



**Eastern Mediterranean University**

*Department of Economics*

*Discussion Paper Series*

**Point Optimal Invariant Tests of a Unit Root in Models  
with Structural Change**

*Mehmet Balcilar*

*Discussion Paper 15-50*

September 1997

*Department of Economics*

*Eastern Mediterranean University*

*Famagusta, North Cyprus*

# Point Optimal Invariant Tests of a Unit Root in Models with Structural Change\*

Mehmet Balcilar

Department of Economics  
Eastern Mediterranean University  
Famagusta, North Cyprus

mehmet.balcilar@emu.edu.tr

September 2007

## Abstract

This study proposes tests for the null hypothesis of a unit root against shifting trend stationary alternatives. In particular, models with structural change in the form of level or trend shifts are studied. Roots local-to-unity are addressed in models with structural change and asymptotic power functions of the tests are derived using local-to-unity asymptotic framework. These tests have good power properties against shifting trend stationary alternatives. Tests in the presence of multiple shifts are also examined. The asymptotic distributions of the tests are derived under exogenous and endogenous shift point assumptions. The asymptotic and finite sample percentiles of the tests are tabulated using Monte Carlo integration. Monte Carlo simulations reveal that tests have good power and size properties against various alternatives.

**JEL Classification:** C22

**Keywords:** Unit roots, structural change, quasi-differencing

Most economic time series show a strong nonreversible tendency to grow. Thus, the original treatment of testing for a unit root against *trend stationary* (TS) alternatives by Nelson and Plosser (1982) can be well justified. It seems reasonable to decompose a time series into stochastic and deterministic components and examine the stochastic nature of the series from the *detrended series*. An assumption made here is, of course, that the detrending leaves only the *stochastic component*. The behavior of the detrended series will heavily depend on the specification of the deterministic component. If deterministic component is misspecified inference from unit root tests will be misleading. A misspecified trend function heavily distorts the test results. It is well known that many economic time series display heterogeneous behavior in their deterministic component. It is observed that this heterogeneous behavior display itself in the form of *level shifts*, *trend shifts*, or both. A level shift corresponds to a change in the mean of series and a trend shift corresponds to a change in the growth rate of series. This type of

---

\*The paper is largely based on my PhD dissertation, Balcilar (1996).

behavior is described as “*variable trend*” by Stock and Watson (1988) and “*breaking trends*” by Perron (1989). What is meant by this terminology is that the time series goes a structural change in the form of level or trend shifts, yielding a more general class of trend stationary processes. We will call this class of trend stationary time series *shifting trend stationary* (STS), hereafter.

Unit root tests against linear trend stationary alternatives have been studied extensively in the literature, following Dickey (1976), Fuller (1976), Dickey and Fuller (1979, 1981). The Dickey-Fuller tests were extended to more general cases by Said and Dickey (1984), Phillips (1987a), Phillips (1988), and Phillips and Perron (1988). A profusion of theoretical as well as empirical research followed these papers. On the theoretical side, many tests of general  $I(1)$  null against stationary,  $I(0)$ , alternatives are proposed and properties of these tests are examined. On the empirical side, researchers tested many economic time series for the existence of unit roots. In some cases, however, it appears that the constant growth rate implied by linear trend may not be appropriate and hard to justify. In these cases, higher order polynomial time trends seem to better approximate the behavior of the time series. The models that allow for higher order polynomial time trends in deterministic component have been examined by Park and Choi (1988) and Ouliaris, Park and Phillips (1989). Diebold and Nerlove (1990) argue that higher order polynomial time trends can indeed approximate STS models, and therefore, unit root models. Indeed, once a quadratic trend is included in the regression, distribution of the test statistics will be different and inference based on critical values in Fuller (1976) will be unqualified. Thus, there is an uncertainty about what higher order polynomial time trends represent.

Perron (1989) is the first study raising the possibility of STS models. He suggests that the null hypothesis of unit root should be tested against an STS alternative. He finds that once a one time change in the trend function is allowed, the null of a unit root is rejected for 11 out of 14 time series examined by Nelson and Plosser (1982).

Perron (1989) treats the timing of shift as exogenous. Although, this is subject to various criticism it often looks appealing. Recently Perron (1990a), Banerjee, Lumsdaine and Stock (1992), Perron and Vogelsang (1992), and Zivot and Andrews (1992) extended unit root tests against STS alternatives with endogenous shift point. A detailed exposition of Perron’s approach as well as a different interpretation of his models based on intervention analysis is made by Noriega-Muro (1993). Noriega-Muro (1993) also has analytic results as well as finite sample analysis of the models under this interpretation. Recently, Park and Sung (1994) extended Perron’s models allowing higher order time polynomials in the deterministic component. Although,

Park and Sung (1994) maintain the exogenous shift point assumption they allow for finite but many trend shifts.

Unit root tests in general are shown to have low power. The power issue becomes more serious under the STS alternatives. Recent research on unit root tests have concentrated on developing test statistics with good power properties. Sargan and Bahargava (1983) and Bahargava (1986) develop most powerful invariant (MPI) test of a unit root hypothesis. Nabeya and Tanaka (1990) derive both locally best invariant (LBI) and locally best invariant unbiased (LBIU) tests, although, LBI test is no use in some cases. Dufour and King (1991) also develop LBI and point optimal invariant (POI) tests, later one being constructed as MPI. Saikkonen and Luukkonen (1993) constructs POI test of a unit root in a first-order Gaussian moving average model and generalize the result to more general *ARIMA* models. Finally, Elliot, Rothenberg and Stock (1992) develop POI tests under local alternatives with generalized least squares (GLS) detrended data.

Although, no uniformly most powerful (UMP) unit root test exists, Elliot et al. (1992) (ERS hereafter) correct POI test of Dufour and King (1991) for nuisance parameter dependency under a Gaussian  $AR(p + 1)$  and construct new statistics. One of the test statistics, which is constructed with locally GLS demeaned series under the alternative, is approximately asymptotically UMP. ERS also show that the Dickey-Fuller test statistics constructed from locally GLS demeaned series under the alternative is also approximately asymptotically UMP. Two other tests, which are constructed from locally demeaned and linearly detrended series under the alternative, are asymptotically uniformly most powerful invariant (UMPI).

All these tests have good power properties compared to the Dickey-Fuller tests. Unfortunately, they are constructed under the TS alternatives. It is well known that, unit root tests have worse power properties when the data is generated by an STS process. Balcilar (1996) investigates power of these tests under structural changes and shows that good power properties of these tests are not preserved when there is a shift in the trend function. In a simulation exercise, Balcilar (1996) shows that as the size of either a level shift or a trend shift increases these tests essentially have zero power after the size of the shift reaches a threshold value. In this study, we construct new test statistics in the lines of Elliot et al. (1992) under local-to-unity alternatives that incorporate level or trend shifts in the data. An important feature of these tests is that they are constructed under roots local-to-unity. The local-to-unity asymptotic framework allows us to establish asymptotic power functions of the tests. We maximize the power of the tests using the asymptotic power functions. As in the previous studies, we examine both additive and innovational outlier models. Further, multiple level and trend shifts are also

allowed. Asymptotic distribution of the tests statistics when the alternative is an STS process is analytically derived under exogenous and endogenous shift point assumptions. Finite sample properties of the estimators used to construct the tests are extensively investigated using Monte Carlo simulations. Finite sample and asymptotic percentiles of the tests are tabulated using Monte Carlo integration. Monte Carlo investigations show that tests have good power and size properties.

The plan of this paper is as follows: Section 1 discusses the models we consider and states some assumptions and conditions that are used in deriving the asymptotic distributions of the tests. Section 2 derives the asymptotic distribution of the test statistics under the assumption of a one time structural change at a known time. In Section 3, we generalize the tests to a more class of models. In particular, the innovational outlier model is considered, multiple structural changes are allowed, and the assumption of exogenous time of structural change is relaxed. The finite sample properties of the estimators are examined in Section 4. The asymptotic and finite sample critical values of the tests are derived in Section 5. The finite sample power and size properties of the tests are investigated in Section 6. Section 7 concludes the paper.

## 1 The Hypothesis, Models and Assumptions

Our approach is based on the local-to-unity asymptotic framework of Bobkoski (1983), Cavanagh (1985), Chan and Wei (1987), and Phillips (1987b). The local-to-unity asymptotic framework allows us to derive the asymptotic power functions of the tests and makes power comparisons among the competing tests possible. we propose tests of a unit root when the alternative is an STS process. Test statistics we propose are useful modifications of the Dickey-Fuller unit root tests under STS alternatives. The modifications to the Dickey-Fuller unit root tests are achieved using the transformation and method developed by Elliot et al. (1992). We construct the test statistics along the lines of Elliot et al. (1992) under local-to-unity alternatives that incorporate the level and trend shifts in the trend function. An important feature of these tests is that they are constructed under roots local-to-unity. The local-to-unity asymptotic framework allows us to establish asymptotic power functions of the tests. We maximize the power of the tests using these asymptotic power functions. As in the previous studies we examine both the additive and innovational outliers models. Further, multiple level and trend shifts are also allowed. The asymptotic distributions of the test statistics when the alternative is an STS process are analytically derived under both exogenous and endogenous shift point assumptions. In the demeaned case, tests are approximately asymptotically UMP, whereas in the demeaned and linearly detrended case tests are approximately asymptotically UMPI.

We consider time series  $\{y_t\}$  which are the sum of a purely deterministic component  $d_t$  and a purely stochastic component  $u_t$ :

$$y_t = d_t + u_t, \quad (1.1)$$

where

$$(1 - \rho L)u_t = v_t. \quad (1.2)$$

The literature on unit roots mostly assumes that  $v_t$  is an  $AR(p)$ . We impose a weaker condition on  $v_t$  and assume that  $v_t$  satisfy the conditions in Assumption 1.1.

**Assumption 1.1.** Let  $\{v_t\}_{t=1}^{\infty}$  be a stochastic process and define  $S_T = \sum_{j=1}^T v_j$  such that

- (a)  $\mathbf{E}(v_t) = 0 \quad \forall t$ ,
- (b)  $\sup_t \mathbf{E} |u_t|^{2\beta} < \infty$  for some  $\beta > 2$ ,
- (c)  $\sigma^2 = \lim_{T \rightarrow \infty} \mathbf{E} [T^{-1} S_T^2]$  exists, and  $\sigma^2 > 0$ ,
- (d)  $\{v_t\}$  is strong mixing with mixing coefficients  $\{\alpha_m\}$  satisfying

$$\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty. \quad (1.3)$$

These conditions allow a wide variety of weakly dependent and heterogeneously distributed DGP's for  $\{v_t\}$ . These include the finite order *ARIMA* models with quite general conditions on the innovations. (b) basically controls allowable heterogeneity. It is also a sufficient condition for the existence of the variance and the non-integer higher order moments of  $\{v_t\}$ . However, it is also sufficiently weak to allow some heterogeneity since  $\mathbf{E} |v_t|^{\beta+\xi}$  is not a constant. (c) is needed to ensure nondegenerate limiting distributions. (d) controls the allowable temporal dependence in relation to the probability of outlier occurrences controlled by (b) (see Phillips (1987a) for further details).

Usually the first observation  $u_1$  is assumed to be zero to simplify derivations. However, integrated processes heavily depend on initial conditions. Therefore, we will assume that  $u_1$  is unknown but satisfy the following condition:

**Condition 1.1.**  $\mathbf{E} [u_1^2] < \infty$ .

That is,  $u_1$  has a finite second moment. This includes the special case that  $u_1$  is fixed.

One of the conditions in the following assumption will be very useful to simplify some of the derivations.

**Assumption 1.2.** Each of the following alternatives are equivalent in terms of the result they lead to:

- (a)  $S_0 = 0$ ,

- (b)  $y_0 = c$  with probability one, where  $c$  is a constant,
- (c)  $y_0$  has a certain distribution,
- (d)  $y_0 = y_T$  where  $T$  is finite sample size.

Further, let

$$\rho = \exp\left(\frac{c}{T}\right) \approx 1 + \frac{c}{T}. \quad (1.4)$$

The crucial role of the expression in (1.4) is that it allows us to derive the asymptotic power functions of the tests. Then, one can maximize the power of the tests by a suitable choice of  $c$  under the alternative, say  $\bar{c}$ . Note also that the models are constructed to be near-stationary in finite samples, thus, roots local-to-unity are explicitly addressed. However, models are asymptotically nonstationary since  $\lim_{T \rightarrow \infty} (1 + c/T) = 1$ .

Following Phillips (1987b), let  $B_c(r)$  defined by

$$B_c(r) = \int_0^r e^{(r-s)c} dB(s). \quad (1.5)$$

The functional  $B_c(r)$  is known as an Ornstein-Uhlenbeck process and will play an important role in our derivations. For fixed  $r > 0$ ,  $B_c(r)$  is distributed as

$$B_c(r) \sim \mathbf{N}\left(0, \frac{1}{2} (e^{2rc} - 1) / c\right).$$

Here,  $B_c(r)$  is generated by the stochastic differential equation:

$$dB_c(r) = cB_c(r)dr + dB(r), \quad B_c(0) = 0.$$

Sometimes it is useful to use the following definition of  $B_c(r)$ :

$$B_c(r) = B(r) + c \int_0^r e^{(r-s)c} B(s) ds, \quad (1.6)$$

which is simple to derive from (1.5). From Phillips (1987b), we obtain the following result:

$$B_c(1)^2 = 1 + 2c \int_0^1 B_c(r)^2 dr + 2 \int_0^1 B_c(r) dB(r), \quad (1.7)$$

which is obtained by stochastically integrating  $(\int_0^r e^{-sc} dB(s))^2$ .

