[T^2(1 - \lambda^2) + T(1 - \lambda)],$$

$$\sum_{t=1}^{T} DU_t = T - T_B = T(1 - \lambda).$$

Using the DGP (3') we can express the sample moments of y_t as follows:

$$\sum_{t=1}^{T} y_{t-1} = (1/2)\beta T^2 + (T^{-3/2}\sum_{t=1}^{T} S_{t-1})T^{3/2}$$

$$+ [\mu - (1/2)\beta + \theta^A(1-\lambda)]T - \theta^A$$

$$\sum_{t=1}^{T} t y_{t-1} = (1/3)\beta T^3 + (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})T^{5/2}$$

$$+ (1/2)[\mu + \theta^A(1-\lambda^2)]T^2 + (T^{-3/2}\sum_{t=1}^{T} S_{t-1})T^{3/2}$$

$$+ \{(1/2)[\mu + \theta^A(1-3\lambda)] - (1/3)\beta\}T - \theta^A$$

$$\sum_{t=T_B+1}^{T} y_{t-1} = \sum_{t=1}^{T} y_{t-1} - \sum_{t=1}^{T_{iB}} y_{t-1} = (1/2)\beta(1-\lambda^2)T^2$$

$$+ [T^{-3/2}(\sum_{t=1}^{T} S_{t-1} - \sum_{t=1}^{T_B} S_{t-1})]T^{3/2}$$

$$+ (1-\lambda)[\mu + \theta^A - (1/2)\beta]T - \theta^A$$

$$y_{T_B} = \beta\lambda T + (T^{-1/2}\sum_{t=1}^{T_B}\varepsilon_t)T^{1/2} + \mu$$

$$\sum_{t=1}^{T} y_{t-1}^2 = \sum_{t=1}^{T} [\mu + \beta(t-1) + \theta^A DU_{t-1} + S_{t-1}]^2$$

$$= (1/3)\beta^2 T^3 + 2\beta(T^{-5/2}\sum_{t=1}^{T_B}(t-1)S_{t-1})T^{5/2}$$

$$+ \{\beta[\mu + \theta^A(1-\lambda^2)] - (1/2)\beta^2 + (T^{-2}\sum_{t=1}^{T} S_{t-1}^2)\}T^2$$

$$+ \{2\mu(T^{-3/2}\sum_{t=1}^{T} S_{t-1}) + 2\theta^A[T^{-3/2}(\sum_{t=1}^{T} S_{t-1} - \sum_{t=1}^{T_B} S_{t-1})]T^{3/2}$$

$$+ (1/6)\{6\mu(\mu - \beta) + \beta^2 + 6\theta^A(1 - \lambda)[\theta^A + 2\mu - 2\beta] + 6\beta\theta^A(1 - 3\lambda)\}T$$

$$- \theta^A[\theta^A + 2(\mu + \beta)].$$

For (2') we also have that:

$$(X'\underline{\varepsilon}) = \begin{bmatrix} \sum_{t=1}^{T} \varepsilon_t \\ \sum_{t=1}^{T} t\varepsilon_t \\ \sum_{t=T_B+1}^{T} \varepsilon_t \\ \varepsilon_{T_B+1} \\ \sum_{t=1}^{T} y_{t-1}\varepsilon_t \end{bmatrix} = \begin{bmatrix} (T^{-1/2}\sum_{t=1}^{T} \varepsilon_t)T^{1/2} \\ (T^{-3/2}\sum_{t=1}^{T} t\varepsilon_t)T^{3/2} \\ (T^{-1/2}\sum_{t=T_B+1}^{T} \varepsilon_t)T^{1/2} \\ \varepsilon_{T_B+1} \\ \sum_{t=1}^{T} y_{t-1}\varepsilon_t \end{bmatrix}$$

where

$$\sum_{t=1}^{T} \varepsilon_t Y_{t-1} = \sum_{t=1}^{T} \varepsilon_t[\mu + \beta(t - 1) + \theta^A DU_{t-1} + S_{t-1}]$$

$$= \beta(T^{-3/2}\sum_{t=1}^{T} t\varepsilon_t)T^{3/2} + (T^{-1}\sum_{t=1}^{T} \varepsilon_t S_{t-1})T$$

$$+ [(\mu - \beta + \theta^A)(T^{-1/2}S_T) - \theta^A(T^{-1/2}S_{T_B+1})]T^{1/2}$$