Since the model involves detrending, results will depend on the trend estimation error which we denote by

$$\xi_t = \tilde{d}_t - d_t. \quad (1.8)$$

Now let,  $U_T(r) = T^{-1/2}u_{[Tr]}$ , and  $\Upsilon_T(r) = T^{-1/2}\xi_{[Tr]}$ . Following Stock (1994), we impose the following condition:

**Condition 1.2** (Detrending conditions). Let  $\{y_t\}$  and  $\{u_t\}$  be defined as in (1.1)-(1.2), if  $u_t$  is  $I(1)$ :

- (a)  $(U_T, \Upsilon_T) \implies \sigma(B_c, \Lambda_c)$  where  $\Lambda_c \in C[0, 1]$ ,
- (b)  $\|\Delta\xi_t\| = O_p(1)$ .

The hypothesis of interest is

$$H_0 : \quad \rho = 1 \quad \text{against the alternative} \quad H_1 : \quad \rho = \bar{\rho}.$$

where  $\bar{\rho} = 1 + \bar{c}/T$ .

We also allow a one time shift in the trend function  $d_t$  at a known time  $T_B$ . For the level shift model we define

$$d_t = \mu + \delta t + \theta D_t, \quad D_t = \{0|t \leq T_B, 1|t > T_B\}. \quad (1.9)$$

A trend shift is modeled as

$$d_t = \mu + \delta t + \theta D_t, \quad D_t = \{0|t \leq T_B, t - T_B|t > T_B\}. \quad (1.10)$$

To simplify the derivations let

$$d_t = \beta' z_t, \quad (1.11)$$

where  $\beta = (\mu, \delta, \theta)'$  and  $z_t = (1, t, D_t)'$ .

Now, in order to obtain the locally GLS detrended series we apply the following transformations:

$$(\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_T) = (y_1, (1 - \bar{\rho}L)y_2, \dots, (1 - \bar{\rho}L)y_T) \quad (1.12a)$$

$$(\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_T) = (z_1, (1 - \bar{\rho}L)z_2, \dots, (1 - \bar{\rho}L)z_T) \quad (1.12b)$$

This is a valid transformation as discussed by Dufour and King (1991). Briefly, when testing for the null of a general  $I(1)$  ( $\rho = 1$ ) against the point alternative  $\rho = \bar{\rho}$ , a general requirement is that the tests should be invariant to the transformations of the form  $y_t \mapsto ay_t + \mathbf{b}'z_t$ , where  $a$  is a positive scalar and  $\mathbf{b}$  is a finite vector with dimension conformable to  $z_t$ . All unit root tests meet this invariance requirement.

Note that with the trend function specified as in (1.9) or (1.10), we cannot construct asymptotically point optimal tests. The reason here is not the introduction of the level or trend shift dummy  $D_t$  (at least for the level shift model), but the presence of the linear trend. For the trend shift model a linear trend already exists and  $D_t$  does not change the order of the existing trend, only the slope of the trend variable is affected by  $D_t$ . In the absence of a linear trend (this includes dropping  $t$  in (1.9) and both  $t$  and  $D_t$  in (1.10)) one can still construct point optimal

unit root tests since in this case unit root test statistics constructed from both  $y_t$  and  $u_t$  would have the same asymptotic power function, an observation initially made by Elliot et al. (1992). However, when a linear or higher order trend is present test statistics that are constructed from  $y_t$  and  $u_t$  will not have the same limiting distribution. Therefore, tests constructed assuming either (1.9) or (1.10) will be necessarily restricted to the class of point optimal invariant tests which are invariant to nuisance parameters.

The results from the following two lemmas will be extensively used to derive the asymptotic distributions of the test statistics.

**Lemma 1.1.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the condition (c) of Assumption 1.2, and assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $\rho = 1+c/T$ , then,*

$$(a) \quad T^{-1/2} \sum_{t=2}^T v_t \implies \sigma B(1),$$

$$(b) \quad T^{-1} \sum_{t=2}^T u_{t-1} v_t \implies \sigma^2 \int_0^1 B_c(r) dB(r) + (1/2)(\sigma^2 - \sigma_v^2),$$

$$(c) \quad T^{-3/2} \sum_{t=2}^T t v_t \implies \sigma \left( B(1) - \int_0^1 B(r) dr \right),$$

$$(d) \quad T^{-3/2} \sum_{t=2}^T u_{t-1} \implies \sigma \int_0^1 B_c(r) dr,$$

$$(e) \quad T^{-2} \sum_{t=2}^T u_{t-1}^2 \implies \sigma^2 \int_0^1 B_c(r)^2 dr,$$

$$(f) \quad T^{-5/2} \sum_{t=2}^T t u_{t-1} \implies \sigma \int_0^1 r B_c(r) dr,$$

$$(g) \quad T^{-1/2} \sum_{t=2}^T \Delta u_t \implies \sigma B_c(1),$$

$$(h) \quad T^{-3/2} \sum_{t=2}^T t \Delta u_t \implies \sigma \left( B_c(1) - \int_0^1 B_c(r) dr \right).$$

*Proof.* See Phillips (1987b) for (a)-(f) and Appendix for (g) and (h). □

**Lemma 1.2.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the condition (c) of Assumption 1.2. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $\lambda = T_B/T$  and  $\rho = 1 + c/T$ , then,*

$$(a) \quad T^{-1/2} u_{T_B} = T^{-1/2} u_{T_B+1} = T^{-1/2} u_{T_B+2} \implies \sigma B_c(\lambda),$$

$$(b) \quad T^{-1/2} \sum_{t=T_B+2}^T \Delta u_t \implies \sigma (B_c(1) - B_c(\lambda)),$$

$$(c) \quad T^{-3/2} \sum_{t=T_B+2}^T u_{t-1} \implies \sigma \int_0^\lambda B_c(r) dr,$$

$$(d) \quad T^{-3/2} \sum_{t=T_B+2}^T t \Delta u_t \implies \sigma \left( \left( B_c(1) - \lambda B_c(\lambda) \right) - \left( \int_0^1 B_c(r) dr - \int_0^\lambda B_c(r) dr \right) \right),$$

$$(e) \quad T^{-5/2} \sum_{t=T_B+2}^T t u_t \implies \sigma \left( \int_0^1 r B_c(r) dr - \int_0^\lambda r B_c(r) dr \right).$$

*Proof.* See Appendix. □

## 2 Efficient and Near Efficient Tests

The test statistics we study are based on the locally GLS detrended data. We define the estimated trend using the transformed data by

$$\tilde{d}_t = \tilde{\beta}' \tilde{z}_t, \tag{2.13}$$

where

$$\tilde{\beta} = \left( \sum_{t=1}^T \tilde{z}_t \tilde{z}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{z}_t \tilde{y}_t \right). \tag{2.14}$$

Now we derive the GLS biases:

$$(\tilde{\beta} - \beta) = \left( \sum_{t=1}^T \tilde{z}_t \tilde{z}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{z}_t \tilde{u}_t \right), \tag{2.15}$$

where  $(\tilde{\beta} - \beta) = ((\tilde{\mu} - \mu), (\tilde{\delta} - \delta), (\tilde{\theta} - \theta))'$ .

First, we consider trend estimation with a level shift.

**Theorem 2.1.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (1.9),  $\lambda = T_B/T$ , and  $\rho = 1 + c/T$ , then,*

$$T^{1/2}(\tilde{\mu} - \mu - u_1) \implies \sigma \left( -\frac{3\bar{c}^2(\bar{c}^2 - 2\bar{c} + 2)}{2\bar{c}^2 - 3\bar{c} + 3} \int_0^1 r B_c(r) dr + \bar{c}^2 \int_0^1 B_c(r) dr + \frac{1}{2} \frac{(\bar{c}^3 - 3\bar{c}^2 + 6\bar{c} - 6)}{\bar{c}^2 - 3\bar{c} + 3} B_c(1) \right), \tag{2.16}$$

$$T^{1/2}(\tilde{\delta} - \delta) \implies \sigma \left( 3 \frac{\bar{c}^2}{\bar{c}^2 - 3\bar{c} + 3} \int_0^1 r B_c(r) dr + 3 \frac{1 - \bar{c}}{\bar{c}^2 - 3\bar{c} + 3} B_c(1) \right), \tag{2.17}$$

$$(\tilde{\theta} - \theta) \implies v_{T_B+1}. \tag{2.18}$$

*Proof.* See Appendix. □

Note that,  $\tilde{\theta}$  is inconsistent even in large samples. This means that an uncertainty remains about the size of the level shift. This is an expected result since  $(1 - \bar{\rho}L)D_t$  is an ‘‘ill-behaving’’ variable. Inconsistency of  $\tilde{\theta}$ , however, will not affect the asymptotic distributions of the test statistics since the test statistics are invariant to  $\tilde{\theta}$ .

Now, we will consider detrending under a trend shift.

**Theorem 2.2.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (1.10),  $\lambda = T_B/T$ , and  $\rho = 1 + c/T$ , then,*

$$\begin{aligned}
T^{1/2}(\tilde{\mu} - \mu - u_1) \implies & \sigma \left( \frac{\bar{c}^4 \lambda^2 (\lambda - 1) (\bar{c} \lambda - \bar{c} + 3)}{m_1(\bar{c}, \lambda)} B_c(1) + 3 \frac{m_2(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} B_c(\lambda) \right. \\
& + \frac{\bar{c}^4 \lambda^2 (\bar{c}^2 \lambda^2 + \bar{c}^2 \lambda - 2 \bar{c}^2 - 6 + 6 \bar{c})}{m_1(\bar{c}, \lambda)} \int_0^1 B_c(r) dr \\
& - 3 \frac{\bar{c}^2 m_2(\bar{c}, \lambda)}{m_1(\bar{c}, \lambda)} \int_0^\lambda B_c(r) dr - 3 \frac{\bar{c}^5 \lambda^2 (\bar{c} \lambda + 2 - \bar{c})}{m_1(\bar{c}, \lambda)} \int_0^1 r B_c(r) dr \\
& \left. + 3 \frac{\bar{c}^2 m_2(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} \int_0^\lambda r B_c(r) dr \right), \tag{2.19}
\end{aligned}$$

$$\begin{aligned}
T^{1/2}(\tilde{\delta} - \delta) \implies & \sigma \left( 6 \frac{\bar{c}^2 (\lambda - 1) (\bar{c} \lambda - \bar{c} + 3)}{m_1(\bar{c}, \lambda)} B_c(1) - 6 \frac{m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} B_c(\lambda) \right. \\
& - 6 \frac{\bar{c}^2 m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} \int_0^1 B_c(r) dr + 6 \frac{\bar{c}^2 m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} \int_0^\lambda B_c(r) dr \\
& \left. + 18 \frac{\bar{c}^3 (\bar{c} \lambda + 2 - c)}{m_1(\bar{c}, \lambda)} \int_0^1 r B_c(r) dr - 6 \frac{\bar{c}^2 m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} \int_0^\lambda r B_c(r) dr \right), \tag{2.20}
\end{aligned}$$

$$\begin{aligned}
T^{1/2}(\tilde{\theta} - \theta) \implies & \sigma \left( -6 \frac{3\bar{c}^2 - \bar{c}^3 + 6\bar{c}^3 \lambda^2 - 6\bar{c} - \bar{c} 2 \lambda^2}{m_1(\bar{c}, \lambda) (\lambda - 1)} B_c(1) \right. \\
& + 12 \frac{3 - 3\bar{c} + \bar{c}^2}{\lambda m_1(\bar{c}, \lambda) (\lambda - 1)} B_c(\lambda) + 12 \frac{\bar{c}^2 (3 - 3\bar{c} + \bar{c}^2)}{m_1(\bar{c}, \lambda) (\lambda - 1)} \int_0^1 B_c(r) dr \\
& - 12 \frac{\bar{c}^2 (3 - 3\bar{c} + \bar{c}^2)}{\lambda m_1(\bar{c}, \lambda) (\lambda - 1)} \int_0^\lambda B_c(r) dr + 18 \frac{\bar{c}^3 (\bar{c} \lambda + 2 - c)}{m_1(\bar{c}, \lambda)} \\
& \times \int_0^\lambda r B_c(r) dr + 6 \frac{\bar{c}^2 (-3\bar{c}^2 + 6\bar{c} - 6 + \bar{c}^2 \lambda^2)}{m_1(\bar{c}, \lambda) (\lambda - 1)} \int_0^\lambda r B_c(r) dr \\
& \left. + 12 \frac{\bar{c}^2 (3 - 3\bar{c} + \bar{c}^2)}{\lambda m_1(\bar{c}, \lambda) (\lambda - 1)} \int_0^\lambda r B_c(r) dr \right). \tag{2.21}
\end{aligned}$$

where

$$\begin{aligned}
m_1(\bar{c}, \lambda) = & \bar{c}^4 \lambda^4 + \bar{c}^4 \lambda^3 - 5 \bar{c}^4 \lambda^2 - 12 \bar{c}^2 \lambda^2 + 12 \bar{c}^3 \lambda^2 + 3 \lambda \bar{c}^4 + 12 \bar{c}^2 \lambda \\
& - 12 \bar{c}^3 \lambda + 12 \bar{c}^2 - 36 \bar{c} + 36,
\end{aligned}$$

$$m_2(\bar{c}, \lambda) = -\lambda \bar{c}^4 - 4 \bar{c}^2 \lambda + 12 \bar{c} + 2 \bar{c}^2 \lambda^2 - 2 \bar{c}^3 \lambda^2 + \bar{c}^4 \lambda^2 + 4 \bar{c}^3 \lambda - 4 \bar{c}^2 - 12.$$

$$m_3(\bar{c}, \lambda) = \bar{c}^2 \lambda^2 + \bar{c}^2 \lambda - 2 \bar{c}^2 - 6 + 6 \bar{c}.$$

*Proof.* See Appendix. □

Based on the estimated trend parameters we define the following detrended series:

$$y_t^{\tau\mu} = y_t - \tilde{\mu} - \tilde{\delta}t - \tilde{\theta}D_t, \quad D_t = \{0|t \leq T_B, 1|t > T_B\}, \quad (2.22)$$

$$y_t^{\tau\delta} = y_t - \tilde{\mu} - \tilde{\delta}t - \tilde{\theta}D_t, \quad D_t = \{0|t \leq T_B, t - T_B|t > T_B\}. \quad (2.23)$$

From Theorem 2.1 and Theorem 2.2 following lemma is easily proven.

**Lemma 2.1.** *Under local GLS detrending:*

$$T^{-1/2}y_{[Tr]}^{\tau\mu} \implies \sigma B_c^{\tau\mu}(r), \quad (2.24)$$

$$T^{-1/2}y_{[Tr]}^{\tau\delta} \implies \sigma B_c^{\tau\delta}(r), \quad (2.25)$$

where

$$B_c^{\tau\mu}(r) = B_c(r) - r \left( 3 \frac{\bar{c}^2}{\bar{c}^2 - 3\bar{c} + 3} \int_0^1 r B_c(r) dr + 3 \frac{1 - \bar{c}}{\bar{c}^2 - 3\bar{c} + 3} B_c(1) \right), \quad (2.26)$$

$$\begin{aligned} B_c^{\tau\delta}(r) = & B_c(r) - r \left( 6 \frac{\bar{c}^2(\lambda - 1)(\bar{c}\lambda - \bar{c} + 3)}{m_1(\bar{c}, \lambda)} B_c(1) - 6 \frac{m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} B_c(\lambda) \right. \\ & \left. - 6 \frac{\bar{c}^2 m_3(\bar{c}, \lambda)}{\lambda m_1(\bar{c}, \lambda)} \int_0^1 B_c(r) dr \right) - (r - \lambda) \mathbf{1}(r > \lambda) \\ & \times \left( -6 \frac{3\bar{c}^2 - \bar{c}^3 + 6\bar{c}^3\lambda^2 - 6\bar{c} - \bar{c}2\lambda^2}{m_1(\bar{c}, \lambda)(\lambda - 1)} B_c(1) + 12 \frac{3 - 3\bar{c} + \bar{c}^2}{\lambda m_1(\bar{c}, \lambda)(\lambda - 1)} B_c(\lambda) \right. \\ & + 12 \frac{\bar{c}^2(3 - 3\bar{c} + \bar{c}^2)}{m_1(\bar{c}, \lambda)(\lambda - 1)} \int_0^1 B_c(r) dr - 12 \frac{\bar{c}^2(3 - 3\bar{c} + \bar{c}^2)}{\lambda m_1(\bar{c}, \lambda)(\lambda - 1)} \int_0^\lambda B_c(r) dr \\ & + 18 \frac{\bar{c}^3(\bar{c}\lambda + 2 - c)}{m_1(\bar{c}, \lambda)} \int_0^\lambda r B_c(r) dr + 6 \frac{\bar{c}^2(-3\bar{c}^2 + 6\bar{c} - 6 + \bar{c}^2\lambda^2)}{m_1(\bar{c}, \lambda)(\lambda - 1)} \int_0^\lambda r B_c(r) dr \\ & \left. + 12 \frac{\bar{c}^2(3 - 3\bar{c} + \bar{c}^2)}{\lambda m_1(\bar{c}, \lambda)(\lambda - 1)} \int_0^\lambda r B_c(r) dr \right), \end{aligned} \quad (2.27)$$

where  $\mathbf{1}(\cdot)$  is the indicator function.

We will consider two unit root test statistics, one is based on the autoregressive coefficient and the other is the usual  $t$  statistic. There are two ways to construct these test statistics.

First, we can use the Said-Dickey tests from the following regressions:

$$\Delta y_t^{\tau\mu} = \alpha y_{t-1}^{\tau\mu} + \sum_{j=1}^p a_j \Delta y_{t-j}^{\tau\mu} + \xi_t, \quad (2.28)$$

$$\Delta y_t^{\tau\delta} = \alpha y_{t-1}^{\tau\delta} + \sum_{j=1}^p a_j \Delta y_{t-j}^{\tau\delta} + \xi_t, \quad (2.29)$$

where  $\alpha = \rho - 1$ . Now, the test statistics are

$$\Gamma_1 = \frac{T(\hat{\rho} - \rho)}{1 - \sum_{j=1}^p \hat{a}_j}, \quad (2.30)$$

$$\Gamma_2 = t_\rho = \frac{\hat{\rho} - \rho}{\hat{s}_\rho}, \quad (2.31)$$

where  $\hat{s}_\rho$  is the standard error of  $\hat{\rho}$ . A second alternative is to use the Phillips-Perron type unit root tests. The Phillips-Perron tests are based on the following regressions:

$$\Delta y_t^{\tau_\mu} = \alpha y_{t-1}^{\tau_\mu} + v_t, \quad (2.32)$$

$$\Delta y_t^{\tau_\delta} = \alpha y_{t-1}^{\tau_\delta} + v_t. \quad (2.33)$$

The Phillips-Perron type unit root test statistics are then defined by

$$\Gamma_1^* = T(\hat{\rho} - \rho) - (1/2)(\hat{\sigma}^2 - \hat{\sigma}_v^2) \left( T^{-2} \sum_{t=1}^T y_{t-1}^{\tau_i^2} \right)^{-1}, \quad (2.34)$$

$$\Gamma_2^* = (\hat{\sigma}/\hat{\sigma}_v)t_\rho - (1/2)(\hat{\sigma}^2 - \hat{\sigma}_v^2) \left[ \left( T^{-2} \sum_{t=1}^T y_{t-1}^{\tau_i^2} \right)^{1/2} \right]^{-1}. \quad (2.35)$$

where  $i = \mu, \delta$ . The Said-Dickey and the Phillips-Perron type unit root test statistics are asymptotically equivalent, but, we will state the asymptotic distributions of the tests in terms of the Phillips-Perron tests which are given by

**Theorem 2.3.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $\lambda = T_B/T$ , and  $\rho = 1 + c/T$ . If  $d_t$  is given by (1.9), then,*

$$\Gamma_1^* \implies c + \left( \int_0^1 B_c^{\tau_\mu}(r) dB(r) \right) \left( \int_0^1 B_c^{\tau_\mu}(r)^2 dr \right)^{-1}, \quad (2.36)$$

$$\Gamma_2^* \implies \left( \left( \int_0^1 B_c^{\tau_\mu}(r) dB(r) \right) \left( \int_0^1 B_c^{\tau_\mu}(r)^2 dr \right)^{-1} + c \right) \left( \int_0^1 B_c^{\tau_\mu}(r)^2 dr \right)^{1/2}. \quad (2.37)$$

If  $d_t$  is given by (1.10), then,

$$\Gamma_1^* \implies c + \left( \int_0^1 B_c^{\tau_\delta}(r) dB(r) \right) \left( \int_0^1 B_c^{\tau_\delta}(r)^2 dr \right)^{-1}, \quad (2.38)$$

$$\Gamma_2^* \implies \left( \left( \int_0^1 B_c^{\tau_\delta}(r) dB(r) \right) \left( \int_0^1 B_c^{\tau_\delta}(r)^2 dr \right)^{-1} + c \right) \left( \int_0^1 B_c^{\tau_\delta}(r)^2 dr \right)^{1/2}. \quad (2.39)$$

*Proof.* See Appendix. □

We observe that the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are invariant to the shift fraction  $\lambda$  in case of a level shift. Thus, the tests have the same asymptotic distributions irrespective of the time of the shift. This is an important result since this invariance property makes the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  applicable to more general cases and easily extendible. We will introduce some of these extensions later in this chapter. Further, because of this invariance property, the same critical values are valid in a variety of cases for the level shift model.

However, in the case of the trend shift the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  do depend on the shift fraction  $\lambda$ . Therefore, the critical values of the test statistics should be tabulated for different values of  $\lambda$  for the trend shift model. Although, the asymptotic distributions of the test statistics in the case of a level shift do not depend on the shift fraction  $\lambda$ , the assumption of an exogenous shift point play an important role in deriving the asymptotic power functions of the test statistics. Namely, asymptotic power functions are obtained over a fixed value of the shift fraction  $\lambda$ .

### 3 Extensions to More General Cases

The unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  can also be applied to more general cases than the case of a single shift at a known time. In this section, we extend the tests to three more cases. First, the models set out in (1.9) and (1.10) are the so called “additive outliers” models in which the shifts in the deterministic component of the series have an instantaneous effect. We generalize the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  to the “innovational outliers” models in which the effect of the shifts in the deterministic component of the series is assumed to be gradual. Second, we relax the assumption of a single shift and consider models with multiple shifts at known times. The assumption of exogenous shift point has been criticized for several reasons. For instance, Christiano (1992) criticizes the exogenous shift point assumption by arguing that it is “pretesting” the data. He further argues that conventionally computed critical values overstate the likelihood of the STS alternatives. The significance of these type of critics will retain their validity as long as the distributions of the test statistics depend on the preset shift point  $T_B$  and it is assumed that the  $T_B$  is exogenously determined. Our third generalization of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  allows endogenous shift points and, thus, resolves the problems associated with pretesting or the exogenous shift point assumption.

#### 3.1 Innovational Outliers Model

One main drawback of the level and trend shift models examined above is that they assume that the effects of the shifts in the trend function occur instantaneously. Such an assumption may not be appealing for some applications. For instance, the Great Crash was not an instantaneous event. In order to overcome this difficulty, one may use the so called *innovational outliers model* in which the response of the economy to the shifts in the deterministic component of the series is modeled as a gradual event. Perron (1989) also addresses the same issue. He suggests the following specification to model the gradual response of the economy to the shifts in the trend

function:

$$d_t = \mu + \delta t + \theta\gamma(L)D_t, \quad (3.40)$$

where  $\gamma(L)$  is a stationary and invertible lag polynomial in the lag operator  $L$ .

Perron (1989) assumes that the economy responds to shifts in the trend function in the same way as it responds to any other shock. This amounts to the imposition of some restrictions on the lag structure  $\gamma(L)$ . Since there is no reason to believe that the transitional dynamics of the economy does not change in times of the extreme events such as the Great Crash and oil price shocks, we relax Perron's assumption and impose no restriction on the lag structure  $\gamma(L)$ .

Let  $\gamma(L)$  be given by

$$\gamma(L) = \gamma_0 + \gamma_1 L + \dots + \gamma_k L^k.$$

Using this specification, we can define the trend function in the following way:

$$d_t = (\mu, \delta, \boldsymbol{\alpha}') (1, t, \mathbf{c}_t')' = \boldsymbol{\beta}' \mathbf{z}_t, \quad (3.41)$$

where  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_k)'$  is a  $(k+1)$ -dimensional parameter vector,  $\mathbf{c}_t = (D_t, \dots, D_{t-k})'$  is a vector containing  $k+1$  shift dummies  $D_{t-i}$  ( $i = 0, 1, \dots, k$ ),  $\boldsymbol{\beta} = (\mu, \delta, \boldsymbol{\alpha}')'$  with dimension  $k+3$ , and  $\mathbf{z}_t = (1, t, \mathbf{c}_t')'$  with dimension  $k+3$ . Let,

$$\mathbf{h}_t = \boldsymbol{\phi}_T^{-1} \mathbf{z}_t, \quad (3.42)$$

where  $\boldsymbol{\phi}_T^{-1}$  is a  $(k+3) \times (k+3)$  diagonal scaling matrix such that

$$\boldsymbol{\phi}_T^{-1} \mathbf{z}_{[Tr]} \implies \mathbf{h}(r). \quad (3.43)$$

For instance, if  $D_{t-i}$  are  $k+1$  level shift dummies, then,

$$\boldsymbol{\phi}_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 & 0 & \dots & 0 \\ 0 & T^{-3/2} & 0 & 0 & \dots & 0 \\ 0 & 0 & T^{-1/2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & T^{-1/2} \end{bmatrix} \quad (3.44)$$

If  $D_{t-i}$  are  $k+1$  trend shift dummies, then, we define  $\boldsymbol{\phi}_T^{-1}$  by

$$\boldsymbol{\phi}_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 & \dots & 0 \\ 0 & T^{-3/2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T^{-3/2} \end{bmatrix} \quad (3.45)$$

The elements of the matrix  $\mathbf{h}_t$  involve sums of various powers of deterministic trends. We state the following results that apply to these sums without proof.

**Fact 3.1.**

$$\sum_{t=1}^T t = (1/2)T(T+1) = O(T^2) \qquad T^{-2} \sum_{t=1}^T t \longrightarrow 1/2 \qquad (3.46a)$$

$$\sum_{t=1}^T t^2 = (1/6)T(T+1)(2T+1) = O(T^3) \qquad T^{-3} \sum_{t=1}^T t^2 \longrightarrow 1/3 \qquad (3.46b)$$

$$\sum_{t=1}^T t^3 = ((1/2)T(T+1))^2 = O(T^4) \qquad T^{-4} \sum_{t=1}^T t^3 \longrightarrow 1/4 \qquad (3.46c)$$

$$\begin{aligned} \sum_{t=1}^T t^4 &= (1/6)T(T+1)(2T+1) \\ &\times \left[ (1/5)(3T + (T+1) - 1) \right] = O(T^5) \end{aligned} \qquad T^{-5} \sum_{t=1}^T t^4 \longrightarrow 1/5 \qquad (3.46d)$$

The expressions in Fact 3.1 can be generalized as

$$\sum_{t=1}^T t^{n-1} = v_1 T + v_2 T^2 + \dots + v_n T^n, \qquad (3.47a)$$

$$T^{-n} \sum_{t=1}^T t^{n-1} \longrightarrow 1/n. \qquad (3.47b)$$

This generalization can easily be verified using the transformation  $r = 1/T, 2/T, \dots, 1, r \in [0, 1]$ .