Now, making use of *REDUCE*, we get the following expression for the determinant of $(X'X)$:

$$\det(X'X) \equiv D_{A_T} = \{T^7[\delta_1] + O_p(T^6)\}/12,$$

where

$$\delta_1 = \lambda(3\lambda^3 - 4)(T^{-3/2}\sum_{t=1}^{T} S_{t-1})^2 + 12\lambda(1 - \lambda^2)(T^{-3/2}\sum_{t=1}^{T} S_{t-1})$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}) + 2\lambda(4-3\lambda)(T^{-3/2}\sum_{t=1}^{T}S_{t-1})$$

$$\times (T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}) + 12\lambda(\lambda-1)(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1} + T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}) - (T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})^2$$

$$+ \lambda(\lambda-1)(3\lambda^2 - 3\lambda + 1)(T^{-2}\sum_{t=1}^{T}S_{t-1}^2)$$

Then, normalizing D_{A_T} by T^7 and using the weak convergence results of the relevant moments from Lemma 2, we have that, as $T \to \infty$,

$$T^{-7}(12)D_{A_T} \overset{D}{\Rightarrow} -\lambda\sigma^2 D_A, \tag{A1.1}$$

with D_A as defined in Theorem 1.

Next, we instruct *REDUCE* to invert the $(X'X)$ matrix in order to get the 5×1 vector $(\hat{\underline{\beta}} - \underline{\beta})^A = (X'X)^{-1}X'\underline{\varepsilon}$ from which the fifth element is:

$$\hat{\alpha}_*^A \equiv (\hat{\alpha} - 1) = \{T^6[\psi_1] + O_p(T^{5.5})\}/(-12D_{A_T}) \equiv A_{A_T}/-12D_{A_T}$$

where

$$\psi_1 = -\{\lambda T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\,[(3\lambda^3 - 4)(T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + 6(1-\lambda^2)$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}) + (4-3\lambda)(T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})]$$

$$+ 6\lambda(T^{-3/2}\sum_{t=1}^{T}t\varepsilon_t)[(1-\lambda^2)(T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + (\lambda-1)$$

$$\times (2T^{-5/2}\sum_{t=1}^{T_b}(t-1)S_{t-1} + T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})] + (T^{-1/2}\sum_{t=1}^{T_B}\varepsilon_t)$$

$$\times [\lambda(4-3\lambda)(T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + 6\lambda(\lambda-1)(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})$$

$$- (T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})] + \lambda(1-\lambda)(3\lambda^2 - 3\lambda + 1)(T^{-1}\sum_{t=1}^{T}\varepsilon_t S_{t-1})\}.$$

Normalizing by T^{-6} and using the weak convergence results of the relevant moments from Lemmas 1 and 2, we have that, as $T \to \infty$,

$$T^{-6}(A_{A_T}) \overset{D}{\Rightarrow} \lambda\sigma^2 A_A , \qquad (A1.2)$$

with A_A as defined in Theorem 1. Part a) of the Theorem now follows immediately from (A1.1), (A1.2) and the continuous mapping theorem.

To show part a'), note that

$$t_{\hat{\alpha}_*^A} = \hat{\alpha}_*^A / \hat{s}[(X'_5 X_5)^{-1}]^{1/2}.$$

From REDUCE, we get the expression

$$(X'_5 X_5)^{-1} = \{T^5\lambda(\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O(T^4)\} / (-12D_{A_T})$$

so that

$$t_{\hat{\alpha}_*^A} = (\hat{\alpha}_*^A / \hat{s})\{T^5\lambda(\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O(T^4)\}^{-1/2}(-12D_{A_T})^{1/2}$$

From (A1.1) and (A1.2), appropriately normalizing, and since $\hat{s}^2$ converges in probability to σ_ε^2, we obtain:

$$(T\hat{\alpha}_*^A / \hat{s})T^{2.5}\{T^5\lambda(\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O(T^4)\}^{-1/2}T^{-3.5}(-12D_{A_T})^{1/2}$$

$$= t_{\hat{\alpha}_*^A} \overset{D}{\Rightarrow} (A_A / D_A)(1/\sigma_\varepsilon)[\lambda\Delta_A]^{-1/2}(\lambda\sigma^2 D_A)^{1/2} = (\sigma/\sigma_\varepsilon)(\Delta_A D_A)^{-1/2}A_A,$$

thus proving a').