Then,

$$T^{-1} \sum_{t=1}^T \left[ \frac{t}{T} \right]^{n-1} \longrightarrow \int_0^1 r^n dr = \left[ \frac{r^n}{n} \right]_0^1 = \frac{1}{n} \quad r \in [0, 1], \quad t = 1, \dots, T. \qquad (3.48)$$

Using the results from Fact 3.1 and (3.48), it is simple to establish that

$$\sum_{t=1}^T \mathbf{h}_t \mathbf{h}_t' = \int_0^1 \mathbf{h}_{[Tr]} \mathbf{h}'_{[Tr]} dr + o(1) \implies \int_0^1 \mathbf{h}(r) \mathbf{h}(r)' dr. \qquad (3.49)$$

From Lemma 1.1 and Lemma 1.2, we have

$$T^{-1} \sum_{t=1}^T \mathbf{h}_t u_t \implies \sigma \int_0^1 B_c(r) \mathbf{h}(r) dr, \qquad (3.50)$$

$$T^{-1} \sum_{t=1}^T \mathbf{h}_t \Delta u_t \implies \sigma \int_0^1 dB_c(r) \mathbf{h}(r) dr, \qquad (3.51)$$

$$T^{-1} \sum_{t=1}^T \mathbf{h}_t v_t \implies \sigma \int_0^1 dB_c(r) \mathbf{h}(r) dr. \qquad (3.52)$$

Since models involve detrending, first we derive GLS biases for the trend estimation:

$$\phi_T(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t \tilde{\mathbf{h}}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t \tilde{u}_t \right). \qquad (3.53)$$

Define the detrended process:

$$y_t^{\tilde{\mu}} = u_t^{\mu} - \left( \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t^{\mu} \tilde{\mathbf{h}}_t^{\mu'} \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t^{\mu} \tilde{u}_t^{\mu} \right) \right)' \tilde{\mathbf{h}}_t^{\mu} \qquad (3.54)$$

for the level shift model, and

$$y_t^{\tau\delta} = u_t^\delta - \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t^\delta \tilde{\mathbf{h}}_t^{\delta'} \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{h}}_t^\delta \tilde{u}_t^\delta \right)' \tilde{\mathbf{h}}_t^\delta \quad (3.55)$$

for the trend shift model. Using (3.49)-(3.52) we can establish that

$$T^{1/2} \phi_T(\tilde{\beta} - \beta) \implies \left( \int_0^1 \tilde{\mathbf{h}}(r) \tilde{\mathbf{h}}(r)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{h}}(r) dr \right). \quad (3.56)$$

Having established the limiting distributions of the GLS bias terms, we can derive the limiting values of the detrended processes  $y_t^{\tau\mu}$  and  $y_t^{\tau\delta}$ , then, the asymptotic distributions of the test statistics are simple to obtain from these limits.

**Theorem 3.1.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (3.41) with  $D_{t-i}$  defined by  $D_{t-i} = \{0|t \leq (T_B + i), 1|t > (T_B + i)\}$  for  $i = 0, 1, \dots, k$ ,  $\lambda = T_B/T$ , and  $\rho = 1 + c/T$ , then,  $T^{-1/2} y_t^{\tau\mu} \implies \sigma B_c^{\tau\mu}(r)$ , where  $B_c^{\tau\mu}(r)$  is defined in (2.26), and the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by (2.36) and (2.37).*

*Proof.* See Appendix. □

It was shown before that the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are invariant to  $\theta$  and  $\lambda$  in the case of a level shift assuming that the economy responds to shifts in the trend function instantaneously. The result in Theorem 3.1 shows that the invariance property of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  to the nuisance parameters  $\lambda$  and  $\alpha$  still holds, if we assume that the economy responds to the shifts in the trend function in the form of a change in the mean of the series gradually rather than instantaneously. Therefore, the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are still given by (2.36) and (2.37). This means that we can still use the same critical values even though the detrending regression includes additional lags of the level shift dummy  $D_t$ .

In contrary to the level shift model, the invariance property of the tests is not preserved for the trend shift model. The asymptotic distributions of the tests depend on the shift fraction  $\lambda$  and in general additional lags of the trend shift dummy. However, the dependence on the additional lags can be removed by a simple modification. Let

$$D_t^\delta = \{0|t \leq T_B, t - T_B|t > T_B\}, \quad (3.57)$$

and,

$$D_t^\mu = \{0|t \leq T_B, 1|t > T_B\}. \quad (3.58)$$

Using the relationship  $D_t^\delta - D_{t-1}^\delta = D_t^\mu$ , we can replace the additional lags of the trend shift dummy  $D_t^\delta$  by the additional lags of the level shift dummy  $D_t^\mu$ . Thus, we redefine  $\mathbf{c}_t$  by

$\mathbf{c}_t = (D_t^\delta, D_t^\mu, D_{t-1}^\mu, \dots, D_{t-k+1}^\mu)'$ . Since the tests are invariant to the level shift dummies  $D_{t-i}^\mu$  ( $i = 0, 1, \dots, k-1$ ) a significant simplification is achieved.

**Theorem 3.2.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (3.41) with  $\mathbf{c}_t$  defined by  $\mathbf{c}_t = (D_t^\delta, D_t^\mu, D_{t-1}^\mu, \dots, D_{t-k+1}^\mu)'$ ,  $\lambda = T_B/T$ , and  $\rho = 1 + c/T$ , then,  $T^{-1/2}y_t^{\tau\delta} \implies \sigma B_c^{\tau\delta}(r)$ , where  $B_c^{\tau\delta}(r)$  is defined in (2.27), and the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by (2.38) and (2.39).*

*Proof.* See Appendix. □

The important result of Theorem 3.1 and Theorem 3.2 is that the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  have the same asymptotic distributions under the innovational outliers model for both the level and trend shift models. For the level shift model invariance to the innovational outliers assumption is a natural property of the test statistics. However, the tests are in general not invariant to the inclusion of additional lags of the trend shift dummy. The invariance in Theorem 3.2 is achieved by transforming the regression equations using the relationship  $D_t^\delta - D_{t-1}^\delta = D_t^\mu$ .

### 3.2 Multiple Level and Trend Shifts

Another extension of the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  is to allow multiple shifts. Since the test statistics are invariant to level shifts, extension to multiple shifts for the level shift model does not pose any difficulty. Allowing multiple shifts in the trend shift model, however, requires further considerations.

Assume that there are  $q$  level or trend shifts at times  $T_{B_1}, T_{B_2}, \dots, T_{B_q}$ . Since  $T_{B_k}/T \rightarrow \lambda_k$  for  $k = 1, 2, \dots, q$ , let  $\lambda = (\lambda_1, \dots, \lambda_q)'$  where  $\lambda_k \in (0, 1)$  for  $k = 1, 2, \dots, q$ . Now, let the deterministic component be given by

$$d_t = (\mu, \delta, \boldsymbol{\alpha}') (1, t, \mathbf{c}_t)' = \boldsymbol{\beta}' \mathbf{z}_t, \quad (3.59)$$

where  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_q)'$  is a  $q$ -dimensional parameter vector,  $\mathbf{c}_t = (D_{T_{B_1}}, \dots, D_{T_{B_q}})'$  is a vector containing  $q$  shift dummies  $D_{T_{B_k}}$  ( $k = 1, 2, \dots, q$ ),  $\boldsymbol{\beta} = (\mu, \delta, \boldsymbol{\alpha}')$  with dimension  $q + 2$ , and  $\mathbf{z}_t = (1, t, \mathbf{c}_t)'$  with dimension  $q + 2$ . Let,

$$\mathbf{g}_t = \boldsymbol{\phi}_T^{-1} \mathbf{z}_t \quad (3.60)$$

where  $\boldsymbol{\phi}_T^{-1}$  is a  $(q + 2) \times (q + 2)$  diagonal scaling matrix such that

$$\boldsymbol{\phi}_T^{-1} \mathbf{z}_{[T\tau]} \implies \mathbf{g}(r). \quad (3.61)$$

For the level shift model where  $D_{T_{B_k}}$  are  $q$  level shift dummies,  $\phi_T^{-1}$  is given by

$$\phi_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 & 0 & \dots & 0 \\ 0 & T^{-3/2} & 0 & 0 & \dots & 0 \\ 0 & 0 & T^{-1/2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & T^{-1/2} \end{bmatrix} \quad (3.62)$$

For the  $q$  trend shift dummies  $D_{T_{B_k}}$ , we define  $\phi_T^{-1}$  by

$$\phi_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 & \dots & 0 \\ 0 & T^{-3/2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T^{-3/2} \end{bmatrix} \quad (3.63)$$

Then, from the results of Fact 3.1 and (3.48), we can readily show that

$$\sum_{t=1}^T \mathbf{g}_t \mathbf{g}_t' = \int_0^1 \mathbf{g}_{[Tr]} \mathbf{g}'_{[Tr]} dr + o(1) \implies \int_0^1 \mathbf{g}(r) \mathbf{g}(r)' dr. \quad (3.64)$$

As in the innovational outliers model, we have

$$T^{-1} \sum_{t=1}^T \mathbf{g}_t u_t \implies \sigma \int_0^1 B_c(r) \mathbf{g}(r) dr, \quad (3.65)$$

$$T^{-1} \sum_{t=1}^T \mathbf{g}_t \Delta u_t \implies \sigma \int_0^1 dB_c(r) \mathbf{g}(r) dr, \quad (3.66)$$

$$T^{-1} \sum_{t=1}^T \mathbf{g}_t v_t \implies \sigma \int_0^1 dB_c(r) \mathbf{g}(r) dr, \quad (3.67)$$

from Lemma 1.1 and Lemma 1.2. The GLS biases for trend estimation are given by

$$\phi_T(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t \tilde{\mathbf{g}}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t \tilde{u}_t \right). \quad (3.68)$$

Now, the detrended processes  $y_t^{\tau_\mu}$  and  $y_t^{\tau_\delta}$  can be obtained from

$$y_t^{\tau_\mu} = u_t^\mu - \left( \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\mu \tilde{\mathbf{g}}_t^{\mu'} \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\mu \tilde{u}_t^\mu \right) \right)' \tilde{\mathbf{g}}_t^\mu \quad (3.69)$$

for the level shift model, and

$$y_t^{\tau_\delta} = u_t^\delta - \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\delta \tilde{\mathbf{g}}_t^{\delta'} \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\delta \tilde{u}_t^\delta \right)' \tilde{\mathbf{g}}_t^\delta \quad (3.70)$$

for the trend shift model. The results of (3.64)-(3.67) allow us to establish

$$T^{1/2} \phi_T(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \implies \left( \int_0^1 \tilde{\mathbf{g}}(r) \tilde{\mathbf{g}}(r)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{g}}(r) dr \right). \quad (3.71)$$

Having established the limiting distribution of the GLS bias terms, we can derive the limiting values of the detrended processes  $y_t^{\tau\mu}$  and  $y_t^{\tau\delta}$  under multiple shifts. Then, the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are simple to obtain from these limits.

**Theorem 3.3.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (3.59) with  $D_{T_{B_k}}$  defined by  $D_{T_{B_k}} = \{0|t \leq T_{B_k}, 1|t > T_{B_k}\}$  for  $k = 1, 2, \dots, q$ . Define  $\lambda_k = T_{B_k}/T$  and  $\rho = 1 + c/T$ , then,  $T^{-1/2}y_t^{\tau\mu} \implies \sigma B_c^{\tau\mu}(r)$ , where  $B_c^{\tau\mu}(r)$  is defined in (2.26), and the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by (2.36) and (2.37).*

*Proof.* See Appendix. □

The result in Theorem 3.3 shows that the invariance property of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  is preserved, if a single level shift at a known time is replaced by multiple level shifts at known times. Thus, the asymptotic distributions of the test statistics are still given by (2.36) and (2.37) and the same critical values can be used even though we allow multiple level shifts.

The invariance property of the tests is not preserved for the trend shift model once multiple trend shifts are allowed. The asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  depend on the shift fractions  $\lambda_k$ .

**Theorem 3.4.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (3.59) with  $\mathbf{c}_t$  defined by  $\mathbf{c}_t = (D_{T_{B_1}}, \dots, D_{T_{B_q}})'$  where  $D_{T_{B_k}} = \{0|t \leq T_{B_k}, t - T_{B_k}|t > T_{B_k}\}$  for  $k = 1, 2, \dots, q$ . Define  $\lambda_k = T_{B_k}/T$  and  $\rho = 1 + c/T$ , then,  $T^{-1/2}y_t^{\tau\delta} \implies \sigma \tilde{B}_c^{\tau\delta}(r)$ , and the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by*

$$\Gamma_1^* \implies c + \left( \int_0^1 \tilde{B}_c^{\tau\delta}(r) dB(r) \right) \left( \int_0^1 \tilde{B}_c^{\tau\delta}(r)^2 dr \right)^{-1}, \quad (3.72)$$

$$\Gamma_2^* \implies \left( \left( \int_0^1 \tilde{B}_c^{\tau\delta}(r) dB(r) \right) \left( \int_0^1 \tilde{B}_c^{\tau\delta}(r)^2 dr \right)^{-1} + c \right) \left( \int_0^1 \tilde{B}_c^{\tau\delta}(r)^2 dr \right)^{1/2}, \quad (3.73)$$

where

$$\tilde{B}_c^{\tau\delta}(r) = B_c(r) - \left( \left( \int_0^1 \tilde{\mathbf{g}}^\delta(r) \tilde{\mathbf{g}}^\delta(r)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{g}}^\delta(r) dr \right) \right)' \tilde{\mathbf{g}}^\delta(r). \quad (3.74)$$

*Proof.* See Appendix. □

As in the single trend shift case, the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  depend on the shift fractions  $\lambda_k$ . This is an unappealing feature since tabulating the critical values of the test statistics becomes quite burdensome for large  $q$ . Park and Sung (1994) describe a simple method to remove the shift fraction dependency of the Perron's tests. Unfortunately, this is only valid for  $c = 0$ . There is no useful transformation that removes the shift point dependency as long as  $c \neq 0$ . For this reason, it is simpler to treat the shift point as endogenous

and develop sequential unit root tests so that all possible trend shifts are taken into account. This is the subject of the next section.

### 3.3 Endogenous Shift Point

In this section, we relax the exogenous shift point assumption and assume that the process  $\{y_t\}$  can be represented with a level or trend shifts at unknown points in time under the alternative hypotheses. This extension is similar to the endogenous shift point model studied in Perron (1990b), Banerjee et al. (1992), Perron and Vogelsang (1992), and Zivot and Andrews (1992).

Although there are several ways to model the endogenous shift point, the method we use is to compute the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  sequentially over the full sample and choose the shift point such that the null hypothesis of a unit root is given the least weight. Test statistics are sequentially computed over the full sample by incrementing the time of the shift by a factor of one at each step. We use the largest window possible to compute these sequential test statistics. Thus, if the lag truncation parameter in the regression is  $p$ , then, the largest window is  $[p + 2, T - 1]$  for a sample size of  $T$ . Thus, one computes  $T - (k + 2)$  statistics over a sample size of  $T$  sequentially increasing the shift point by a factor of one and chooses the minimum of the sequence of the values of the tests so that the shift point chosen gives the least favorable result to the null hypothesis of a unit root.

Since  $\lambda \in (0, 1)$  is treated to be endogenous, some changes in our notation are required. Let the deterministic component be given by

$$d_t = (\mu(\lambda), \delta(\lambda), \theta(\lambda))(1, t, D_t(\lambda))' = \beta(\lambda)'z_t(\lambda) \quad (3.75)$$

where  $D_t(\lambda) = \{0|t \leq \lambda T, 1|t > \lambda T\}$  for a level shift and  $D_t(\lambda) = \{0|t \leq \lambda T, t - \lambda T|t > \lambda T\}$  for a trend shift. Let,

$$g_t(\lambda) = \phi_T^{-1}z_t(\lambda), \quad (3.76)$$

where  $\phi_T^{-1}$  is a  $3 \times 3$  diagonal scaling matrix such that

$$\phi_T^{-1}z_{[Tr]}(\lambda) \implies g(r, \lambda). \quad (3.77)$$

Here,  $\phi_T^{-1}$  is defined as

$$\phi_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-1/2} \end{bmatrix}$$

for the level shift model, and

$$\phi_T^{-1} = \begin{bmatrix} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-3/2} \end{bmatrix}$$

for the trend shift model.

The GLS biases  $(\tilde{\beta}(\lambda) - \beta(\lambda))$  can be expressed as

$$\phi_T(\tilde{\beta}(\lambda) - \beta(\lambda)) = \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t(\lambda) \tilde{\mathbf{g}}_t(\lambda)' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t(\lambda) \tilde{u}_t(\lambda) \right). \quad (3.78)$$

The detrended processes  $y_t^{\tau\mu}(\lambda)$  and  $y_t^{\tau\delta}(\lambda)$  can be obtained from

$$y_t^{\tau\mu}(\lambda) = u_t^\mu(\lambda) - \left( \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\mu(\lambda) \tilde{\mathbf{g}}_t^\mu(\lambda)' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\mu(\lambda) \tilde{u}_t^\mu(\lambda) \right) \right)' \tilde{\mathbf{g}}_t^\mu(\lambda) \quad (3.79)$$

for the level shift model, and

$$y_t^{\tau\delta}(\lambda) = u_t^\delta(\lambda) - \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\delta(\lambda) \tilde{\mathbf{g}}_t^\delta(\lambda)' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{g}}_t^\delta(\lambda) \tilde{u}_t^\delta(\lambda) \right)' \tilde{\mathbf{g}}_t^\delta(\lambda) \quad (3.80)$$

for the trend shift model.

As in the previous sections, we have the following result:

$$T^{1/2} \phi_T(\tilde{\beta}(\lambda) - \beta(\lambda)) \implies \left( \int_0^1 \tilde{\mathbf{g}}(r, \lambda) \tilde{\mathbf{g}}(r, \lambda)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{g}}(r, \lambda) dr \right). \quad (3.81)$$

For a one sided test of the null hypothesis  $H_0 : \rho = 1$ , let  $\lambda_{\inf}$  be the value of  $\lambda$  that gives the least favorable result to the null. Then, we define the test statistics as follows:

$$\Gamma_1^*(\lambda_{\inf}) = \inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda), \quad (3.82)$$

$$\Gamma_2^*(\lambda_{\inf}) = \inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda), \quad (3.83)$$

where  $\Lambda$  is a closed subset of  $(0, 1)$ .

**Theorem 3.5.** *Let  $\{u_t\}$  be a stochastic process defined in (1.2) satisfying the Condition 1.1. Assume that  $\{v_t\}$  satisfy the conditions in Assumption 1.1. Let  $d_t$  be as defined in (3.75) with  $D_t(\lambda) = \{0|t \leq \lambda T, 1|t > \lambda T\}$  for a level shift and  $D_t(\lambda) = \{0|t \leq \lambda T, t - \lambda T|t > \lambda T\}$  for a trend shift. Let  $\lambda_{\inf}$  be the value of the  $\lambda \in \Lambda$  that minimizes the test when a small value of the statistic leads to the rejection of the null  $H_0 : \rho = 1$ , and assume that  $\rho = 1 + c/T$ , then,*

$$\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda) \implies \inf_{\lambda \in \Lambda} \left( c + \left( \int_0^1 B_c^{\tau\mu}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{-1} \right), \quad (3.84)$$

$$\begin{aligned} \inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( \left( \int_0^1 B_c^{\tau\mu}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{-1} + c \right) \\ &\times \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{1/2} \end{aligned} \quad (3.85)$$

for the level shift model, and

$$\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda) \implies \inf_{\lambda \in \Lambda} \left( c + \left( \int_0^1 B_c^{\tau_\delta}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau_\delta}(r, \lambda)^2 dr \right)^{-1} \right), \quad (3.86)$$

$$\begin{aligned} \inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( \left( \left( \int_0^1 B_c^{\tau_\delta}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau_\delta}(r, \lambda)^2 dr \right)^{-1} + c \right) \right. \\ &\quad \left. \times \left( \int_0^1 B_c^{\tau_\delta}(r, \lambda)^2 dr \right)^{1/2} \right) \end{aligned} \quad (3.87)$$

for the trend shift model, where

$$\tilde{B}_c^{\tau_\mu}(r) = B_c(r, \lambda) - \left( \left( \int_0^1 \tilde{\mathbf{g}}^\mu(r, \lambda) \tilde{\mathbf{g}}^\mu(r, \lambda)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{g}}^\mu(r, \lambda) dr \right) \right)' \tilde{\mathbf{g}}^\mu(r, \lambda), \quad (3.88)$$

$$\tilde{B}_c^{\tau_\delta}(r) = B_c(r, \lambda) - \left( \left( \int_0^1 \tilde{\mathbf{g}}^\delta(r, \lambda) \tilde{\mathbf{g}}^\delta(r, \lambda)' dr \right)^{-1} \left( \int_0^1 B_c(r) \tilde{\mathbf{g}}^\delta(r, \lambda) dr \right) \right)' \tilde{\mathbf{g}}^\delta(r, \lambda). \quad (3.89)$$

*Proof.* See Appendix. □

## 4 Finite Sample Analysis

To investigate the finite sample properties and sensitivity of the estimators to the shift fraction  $\lambda$  and different sample sizes we designed some Monte Carlo experiments under several configurations for both level and trend shift models.

For the level shift model 10,000 replications of a series  $\{y_t\}$  are generated by

$$y_t = \mu_1 + \delta t + (\mu_2 - \mu_1)D_t + u_t, \quad (4.90a)$$

$$u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \mathbf{N}(0, \sigma_\varepsilon^2), \quad (4.90b)$$

$$D_t = \{0 | t \leq T_B, 1 | t > T_B\}. \quad (4.90c)$$

Similarly, the trend shift model is examined by generating 10,000 replications of a series  $\{y_t\}$  according to

$$y_t = \mu + \delta_1 t + (\delta_2 - \delta_1)D_t + u_t, \quad (4.91a)$$

$$u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \mathbf{N}(0, \sigma_\varepsilon^2), \quad (4.91b)$$

$$D_t = \{0 | t \leq T_B, t - T_B | t > T_B\}. \quad (4.91c)$$

We designed two types of experiments. Experiment 1 is designed to analyze the sensitivity of the estimators to the shift fraction  $\lambda$  for a given sample size  $T$ . Experiment 2 is designed to analyze the finite sample properties of the estimators by varying  $T$  and keeping  $\lambda$  constant. These experiments are parameterized as follows :

Experiment 1:

$$\begin{aligned}\mu &= \mu_1 = 0, & \delta &= \delta_1 = 1, & T &= 100, & \rho &= 1, \\ \mu_2 &= -1, & \delta_2 &= .8, \\ \lambda &= \{.15, .50, .85\}.\end{aligned}$$

Experiment 2:

$$\begin{aligned}\mu &= \mu_1 = 0, & \delta &= \delta_1 = 1, & \lambda &= .50, & \rho &= 1, \\ \mu_2 &= -1, & \delta_2 &= .8, \\ T &= \{50, 100, 250, 500\}.\end{aligned}$$

For each configuration estimates of the parameters are obtained from the following regressions:

$$y_t = \hat{\mu} + \hat{\delta}t + \hat{\theta}D_t + \hat{u}_t, \quad (4.92)$$

$$\hat{u}_t = \hat{\rho}\hat{u}_{t-1} + \hat{e}_t. \quad (4.93)$$

Frequency plots of the Monte Carlo estimates are given in Figure 1, Figure 2, Figure 3, and Figure 4 for  $\hat{\mu}$ ,  $\hat{\delta}$ ,  $\hat{\theta}$ , and  $\hat{\rho}$ , respectively. Several observations are in stand from these frequency plots.

None of the estimators are sensitive to the break fraction  $\lambda$  in case of a level shift. The estimators  $\hat{\mu}$  and  $\hat{\rho}$  are also insensitive to  $\lambda$  when there is a trend shift. However,  $\hat{\delta}$  and  $\hat{\theta}$  are sensitive to  $\lambda$  in terms of the variance when there is a trend shift. This is an expected result since for small  $\lambda$  these two estimators would be highly correlated. All estimators expect  $\hat{\rho}$  has a symmetric distribution for both level and trend shift models and across all  $\lambda$  and  $T$ . All estimators except  $\hat{\rho}$  show consistency and efficiency and converge to their true values as the sample size increases. As expected there is an efficiency gain from the increased sample size for all estimators.

The estimator  $\hat{\mu}$  is not sensitive to the type of shift. It has the same distribution for both level and trend shift models. The estimators  $\hat{\delta}$  and  $\hat{\theta}$  show sensitivity to the type of shift. They have some size sensitivity to  $\lambda$  when there is a trend shift, but they are insensitive to  $\lambda$  when the shift is a level shift. This confirms our analytical results: The estimators will depend on  $\lambda$  for a trend shift, but are asymptotically independent of  $\lambda$  for a level shift. It is surprising that the independence from  $\lambda$  is preserved even in sample sizes as small as 50 for the level shift model.

The estimator  $\hat{\rho}$  shows very little sensitivity to the shift fraction  $\lambda$  when there is a trend shift. However, it is invariant to the shift fraction  $\lambda$  when there is a level shift. As it is well









known,  $\hat{\rho}$  shows a large bias in small samples and right skewed. It is interesting that a sample size as large as 500 is needed for  $\hat{\rho}$  to get close to its true value. However, a small bias still remains even with a sample size of 500 or more.

## 5 Asymptotic and Finite Sample Critical Values of the Tests

In this section, we tabulate the asymptotic and finite sample critical values of the test statistics  $\Gamma_1^*$ ,  $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$ ,  $\Gamma_2^*$ , and  $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$ .

In order to tabulate the critical values the value of  $\bar{c}$  should be determined. Elliot et al. (1992) derived asymptotic power function of a POI test under local-to-unity asymptotics. When the deterministic component is composed of a constant and a linear trend they show that the local asymptotic power function is given by  $\pi^{\tau\mu}(c, \bar{c}) = P[\Psi^{\tau\mu}(c, \bar{c}) < b^{\tau\mu}(\bar{c})]$ , where  $b^{\tau\mu}(\bar{c})$  is a constant and  $\Psi^{\tau\mu}(c, \bar{c})$  is given by

$$\Psi^{\tau\mu}(c, \bar{c}) = \bar{c}^2 \int_0^1 B_c^{\tau\mu}(r) dr + (1 - \bar{c}) B_c^{\tau\mu}(1)^2, \quad (5.94)$$

where

$$B_c^{\tau\mu}(r) = B_c(r) - r \left( 3 \frac{\bar{c}^2}{\bar{c}^2 - 3\bar{c} + 3} \int_0^1 r B_c(r) dr + 3 \frac{1 - \bar{c}}{\bar{c}^2 - 3\bar{c} + 3} B_c(1) \right).$$

The power envelope  $\Pi^{\tau\mu}(c) = \pi^{\tau\mu}(c, c)$  defines an upper bound for the local asymptotic power function. This is an upper bound for any one sided test of  $H_0 : \rho = 1$ . For the level shift model the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  have the same asymptotic power function as the one derived by Elliot et al. (1992). For the trend shift model the local asymptotic power function is given by  $\pi^{\tau\delta}(c, \bar{c}) = P[\Psi^{\tau\delta}(c, \bar{c}) < b^{\tau\delta}(\bar{c})]$ , where  $b^{\tau\delta}(\bar{c})$  is a constant.  $\Psi^{\tau\delta}(c, \bar{c})$  is given by

$$\Psi^{\tau\delta}(c, \bar{c}) = \bar{c}^2 \int_0^1 B_c^{\tau\delta}(r) dr + (1 - \bar{c}) B_c^{\tau\delta}(1)^2, \quad (5.95)$$

where  $B_c^{\tau\delta}(r)$  is defined in (2.27). Thus, the relevant power envelope for the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  in the presence of a trend shift is given by  $\Pi^{\tau\delta}(c) = \pi^{\tau\delta}(c, c)$ .