Next, from the (5×1) vector $(\hat{\beta} - \beta)^A$ we take the second element:

$$\hat{\beta}_*^A = \{T^6\beta[-\psi_1] + O_p(T^{5.5})\} / (-12D_{A_T}),$$

where ψ_1 has been defined above for the numerator of $\hat{\alpha}_*^A$. Therefore, normalizing and using both (A1.1), (A1.2) and Lemmas 1 and 2 we obtain the required result:

$$T\hat{\beta}_*^A \overset{D}{\Rightarrow} -\beta\lambda\sigma^2 A_A / \lambda\sigma^2 D_A = -\beta D_A^{-1}A_A .$$

For part b'), we have that

$$t_{\hat{\beta}_*^A} = \hat{\beta}_*^A / [var(\hat{\beta}_*^A)]^{1/2} = \hat{\beta}_*^A / \hat{s}(X'_2 X_2)^{-1}]^{1/2}$$

where, from the *REDUCE* output,

$$(X'_2 X_2)^{-1} = \{T^5 \lambda \beta^2 (\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O_p(T^{4.5})\} / (-12 D_{A_T}).$$

Then,

$$t_{\hat{\beta}_*^A} = (\hat{\beta}_*^A / \hat{s})\{T^5 \lambda \beta^2 (\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O_p(T^{4.5})\}^{-1/2}(-12 D_{A_T})^{1/2},$$

which upon appropriately normalizing, and using the previous results, gives

$$(T\hat{\beta}_*^A / \hat{s})T^{2.5}\{T^5 \lambda \beta^2 (\lambda - 1)(3\lambda^2 - 3\lambda + 1) + O_p(T^{4.5})\}^{-1/2}$$

$$\times T^{-3.5}(-12 D_{A_T})^{1/2} = t_{\hat{\beta}_*^A} \overset{D}{\Rightarrow} (1/\sigma_\varepsilon)(-\beta D_a^{-1} A_A) \times \{\lambda\beta^2 \Delta_A\}$$

$$^{-1/2}(\lambda\sigma^2 D_A)^{1/2} = -(\sigma/\sigma_\varepsilon)(\Delta_A D_A)^{-1/2}A_A,$$

proving b').

Part c) is established as follows. Making use of *REDUCE*, the first element of the (5×1) vector $(\hat{\underline{\beta}} - \underline{\beta})^A$ is:

$$(\hat{\mu}_*^A - \beta) = \{T^{6.5}[\psi_2] + O_p(T^6)\} / (-12 D_{A_T}) \equiv G_T / -12 D_{A_T}.$$

From the *REDUCE* output, we can express ψ_2 as:

$$\psi_2 = T^{-1/2}\sum_{t=1}^{T} \varepsilon_t \quad \{6T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}[\lambda^2(T^{-3/2}\sum_{t=1}^{T}S_{t-1})$$

$$- (2\lambda^2 - 1)(T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})] + 4T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}(T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}$$

$$- T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + 3\lambda^2(\lambda^2 - 1)(T^{-2}\sum_{t=1}^{T}S_{t-1}^2)\} + 6T^{-3/2}\sum_{t=1}^{T} t\,\varepsilon_t$$

$$\times \quad \{T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}[(\lambda^2+1)(T^{-3/2}\sum_{t=1}^{T}S_{t-1})+2(\lambda-1)$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})-T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}]-\lambda^2(T^{-3/2}\sum_{t=1}^{T}S_{t-1})^2$$

$$+\lambda^2(1-\lambda)(T^{-2}\sum_{t=1}^{T}S_{t-1}^2)\}+T^{-1/2}\sum_{t=1}^{T_B}\varepsilon_t\quad\{T^{-3/2}\sum_{t=1}^{T}S_{t-1}$$