The choice of  $\bar{c}$  is made in such a way that the asymptotic power functions of the test statistics are tangent to the power envelope. The value  $\bar{c} = -13.5$  determined by Elliot et al. (1992) is used for the level shift model. For the trend shift model a grid search over  $\bar{c} = [-8, -30]$  reveals that the tangency also occurs at approximately  $\bar{c} = -13.5$ . Thus, we use  $\bar{c} = -13.5$  for both models.

Finite sample critical values of the  $\Gamma_1^*$  and  $\Gamma_2^*$  are obtained from a Gaussian random walk using the finite sample approximations of the corresponding test statistics under the alternative

Table 1: Critical Values of the  $\Gamma_1^*$  and  $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$  for the Level Shift Model

T	1%	5%	10%	25%	75%	90%	95%	99%
A. $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$								
	-36.94	-29.76	-25.88	-20.11	-10.63	-7.48	-6.17	-4.14
B. $\Gamma_1^*$ for a fixed $\lambda$								
50	-23.31	-17.57	-14.93	-11.07	-5.09	-3.37	-2.56	-1.35
100	-23.81	-17.37	-14.57	-10.41	-4.32	-2.70	-1.97	-0.80
200	-23.51	-17.15	-14.06	-9.89	-3.87	-2.29	-1.60	-0.59
500	-23.35	-16.99	-13.94	-9.52	-3.49	-1.98	-1.31	-0.33
$\infty$	-23.39	-16.73	-13.59	-9.30	-3.39	-1.89	-1.24	-0.30

Table 2: Critical Values of the  $\Gamma_2^*$  and  $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$  for the Level Shift Model

T	1%	5%	10%	25%	75%	90%	95%	99%
A. $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$								
	-4.35	-3.90	-3.62	-3.18	-2.28	-1.89	-1.71	-1.38
B. $\Gamma_2^*$ for a fixed $\lambda$								
50	-3.84	-3.21	-2.90	-2.44	-1.57	-1.21	-0.98	-0.56
100	-3.62	-3.03	-2.74	-2.29	-1.40	-1.03	-0.81	-0.38
200	-3.49	-2.95	-2.65	-2.20	-1.30	-0.93	-0.70	-0.30
500	-3.43	-2.90	-2.62	-2.14	-1.23	-0.85	-0.61	-0.18
$\infty$	-3.41	-2.87	-2.57	-2.11	-1.20	-0.82	-0.58	-0.16

Table 3: Critical Values of the  $\Gamma_1^*$  and  $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$  for the Trend Shift Model

$\lambda$	T	1%	5%	10%	25%	75%	90%	95%	99%
A. $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$									
		-38.48	-29.78	-25.99	-19.86	-11.47	-8.93	-7.54	-5.38
B. $\Gamma_1^*$ for a fixed $\lambda$									
.50	50	-28.60	-22.75	-19.87	-15.44	-8.33	-6.10	-5.10	-3.56
.40, .60	50	-28.73	-22.67	-19.69	-15.34	-8.09	-5.87	-4.77	-3.14
.30, .70	50	-28.45	-22.32	-19.34	-14.93	-7.54	-5.28	-4.18	-2.61
.20, .80	50	-27.61	-21.57	-18.63	-14.07	-6.72	-4.56	-3.58	-2.08
.10, .90	50	-26.12	-20.01	-17.09	-12.81	-5.94	-3.93	-2.99	-1.71
.50	100	-30.25	-23.28	-20.05	-15.27	-7.74	-5.57	-4.56	-3.04
.40, .60	100	-30.23	-23.43	-20.07	-15.34	-7.68	-5.46	-4.38	-2.79
.30, .70	100	-30.08	-22.99	-19.86	-15.07	-7.38	-5.08	-4.08	-2.52
.20, .80	100	-29.51	-22.33	-18.99	-14.29	-6.73	-4.54	-3.52	-2.09
.10, .90	100	-27.85	-21.41	-18.03	-13.20	-5.90	-3.86	-2.92	-1.69
.50	200	-30.79	-23.49	-20.00	-15.08	-7.41	-5.29	-4.27	-2.81
.40, .60	200	-31.02	-23.70	-20.17	-15.09	-7.47	-5.23	-4.22	-2.80
.30, .70	200	-30.23	-23.41	-20.02	-15.11	-7.33	-5.03	-3.95	-2.30
.20, .80	200	-30.13	-22.73	-19.38	-14.44	-6.71	-4.46	-3.45	-2.02
.10, .90	200	-29.09	-21.51	-18.23	-13.37	-6.02	-3.98	-2.98	-1.59
.50	500	-31.06	-23.20	-19.65	-14.88	-7.20	-5.03	-4.04	-2.62
.40, .60	500	-32.08	-23.84	-20.17	-15.16	-7.39	-5.14	-4.12	-2.67
.30, .70	500	-30.89	-23.41	-19.92	-14.92	-7.18	-4.95	-3.99	-2.47
.20, .80	500	-31.59	-23.21	-19.65	-14.42	-6.66	-4.46	-3.45	-2.03
.10, .90	500	-29.42	-21.98	-18.54	-13.55	-5.98	-3.88	-2.96	-1.63
.50	$\infty$	-31.19	-23.40	-19.72	-14.77	-7.02	-4.88	-3.88	-2.65
.40, .60	$\infty$	-31.83	-23.73	-20.18	-15.10	-7.27	-5.07	-4.04	-2.62
.30, .70	$\infty$	-31.67	-23.73	-20.17	-14.96	-7.13	-4.90	-3.88	-2.35
.20, .80	$\infty$	-31.13	-23.31	-19.63	-14.51	-6.73	-4.51	-3.44	-1.97
.10, .90	$\infty$	-29.94	-22.19	-18.52	-13.55	-5.95	-3.89	-2.92	-1.56

Table 4: Critical Values of the  $\Gamma_2^*$  and  $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$  for the Trend Shift Model

$\lambda$	T	1%	5%	10%	25%	75%	90%	95%	99%
<i>A. <math>\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)</math></i>									
		-4.45	-3.88	-3.63	-3.16	-2.35	-2.05	-1.92	-1.57
<i>B. <math>\Gamma_2^*</math> for a fixed <math>\lambda</math></i>									
.50	50	-4.42	-3.78	-3.46	-2.97	-2.08	-1.74	-1.56	-1.23
.40, .60	50	-4.42	-3.76	-3.43	-2.94	-2.02	-1.66	-1.46	-1.09
.30, .70	50	-4.38	-3.71	-3.39	-2.89	-1.92	-1.55	-1.33	-0.94
.20, .80	50	-4.28	-3.63	-3.31	-2.79	-1.79	-1.41	-1.19	-0.79
.10, .90	50	-4.12	-3.46	-3.13	-2.62	-1.68	-1.29	-1.06	-0.67
.50	100	-4.19	-3.59	-3.30	-2.83	-1.95	-1.63	-1.45	-1.13
.40, .60	100	-4.19	-3.60	-3.29	-2.83	-1.93	-1.58	-1.39	-1.03
.30, .70	100	-4.15	-3.56	-3.27	-2.80	-1.88	-1.51	-1.31	-0.93
.20, .80	100	-4.10	-3.50	-3.19	-2.72	-1.78	-1.40	-1.19	-0.79
.10, .90	100	-3.98	-3.41	-3.09	-2.60	-1.65	-1.27	-1.05	-0.67
.50	200	-4.06	-3.50	-3.22	-2.77	-1.89	-1.57	-1.40	-1.08
.40, .60	200	-4.07	-3.51	-3.22	-2.77	-1.89	-1.55	-1.36	-1.03
.30, .70	200	-4.01	-3.49	-3.21	-2.76	-1.86	-1.50	-1.29	-0.89
.20, .80	200	-4.00	-3.43	-3.15	-2.69	-1.76	-1.39	-1.18	-0.79
.10, .90	200	-3.93	-3.33	-3.05	-2.58	-1.66	-1.29	-1.07	-0.66
.50	500	-3.99	-3.42	-3.14	-2.72	-1.86	-1.53	-1.36	-1.06
.40, .60	500	-4.06	-3.47	-3.18	-2.74	-1.87	-1.53	-1.34	-1.01
.30, .70	500	-3.97	-3.44	-3.16	-2.72	-1.83	-1.49	-1.30	-0.94
.20, .80	500	-4.02	-3.41	-3.13	-2.67	-1.76	-1.39	-1.19	-0.80
.10, .90	500	-3.87	-3.32	-3.04	-2.58	-1.66	-1.28	-1.07	-0.67
.50	$\infty$	-3.96	-3.42	-3.13	-2.70	-1.83	-1.50	-1.33	-1.04
.40, .60	$\infty$	-4.01	-3.44	-3.17	-2.73	-1.86	-1.52	-1.33	-1.01
.30, .70	$\infty$	-3.99	-3.44	-3.16	-2.71	-1.83	-1.47	-1.28	-0.91
.20, .80	$\infty$	-3.96	-3.40	-3.12	-2.66	-1.77	-1.40	-1.18	-0.80
.10, .90	$\infty$	-3.88	-3.32	-3.02	-2.57	-1.65	-1.27	-1.06	-0.65

hypothesis with 20,000 Monte Carlo replications. Asymptotic critical values of the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests are obtained with discretized realizations of the corresponding asymptotic distributions under the null ( $c = 0$ ), using 20,000 Monte Carlo replications. We provide only asymptotic critical values of the test statistics  $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$  and  $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$ . We used the method described in Zivot and Andrews (1992) to approximate the asymptotic distributions of the test statistics  $\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda)$  and  $\inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda)$ . The approximations are obtained from 1,000 normal realizations with 10,000 Monte Carlo replications.

## 6 Finite Sample Power of the Tests

This section investigates the finite sample size and power of the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  we studied in previous sections of this chapter. The power and size properties of unit root tests in models without structural change have been investigated and it has been shown that the power of the Dickey-Fuller type unit root test statistics are low (see Schwert (1989)). Unit root tests such as the test statistics developed by Elliot et al. (1992) have been shown to have good power properties against the TS alternatives they are designed for. Unfortunately, the size and power properties of unit root tests constructed against the STS alternatives are not known except the test statistics studied by Banerjee et al. (1992). Perron (1989), Perron (1990b), Perron and Vogelsang (1992), Zivot and Andrews (1992), and Park and Sung (1994) did not investigate the power and size properties of the test statistics they developed. We investigated the power of unit root tests against the misspecified TS alternatives when the true DGP's were STS models by extensive Monte Carlo simulations in Balcilar (1996). In addition to investigating the power and size properties of the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  we also investigate the power and size properties of the unit root test statistics of Perron (1989) by Monte Carlo simulations. We limit these Monte Carlo simulations to the models with a single predetermined shift point. Computational cost of the simulations required to investigate the power and size of the tests with endogenous shift point assumption makes it infeasible for this study. The power and size of sequential unit root test statistics constructed against STS alternatives were studied by Banerjee et al. (1992). They found that these sequential unit root test statistics have good power properties against the alternatives they are designed for. The sequential test statistics we examined in §3.3 should have approximately the same size and power properties. We leave the investigation of the power and size properties of these sequential unit root tests for future research.

To investigate the finite sample size and power properties of the unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$ , and Perron's test statistics  $t_{\tilde{\alpha}}$  and  $T(\tilde{\alpha} - 1)$  we designed two separate Monte Carlo

experiments under variety of parameter configurations for the level and trend shift models.

For the level shift model 10,000 replications of a series  $\{y_t\}$  are generated by

$$y_t = \mu_1 + \delta t + (\mu_2 - \mu_1)D_t + u_t, \quad (6.96a)$$

$$u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \mathbf{N}(0, 1), \quad (6.96b)$$

$$D_t = \{0|t \leq T_B, 1|t > T_B\}. \quad (6.96c)$$

Similarly, the trend shift model is examined by generating 10,000 replications of a series  $\{y_t\}$  according to

$$y_t = \mu + \delta_1 t + (\delta_2 - \delta_1)D_t + u_t, \quad (6.97a)$$

$$u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \mathbf{N}(0, 1), \quad (6.97b)$$

$$D_t = \{0|t \leq T_B, t - T_B|t > T_B\}. \quad (6.97c)$$

The parameter configurations used in these simulations are as follows:

$$T = 100, 200,$$

$$\delta = \delta_1 = \mu = \mu_1 = 1.00, \quad T_B/T = .50,$$

$$\rho = \{.20, .50, .80, 1.00\},$$

$$\delta_2 = \{1, .7, .3, 0\},$$

$$\mu_2 = \{0, -5, -15, -30\}.$$

To compute the test statistics first we detrend the data by GLS and obtain the test statistics from a Said-Dickey type regression. Let  $\mathbf{z}_t = (1, t, D_t)'$  and define the transformed variables  $\tilde{\mathbf{z}}_t$  and  $\tilde{y}_t$ . These transformed variables are obtained as defined in (1.12). From these detrended series we obtain the detrended series in the following way

$$\tilde{u}_t = y_t - \tilde{\beta}' \tilde{\mathbf{z}}_t, \quad (6.98)$$

where

$$\tilde{\beta} = \left( \sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{\mathbf{z}}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{y}_t \right). \quad (6.99)$$

The unit root test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are then computed from the following Said-Dickey regression:

$$\tilde{u}_t = \tilde{\alpha} \tilde{u}_{t-1} + \sum_{j=1}^k \tilde{a}_j \Delta \tilde{u}_{t-j} + \tilde{e}_t. \quad (6.100)$$

The test statistic  $\Gamma_1^*$  is computed from the estimate  $\tilde{\alpha}$ , that is,  $\Gamma_1^* = T(\tilde{\alpha} - 1)$ . The test statistic  $\Gamma_2^*$  is the usual  $t$  statistic corresponding to parameter  $\tilde{\alpha}$  in the regression (6.100). Perron's test

statistics are obtained in the same way by replacing  $\tilde{u}_t$  in (6.100) by  $\hat{u}_t$  which is obtained from

$$\hat{u}_t = y_t - \hat{\beta}' z_t, \quad (6.101)$$

where

$$\hat{\beta} = \left( \sum_{t=1}^T z_t z_t' \right)^{-1} \left( \sum_{t=1}^T z_t y_t \right). \quad (6.102)$$

The lag truncation parameter  $k$  in (6.100) should be determined from the data, although, some researchers prefer to prefix  $k$  at a controlled rate. The additional regressors  $\Delta \tilde{u}_{t-j}$  are included to remove the nuisance parameter dependencies of the test statistics due to the auto-correlations in the data. However, it has been found that if  $k$  is chosen to be too high the power of the unit root tests will be adversely affected. There are several methods to determine the lag truncation parameter  $k$  that can be employed in practice, among these we use the BIC which is a common method in applied studies. We choose the value of  $k$  such that

$$BIC(k) = T \ln(\sigma_e^2) + (k+1) + \ln(T)(k+1) \quad (6.103)$$

is minimized subject to  $k < 23$ , that is, we restrict the maximum of  $k$  to be 23. However, this restriction was never binding.