$$\times [4T^{-3/2}\sum_{t=1}^{T}S_{t-1}+6(\lambda^2-2)(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})$$

$$-4T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}]+6T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}]T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}$$

$$+2(1-\lambda)(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})]-(3\lambda^2-4\lambda+1)T^{-2}\sum_{t=1}^{T}S_{t-1}^2\}$$

$$+T^{-1}\sum_{t=1}^{T}\varepsilon_tS_{t-1}\{3\lambda^2(1-\lambda^2)(T^{-3/2}\sum_{t=1}^{T}S_{t-1})+6\lambda^2(\lambda-1)$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})+(3\lambda^2-4\lambda+1)(T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})\}.$$

Suitably normalizing, applying (A1.1) to D_{A_T}, and Lemmas 1 and 2 to ψ_2, we get:

$$T^{1/2}(\hat{\mu}_*^A-\beta)\overset{D}{\Rightarrow}\lambda\sigma^3G/\lambda\sigma^2D_A=\sigma GD_A^{-1},$$

as required.

To prove c') we write $t_{\hat{\mu}_*^A}=\hat{\mu}_*^A/\hat{s}[(X_1'X_1)^{-1}]^{1/2}$, where, from *REDUCE*, we have:

$$(X_1'X)^{-1}=\{T^6[\psi_3]+O_p(T^{5.5})\}/(-12D_{A_T}),$$

where,

$$\psi_3=4T^{-3/2}\sum_{t=1}^{T}S_{t-1}(T^{-3/2}\sum_{t=1}^{T}S_{t-1}-2T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})$$

$$+12(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})[(\lambda^2-1)(T^{-3/2}\sum_{t=1}^{T}S_{t-1}-T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})$$

$$+ (1 - \lambda)(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})] + 4(T^{-3/2}\sum_{t=1}^{T_B}S_{t-1})^2$$

$$+ (3\lambda^4 - 6\lambda^2 + 4\lambda - 1)(T^{-2}\sum_{t=1}^{T}S_{t-1}^2)\}.$$

Therefore,

$$t_{\hat{\mu}_*^A} = (\hat{\mu}_*^A / \hat{s})\{T^6[\psi_3] + O_p(T^{5.5})\}^{-1/2}(-12D_{A_T})^{1/2}.$$

Now, from above we know that $T^{1/2}(\hat{\mu}_*^A - \beta) \overset{\Psi}{\Rightarrow} \sigma GD_A^{-1}$ from which we can deduce that $\hat{\mu}_*^A \overset{D}{\Rightarrow} \beta$. Using this, and suitably normalizing we get:

$$(\hat{\mu}_*^A / \hat{s})T^3\{T^6[\psi_3] + O_p(T^{5.5})\}^{-1/2}T^{-3.5}(-12D_{A_T})^{1/2}$$

$$= T^{-1/2}t_{\hat{\mu}_*^A} \overset{D}{\Rightarrow} (\beta / \sigma_\varepsilon)(\sigma^2\lambda\Gamma_\mu)^{-1/2}(\sigma^2\lambda D_A)^{1/2} = (\beta / \sigma_\varepsilon)(D_A / \Gamma_\mu)^{1/2},$$

by using Lemmas 1 and 2, thus establishing c').

To prove d) we use the third element of the vector $(\hat{\underline{\beta}} - \underline{\beta})^A$, computed by *REDUCE*

$$(\hat{\phi}^A - \theta^A) = \{T^{6.5}[\psi_4] + O_p(T^6)\} / (-12D_{A_T}),$$

where

$$\psi_4 = T^{-1/2}\sum_{t=1}^{T}\varepsilon_t [6\lambda(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})(2T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}$$

$$- \lambda T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}(4T^{-3/2}\sum_{t=1}^{T}S_{t-1}$$

$$- 6T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}) + \lambda(3\lambda - 4)(T^{-2}\sum_{t=1}^{T}S_{t-1}^2)] + 6T^{-3/2}\sum_{t=1}^{T}t\varepsilon_t$$

$$\times [\lambda T^{-3/2}\sum_{t=1}^{T}S_{t-1}(\lambda T^{-3/2}\sum_{t=1}^{T}S_{t-1} - 2T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})$$

$$+ T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}(2T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1} - T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + \lambda(1-\lambda)$$

$$\times (T^{-2}\sum_{t=1}^{T}S_{t-1}^2)] + T^{-1/2}\sum_{t=1}^{T_B}\varepsilon_t [4T^{-3/2}\sum_{t=1}^{T}S_{t-1}(3T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}$$

$$-T^{-3/2}\sum_{t=1}^{T}S_{t-1}) - 12(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})^2 + T^{-2}\sum_{t=1}^{T}S_{t-1}^2]$$

$$+ T^{-1}\sum_{t=1}^{T}\varepsilon_t S_{t-1}[\lambda(4 - 3\lambda(T^{-3/2}\sum_{t=1}^{T}S_{t-1}) + 6\lambda(\lambda - 1)$$

$$\times (T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1}) - T^{-3/2}\sum_{t=1}^{T_B}S_{t-1}].$$

Noting that, by using Lemmas 1 and 2, $\psi_4 \overset{D}{\Rightarrow} \lambda\sigma^3 H$, result d) follows.

For d'), we write the t-statistic as $t_{\hat{\varphi}^A} = \hat{\varphi}^A / \hat{s}[(X'_3 X_3)^{-1}]^{1/2}$, where from the *REDUCE* output we have:

$$(X'_3 X_3)^{-1} = \{T^6[\psi_5] + O_p(T^{5.5})\} / (-12D_{A_T}),$$

where

$$\psi_5 = 4(T^{-3/2}\sum_{t=1}^{T}S_{t-1})^2 - 12(T^{-3/2}\sum_{t=1}^{T}S_{t-1})(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})$$

$$+ 12(T^{-5/2}\sum_{t=1}^{T}(t-1)S_{t-1})^2 - (T^{-2}\sum_{t=1}^{T}S_{t-1}^2).$$

Then,

$$t_{\hat{\varphi}^A} = (\hat{\varphi}^A / \hat{s})\{T^6[\psi_5] + O_p(T^{5.5})\}^{-1/2}(-12D_{A_T})^{1/2}$$

$$= (T^{1/2}\hat{\varphi}^A / \hat{s})T^3\{T^6[\psi_5] + O_p(T^{5.5})\}^{-1/2}T^{-3.5}(-12D_{A_T})^{1/2}.$$

Using previous results and Lemmas 1 and 2 on ψ_5, we obtain

$$t_{\hat{\varphi}^A} \overset{D}{\Rightarrow} (\sigma H D_A^{-1} / \sigma_\varepsilon)\{\lambda\sigma^2\Delta_\varphi\}^{-1/2}(\lambda\sigma^2 D_A)^{1/2} = (\sigma / \sigma_\varepsilon)(\Delta_\varphi D_A)^{-1/2}H,$$

proving d').

To prove e) we use the fourth element of the vector $(\hat{\underline{\beta}} - \underline{\beta})^A$, computed by *REDUCE*

$$(\hat{\varphi}_*^A - \theta^A) = \{T^7[\psi_6]\varepsilon_{T_B+1} + O_p(T^{6.5})\} / (-12D_{A_T}),$$

where $\psi_6 = -[\delta_1]$, defined above for D_{A_T}. Then, since $\psi_6 \overset{D}{\Rightarrow} \lambda\sigma^2 D_A$, result e) follows.

For e'), we write the t-statistic as $t_{\hat{\varphi}^A_*} = \hat{\varphi}^A_* / \hat{s}[(X'_4 X_4)^{-1}]^{1/2}$, where from the *REDUCE* output we have:

$$(X'_4 X_4)^{-1} = \{T^7[\psi_6] + O_p(T^6)\} / (-12 D_{A_T}).$$

Then,

$$t_{\hat{\varphi}^A_*} = (\hat{\varphi}^A_* / \hat{s})T^{3.5}\{T^7[\psi_6] + O_p(T^6)\}^{-1/2} T^{-3.5}(-12 D_{A_T})^{1/2}$$

$$\overset{D}{\Rightarrow} (\theta^A + \varepsilon_{T_B+1}) / \sigma_\varepsilon (\lambda \sigma^2 D_A)^{-1/2} (\lambda \sigma^2 D_A)^{1/2} = (\theta^A + \varepsilon_{T_B+1}) / \sigma_\varepsilon,$$

proving e').