The results of the Monte Carlo simulations for the  $\Gamma_1^*$  and  $\Gamma_2^*$  test statistics when there is a level shift in the trend function are reported in Table 5. These simulations are based on 10,000 replications for sample sizes 100 and 200. For all simulations, we compute the power of the tests at the 5 percent level and report the nominal, not size adjusted, power.

The results in Table 5 show that the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  have a power of 1 for moderate values of  $\rho$ . Test statistics retain a good power for values of  $\rho$  as large as .80, which is a value unit root tests most commonly fail to reject the null. There is a very small reduction in the power of the tests, from .99 to .97, if we raise  $\rho$  from .50 to .80 for a sample size of 200. However, the reduction in the power of the tests when  $\rho = .80$  is more significant when the sample size is 100. The powers of the test statistics for  $\rho = .80$  is .77 in this case. However, a power of .77 is a significant improvement over the power of the other unit root tests. We should also note that the sample sizes as small as 100 are believed to be fairly small for unit root tests. A significant finding from the results of these simulations is that the powers of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are not sensitive to the size of the level shift which shows that the GLS detrending procedure efficiently removes the deterministic component even at sample sizes as small as 100. From the last two columns of Table 5 we note that the test statistics have reasonably good size properties. The difference between the empirical and theoretical sizes is at most in the order of .02 which is negligible.

Table 5: Size and Power of the  $\Gamma_1^*$  and  $\Gamma_2^*$  Tests for the Level Shift Model

$\rho$	.20		.50		.80		1.00		
	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	
T=200									
$\mu_2$									
0	P	1.00	0.95	1.00	0.99	0.97	0.97	0.04	0.03
	M	-112.88	-8.34	-80.37	-6.97	-36.19	-4.39	-6.98	-1.72
	V	1567.96	8.21	411.03	1.67	107.85	0.52	22.51	0.44
-5	P	1.00	0.95	1.00	0.99	0.97	0.97	0.04	0.03
	M	-113.68	-8.38	-79.97	-6.93	-35.94	-4.38	-6.97	-1.72
	V	1554.53	8.20	423.32	1.74	104.64	0.51	22.63	0.43
-15	P	1.00	0.95	1.00	0.99	0.98	0.97	0.03	0.03
	M	-113.10	-8.34	-79.89	-6.94	-35.98	-4.38	-6.90	-1.71
	V	1564.02	8.21	412.26	1.68	103.51	0.51	22.46	0.43
-30	P	1.00	0.95	1.00	0.99	0.97	0.97	0.03	0.03
	M	-113.60	-8.38	-80.07	-6.95	-35.96	-4.38	-6.88	-1.71
	V	1556.25	8.17	410.53	1.67	105.10	0.51	21.83	0.42
T=100									
0	P	1.00	0.99	1.00	0.99	0.77	0.77	0.05	0.06
	M	-68.45	-7.10	-46.46	-5.43	-22.60	-3.50	-8.10	-1.89
	V	215.38	1.89	100.29	0.66	46.32	0.38	25.05	0.48
-5	P	1.00	0.99	1.00	0.99	0.75	0.76	0.06	0.06
	M	-68.63	-7.11	-46.40	-5.43	-22.43	-3.48	-8.19	-1.90
	V	214.12	1.88	99.07	0.65	48.41	0.39	25.30	0.48
-15	P	1.00	0.99	1.00	0.99	0.76	0.77	0.05	0.06
	M	-68.39	-7.09	-46.42	-5.43	-22.46	-3.49	-8.09	-1.89
	V	221.88	1.94	99.09	0.65	45.68	0.37	25.08	0.47
-30	P	1.00	0.99	1.00	0.99	0.76	0.76	0.06	0.06
	M	-68.54	-7.10	-46.46	-5.43	-22.60	-3.50	-8.15	-1.90
	V	217.16	1.90	102.74	0.68	49.04	0.40	24.75	0.47

Notes: P shows the nominal power of the test at 5% level, M is the Monte Carlo mean, and V is the Monte Carlo variance. All simulations are based on 10,000 replications.

Table 6: Size and Power of the  $\Gamma_1^*$  and  $\Gamma_2^*$  Tests for the Trend Shift Model

$\rho$	.20		.50		.80		1.00		
	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	$\Gamma_1^*$	$\Gamma_2^*$	
$\delta_2$		T=200							
1	P	1.00	0.99	1.00	1.00	0.98	0.98	0.05	0.05
	M	-142.71	-10.33	-93.39	-7.73	-41.75	-4.78	-11.85	-2.36
	V	728.53	3.96	262.40	0.95	96.94	0.41	36.48	0.42
.7	P	1.00	0.99	1.00	1.00	0.98	0.98	0.05	0.05
	M	-142.65	-10.32	-93.52	-7.74	-41.75	-4.78	-11.74	-2.35
	V	733.13	3.96	260.32	0.93	97.44	0.42	36.01	0.41
.3	P	1.00	0.99	1.00	1.00	0.98	0.98	0.05	0.05
	M	-143.31	-10.35	-93.46	-7.74	-41.90	-4.79	-11.82	-2.35
	V	707.13	3.85	261.58	0.93	98.70	0.42	37.43	0.42
0	P	1.00	0.99	1.00	1.00	0.98	0.98	0.05	0.05
	M	-143.44	-10.37	-93.38	-7.74	-41.96	-4.79	-11.82	-2.36
	V	721.02	3.85	265.28	0.94	98.05	0.42	36.12	0.41
		T=100							
1	P	1.00	0.99	1.00	0.99	0.55	0.56	0.05	0.05
	M	-75.16	-7.67	-49.95	-5.71	-24.87	-3.71	-12.03	-2.43
	V	162.74	1.44	96.41	0.62	50.67	0.39	34.60	0.43
.7	P	1.00	0.99	1.00	0.99	0.54	0.54	0.05	0.05
	M	-75.09	-7.66	-49.91	-5.71	-24.76	-3.70	-12.08	-2.43
	V	167.09	1.48	97.94	0.62	53.74	0.41	35.10	0.44
.3	P	1.00	0.99	1.00	0.99	0.53	0.54	0.05	0.05
	M	-75.07	-7.66	-49.92	-5.71	-24.75	-3.70	-11.96	-2.42
	V	168.46	1.49	94.93	0.61	51.64	0.39	35.01	0.44
0	P	1.00	0.99	1.00	0.99	0.54	0.55	0.05	0.05
	M	-75.31	-7.69	-50.06	-5.72	-24.93	-3.72	-12.05	-2.43
	V	160.55	1.40	98.32	0.63	54.02	0.42	34.00	0.43

Notes: See notes to Table 5.

Table 7: Size and Power of the Perron's Tests

$\rho$	.20		.50		.80		1.00		
	$T(\tilde{\alpha} - 1)$	$t_{\tilde{\alpha}}$							
<i>A. Level Shift Model</i>									
$\mu_2$									
0	P	1.00	1.00	1.00	1.00	0.60	0.62	0.05	0.06
	M	-83.60	-8.44	-55.45	-6.16	-27.93	-4.00	-12.65	-2.53
	V	104.82	0.80	89.32	0.54	57.83	0.44	44.63	0.63
-5	P	1.00	1.00	1.00	1.00	0.59	0.60	0.05	0.07
	M	-83.68	-8.43	-55.47	-6.16	-27.81	-3.99	-12.84	-2.55
	V	114.01	0.82	90.21	0.54	60.22	0.45	46.15	0.65
-15	P	1.00	1.00	1.00	1.00	0.60	0.62	0.05	0.07
	M	-83.63	-8.43	-55.43	-6.16	-27.91	-3.99	-12.67	-2.54
	V	109.81	0.82	89.38	0.54	59.51	0.45	45.06	0.63
-30	P	1.00	1.00	1.00	1.00	0.60	0.62	0.05	0.06
	M	-83.56	-8.43	-55.58	-6.17	-28.03	-4.00	-12.67	-2.53
	V	109.50	0.82	91.18	0.55	60.89	0.46	44.20	0.63
<i>B. Trend Shift Model</i>									
$\delta_2$									
1	P	1.00	1.00	1.00	1.00	0.38	0.51	0.03	0.06
	M	-83.70	-8.45	-55.45	-6.16	-28.09	-4.01	-14.46	-2.74
	V	102.92	0.78	89.17	0.54	58.12	0.44	43.37	0.56
.7	P	1.00	1.00	1.00	1.00	0.37	0.49	0.03	0.06
	M	-83.71	-8.43	-55.53	-6.17	-27.95	-4.00	-14.56	-2.74
	V	114.49	0.82	92.51	0.55	60.91	0.45	45.24	0.59
.3	P	1.00	1.00	1.00	1.00	0.37	0.50	0.03	0.06
	M	-83.71	-8.44	-55.44	-6.16	-28.02	-4.00	-14.40	-2.73
	V	109.38	0.82	88.74	0.54	58.94	0.44	44.36	0.58
0	P	1.00	1.00	1.00	1.00	0.38	0.51	0.03	0.06
	M	-83.66	-8.43	-55.65	-6.18	-28.23	-4.02	-14.48	-2.73
	V	110.21	0.82	91.70	0.55	59.94	0.45	43.69	0.57

Notes: See notes to Table 5.

The nominal powers of the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests computed at the 5 percent level when there is a trend shift are reported in Table 6. When the sample size is 200, both test statistics have an exact size and a power of 1 across all values of  $\delta_2$  and  $\rho$  except when  $\rho = .80$  for which only the power of the tests drop to .98, but the tests still have exact size. When the sample size is reduced to 100 only the power of the tests when  $\rho = .80$  are affected, that is, the power of the tests drop approximately to .55 across all values of  $\delta_2$ . The sizes of the tests are still exact even for a sample size as small as 100. The worsening of the power when  $\rho$  is high is an expected result for small samples. However, we point out that this power is still a significant improvement over the power of previously developed unit root tests by Perron (1989).

The nominal powers of Perron's test statistics computed at the 5 percent level for a sample

size of 100 are reported in Table 7. The powers of Perron's tests are equivalent to the powers of the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests for small values of  $\rho$ . However, the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests perform significantly better than Perron's tests when  $\rho = .80$  which is the vicinity of the values that the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests are designed to have good power. Particularly, the power gain over the  $T(\tilde{\alpha} - 1)$  test is significant. The  $\Gamma_1^*$  test has 28 percent higher power when there is a level shift and 47 percent higher power when there is a trend shift. Surprisingly, we are able to improve the power significantly for the trend shift model. Although, the power gain from using the  $\Gamma_2^*$  test rather than the  $t_{\tilde{\alpha}}$  test of Perron is not as high as using the  $\Gamma_1^*$  test rather than the  $T(\tilde{\alpha} - 1)$  test of Perron, the improvement is still significant both for the level and trend shift models. The power improvement of  $\Gamma_2^*$  test over the  $t_{\tilde{\alpha}}$  test is around 10 percent for the trend shift model and 25 percent for the level shift model. As the  $\Gamma_1^*$  and  $\Gamma_2^*$  tests, Perron tests too do not show sensitivity to the size of either the level or the trend shifts. As for the size of the tests, the results reveal that the  $\Gamma_2^*$  outperforms the  $t_{\tilde{\alpha}}$  tests in terms of the size stability. The test statistics  $\Gamma_1^*$  and the  $T(\tilde{\alpha} - 1)$  have equivalent size sensitivities. However, the size sensitivities of all the tests we examined in this section are negligible.