References

Bhargava, A. (1986). "On the Theory of Testing for Unit Roots in Observed Time Series", *Review of Economic Studies*, 53, pp. 369-384.

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

———— (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, pp. 1057-1072.

Dickey, D. A., W. R. Bell, and R. B. Miller (1986). "Unit Roots in Time Series Models: Tests and Implications", *The American Statistician*, 40, pp. 13-26.

Dolado, J. J. and T. Jenkinson (1987). *Co-integration: A Survey of Recent Developments*, Applied Economics Discussion Paper no. 39, Institute of Economics and Statistics, University of Oxford.

Engle, R., and C. W. Granger (1987). "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, vol. 55, pp. 251-276.

Engle, R., and B. S. Yoo (1987). "Forecasting and Testing in Co-integrated Systems", *Journal of Econometrics*, 35, pp. 143-159.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, J. Wiley.

Granger C. W. (1986). "Developments in the Study of Co-integrated Economic Variables", *Oxford Bulletin of Economics and Statistics*, 48, pp. 213-228.

Haldrup, N. (1990). "Tests for Unit Roots with a Maintained Trend when the true Data Generating Process is a Random Walk with Drift", Memo 1990-22, Institute of Economics, University of Aarhus.

Hearn, A. C. (1987). *REDUCE User's Manual*, Version 3.3, The RAND Corporation, Santa Monica, CA (RAND Publication CP78 Rev. 7/87). Reproduced by the University of Manchester Regional Computer Centre, NWD 73, 1988.

Nelson, Ch. R., and Ch. I. Plosser (1982). "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, 10, pp. 139-162.

Noriega-Muro A. E. (1992). *Nonstationarity and Structural Breaks in Economic Time Series: Asymptotic Theory and Monte Carlo Simulations,* Ph.D. Thesis, Department of Econometrics, University of Manchester.

———— (1993a). "On the Consistency of Some Dummy Variable Coefficient Estimates", Department of Econometrics, University of Manchester (mimeo).

———— (1993b). *Asymptotic Theory of Statistics from Unit Root Test Regressions when the Alternative is a Breaking-Trend-Stationary Model*, Department of Econometrics, Discussion Paper ES240, University of Manchester.

———— (1993c). *Nonstationarity and Structural Breaks in Economic Time Series*, Avebury, Ashgate Publishing Ltd. England.

Perron. P. (1989). "The Great Crash, The Oil Price Shock, and The Unit Root Hypothesis", *Econometrica*, 57, pp. 1361-1401.

Phillips, P. C. B. (1987). "Time Series Regression with a Unit Root", *Econometrica*, 55, pp. 277-301.

————, and P. Perron (1988). "Testing for a Unit Root in Time Series Regression", *Biometrika*, 75, pp. 335-346.

Rappoport, P., and L. Reichlin (1987). "Segmented Trends and Nonstationary Time Series", European University Institute Working Paper no. 877319.

———— (1989). "Segmented Trends and Nonstationary Time Series", *The Economic Journal*, 99, pp. 168-177.

Rayna, G. (1987). *REDUCE, Software for Algebraic Computation*, New York, Springer-Verlag.

Said, E. S., and D. A. Dickey (1984). "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order", *Biometrika*, 71, pp. 599-607.

Schmidt, P., and P. C. B. Phillips (1989). "Testing for a Unit Root in the Presence of Deterministic Trends" (mimeo).

Stock, J. H., and M. Watson (1988). "Variable Trends in Economic Time Series", *Journal of Economic Perspectives*, 2, pp. 147-174.

Zivot, E., and D. W. K. Andrews (1992). "Further Evidence on the Great Crash, The Oil Price Shock, and The Unit Root Hypothesis", *Journal of Business and Economic Statistics*, vol. 10, no. 3, pp. 251-270.