## 7 Summary and Conclusion

If a time series undergoes a structural change either in the form of a change in the mean of the series, a level shift, or in the form of a change in the growth rate of the series, a trend shift, unit root tests incorrectly fail to reject the null hypothesis of a unit root when the effects of structural change is not taken into account. This study develops two unit root test statistics with good power and size properties against alternatives that incorporate the effects of structural changes. These tests are approximately uniformly most powerful invariant. The asymptotic distributions of the test statistics are obtained by generalizing the method developed by Elliot et al. (1992) to models with structural change. The test statistics are constructed from the locally generalized least squares detrended data. The generalized least squares estimates are obtained by transforming the data according to the local alternatives. We derive the asymptotic distributions of these test statistics using the local-to-unity asymptotic framework of Bobkoski (1983), Cavanagh (1985), Chan and Wei (1987), and Phillips (1987b). The local to-unity-asymptotic framework allows us to derive the asymptotic power functions of the test statistics. Contrary to the previously proposed unit root test statistics, the asymptotic distributions of these test statistics do not depend on the date of the structural change for the level shift model. This property makes the test statistics easily generalizable to more general models. Furthermore, we derive the asymptotic distributions of the test statistics under

multiple shifts and innovational outliers assumptions. The test statistics are also studied under the assumption that the date of the structural change is unknown. Thus, the test statistics are generalized to models in which the time of the structural change is treated to be endogenous. This generalization is achieved by sequentially computing the test statistics over the full sample. The asymptotic distributions of these sequential test statistics are also derived. We tabulate the finite sample and asymptotic percentiles of the test statistics by Monte Carlo integration. The finite sample properties of the estimators involved in computing the test statistics are also examined. We find that the estimators are invariant to the shift fraction for the level shift model even in small samples. We also find that only the estimator of the slope of the linear time trend is sensitive to the shift fraction for the trend shift model. Power properties of the test statistics are examined by extensive simulations. These simulations show that test statistics provide significant power gains over the similar unit root tests.

## Appendix: Proofs

### Proof of Lemma 1.1

To prove parts (g) and (h) of Lemma 1.1 let  $\rho = e^{c/T}$  and generate double array

$$u_{t,T} = e^{c/T} u_{t-1,T} + v_t, \quad t = 1, \dots, T, \quad u_{0T} = 0. \quad (\text{A.1})$$

Now define

$$B_c(r) = \int_0^r e^{(r-s)c} dB(s), \quad (\text{A.2})$$

where  $B(s)$  is a standard Brownian motion and  $B_c(r)$  is an Ornstein-Uhlenbeck process. Solving (A.1), we get

$$u_{t,T} = \sum_{j=1}^t e^{(t-j)c/T} v_j.$$

Let  $X_T(r)$  be defined by

$$X_T(r) = \frac{1}{\sigma\sqrt{T}} S_{j-1}, \quad (j-1)/T \leq r < j/T, \quad r \in [0, 1], \quad (\text{A.3})$$

where  $S_j = v_1 + \dots + v_j$ . Using  $v_j = \sigma\sqrt{T} \int_{(j-1)/T}^{j/T} dX_T(r)$ , as in Phillips (1987b), we derive

$$\begin{aligned} T^{-1/2} u_{[Tr]} &= \sigma \sum_{j=1}^{[Tr]} e^{c([Tr]-j)/T} \int_{(j-1)/T}^{j/T} dX_T(s) + O_p(T^{-1/2}) \\ &= \sigma \sum_{j=1}^{[Tr]} \int_{(j-1)/T}^{j/T} e^{c(r-s)} dX_T(s) + O_p(T^{-1/2}) \\ &= \sigma \int_0^r e^{c(r-s)} dX_T(s) + O_p(T^{-1/2}). \end{aligned} \quad (\text{A.4})$$

Note that

$$\begin{aligned} \int_0^r e^{c(r-s)} dX_T(s) &= e^{c(r-s)} dX_T(s) \Big|_0^r + ce^{cr} \int_0^r e^{-cs} X_T(s) ds \\ &\implies X_T(r) + c \int_0^r e^{c(r-s)} dB(s) \\ &= B_c(r), \end{aligned} \quad (\text{A.5})$$

where last two steps follow from (1.6), Herrndorf's weak convergence result, and continuous mapping theorem. Combining these results we establish that

$$T^{-1/2} u_{[Tr]} \implies \sigma B_c(r). \quad (\text{A.6})$$

To prove part (g) of Lemma 1.1, note that  $T^{-1/2} \sum_{t=2}^T \Delta u_t = u_T$ . Then, the proof follows from setting  $r = 1$  in (A.6).

To prove part (h) of Lemma 1.1, we use the following identity:

$$\begin{aligned} T^{-3/2} \sum_{t=2}^T t \Delta u_t &= T^{-3/2} \left( \sum_{t=2}^T t u_t - \sum_{t=2}^T (t-1+1) u_{t-1} \right) \\ &= T^{-1/2} u_T - T^{-3/2} u_1 - T^{-3/2} \sum_{t=2}^T u_{t-1}. \end{aligned} \tag{A.7}$$

Note that from Condition 1.1 and strong law of McLeish (1975)  $T^{-3/2} u_1 \xrightarrow{a.s.} 0$ . Then, the proof follows from parts (d) and (h) of Lemma 1.1, which is proven above, Herrndorf's weak convergence result, and continuous mapping theorem.

### Proof of Lemma 1.2

To prove part (a) of Lemma 1.2, we note that using  $T_B = \lambda T$  we obtain  $\lim_{T \rightarrow \infty} (T_B + i)/T = \lambda$  for  $i = 0, 1, 2$ . Then the result follows from setting  $r = \lambda$  in (A.6). To prove the rest of the results in Lemma 1.2 we note that for any  $J_t$  we can write  $\sum_{t=T_B+2}^T J_t = \sum_{t=1}^T J_t - \sum_{t=1}^{T_B+1} J_t$ . Then, applying the results from Lemma 1.1 to each part of this expression we obtain the results in parts (b)-(e) of Lemma 1.2.

### Proof of Theorem 2.1

The proof of Theorem 2.1 involves analytic calculation of the GLS bias terms

$$(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \left( \sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{\mathbf{z}}_t' \right)^{-1} \left( \sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{u}_t \right), \tag{A.8}$$

where  $(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = ((\tilde{\mu} - \mu), (\tilde{\delta} - \delta), (\tilde{\theta} - \theta))'$ ,  $\tilde{\mathbf{z}}_t = (1 - \bar{\rho}L)(1, t, D_t)'$ ,  $\tilde{u}_t = (1 - \bar{\rho})u_t$ , and  $D_t = \{0 | t \leq T_B, 1 | t > T_B\}$ .

**Fact .1.** Noting that under GLS transformation first observation is not transformed, we can easily verify that the elements of the moment matrix  $\sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{\mathbf{z}}_t'$  can be expressed as

$$\begin{aligned} \sum_{t=1}^T ((1 - \bar{\rho}L))^2 &= 1 + \frac{(T-1)\bar{c}^2}{T^2}, \\ \sum_{t=1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)t &= 1 - \frac{\bar{c} \left( \left(1 - \frac{1}{2}\bar{c}\right) T - 1 + \frac{1}{2}\bar{c} \right)}{T}, \\ \sum_{t=T_B+1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)D_t &= -\frac{\bar{c}}{T} + \frac{(1-\lambda)\bar{c}^2}{T} - \frac{\bar{c}^2}{T^2}, \\ \sum_{t=1}^T ((1 - \bar{\rho}L)t)^2 &= \left( \frac{1}{3}\bar{c}^2 - \bar{c} + 1 \right) T + \bar{c} - \frac{1}{2}\bar{c}^2 + \frac{1}{6}\frac{\bar{c}^2}{T}, \end{aligned}$$

$$\sum_{t=T_B+1}^T (1 - \bar{\rho}L)t(1 - \bar{\rho}L)D_t = 1 - \bar{c}\lambda - \bar{c} \left( 1 - \frac{1}{2}\bar{c} - \lambda + \frac{1}{2}\bar{c}\lambda^2 \right) - \frac{\bar{c} \left( -1 + \frac{1}{2}\bar{c} + \frac{1}{2}\bar{c}\lambda \right)}{T},$$

$$\sum_{t=T_B+1}^T ((1 - \bar{\rho}L)D_t)^2 = 1 + \frac{(1 - \lambda)c^2}{T} - \frac{\bar{c}^2}{T^2}.$$

**Fact .2.** By direct calculation, elements of matrix  $\sum_{t=1}^T \tilde{z}_t \tilde{u}_t$  are given by

$$\sum_{t=1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)u_t = u_1 - \frac{\bar{c} \sum_{t=2}^T \Delta u_t}{T} + \frac{\bar{c}^2 \sum_{t=2}^T u_{t-1}}{T^2},$$

$$\sum_{t=1}^T (1 - \bar{\rho}L)t(1 - \bar{\rho}L)u_t = u_1 + \sum_{t=2}^T \Delta u_t + \frac{\bar{c} (\sum_{t=2}^T \Delta u_t - \sum_{t=2}^T t \Delta u_t - \sum_{t=2}^T u_{t-1})}{T}$$

$$+ \frac{\bar{c}^2 (\sum_{t=2}^T t u_{t-1} - \sum_{t=2}^T u_{t-1})}{T^2},$$

$$\sum_{t=T_B+1}^T (1 - \bar{\rho}L)D_t(1 - \bar{\rho}L)u_t = u_{T_B+1} - \left( 1 + \frac{\bar{c}}{T} \right) u_{T_B} - \frac{\bar{c} \sum_{t=T_B+2}^T \Delta u_t}{T}$$

$$+ \frac{\bar{c}^2 \sum_{t=T_B+2}^T u_{t-1}}{T^2}.$$

Using the results from Fact .1 we symbolically compute  $\left( \sum_{t=1}^T \tilde{z}_t \tilde{z}_t' \right)^{-1}$ , then the result of this matrix inversion is multiplied by  $\left( \sum_{t=1}^T \tilde{z}_t \tilde{u}_t \right)$  making use of the results from Fact .2. This computation yields the bias terms  $(\tilde{\beta} - \beta) = ((\tilde{\mu} - \mu), (\tilde{\delta} - \delta), (\tilde{\theta} - \theta))'$ . The expressions are very lengthy, for this reason we present these expressions in a compact form. Here, we will use these expressions in a simplified form. Derivation of these expressions involve tedious algebra and omitted.

From the OLS normal equations the expression for  $(\tilde{\mu} - \mu)$  can be written as

$$(\tilde{\mu} - \mu) = \frac{T^6 M_{16} + T^5 M_{15} + T^4 M_{14} + T^3 M_{13} + O_p(T^2)}{T^6 M_T}, \quad (\text{A.9})$$

where  $M_{1j}$  are polynomials in  $\bar{c}$  and  $\lambda$  involving various sums of  $u_t$ ,  $u_{t-1}$ ,  $\Delta u_t$ , etc, and will be defined in shorter forms later in this section, and  $M_T$  is given by

$$M_T = (-12 + 12\bar{c} - 4\bar{c}^2) + \sum_{j=1}^5 m_{Tj} T^{j-6}, \quad (\text{A.10})$$

where  $m_{Tj}$  are polynomials in  $\bar{c}$  and  $\lambda$ . The notation  $m_{ij}$  will be used to mean the polynomials in  $\bar{c}$  and  $\lambda$ , hereafter. The  $M_T$  is a sixth degree polynomial in  $T$  involving various powers of  $\bar{\lambda}$  and  $\bar{c}$ . The full expression for  $M_T$  is very lengthy and omitted. We normalize (A.9) in the following manner

$$T^{1/2}(\tilde{\mu} - \mu) = \frac{T^{1/2} M_{16} + T^{-1/2} M_{15} + T^{-3/2} M_{14} + T^{-5/2} M_{13} + O_p(T^{-7/2})}{M_T}. \quad (\text{A.11})$$

Now, we will consider each term in the numerator and the denominator of (A.11) separately.

It is simple to show that

$$\begin{aligned}\lim_{T \rightarrow \infty} M_T &= (-12 + 12\bar{c} - 4\bar{c}^2) \\ &= M.\end{aligned}\tag{A.12}$$

From the OLS normal equations,  $T^{1/2}M_{16}$  can be written as

$$T^{1/2}M_{16} = T^{1/2}(-12 + 12\bar{c} - 4\bar{c}^2)u_1.$$

From this expression, we note that  $M_{16} = Mu_1$ . Using this equality (A.11) can be rearranged as

$$\begin{aligned}T^{1/2}(\tilde{\mu} - \mu - u_1) &= \\ &= \frac{T^{1/2}M_{16} + T^{-1/2}M_{15} + T^{-3/2}M_{14} + T^{-5/2}M_{13} + O_p(T^{-7/2})}{M_T} \\ &+ \frac{(M - M_T)u_1}{T^{-1/2}M_T}\end{aligned}\tag{A.13}$$

Note that

$$\frac{(H - H_T)u_1}{T^{-1/2}H_T} \rightarrow 0, \quad O_p\left(T^{-7/2}\right) \rightarrow 0 \quad \text{as } T \rightarrow \infty.\tag{A.14}$$

*Remark .1.*  $T^{-\epsilon}u_{T_B} \xrightarrow{a.s.} 0$  for  $\epsilon > 1/2$ ,  $T^{-\epsilon}u_{T_B+1} \xrightarrow{a.s.} 0$  for  $\epsilon > 1/2$ ,  $T^{-\epsilon}u_1 \xrightarrow{a.s.} 0$  for  $\epsilon > 0$ ,  $T^{-\epsilon}v_{T_B+1} \xrightarrow{a.s.} 0$  for  $\epsilon > 0$ ,  $T^{-\epsilon}\sum_{t=2}^T \Delta u_t \xrightarrow{a.s.} 0$  for  $\epsilon > 1/2$ ,  $T^{-\epsilon}\sum_{t=2}^T t\Delta u_t \xrightarrow{a.s.} 0$  for  $\epsilon > 3/2$ ,  $T^{-\epsilon}\sum_{t=2}^T u_{t-1} \xrightarrow{a.s.} 0$  for  $\epsilon > 3/2$ ,  $T^{-\epsilon}\sum_{t=2}^T tu_{t-1} \xrightarrow{a.s.} 0$  for  $\epsilon > 5/2$  by strong law of McLeish (1975).

Using the OLS normal equations and with tedious but straightforward algebra  $T^{-1/2}M_{15}$ ,  $T^{-3/2}M_{14}$ , and  $T^{-5/2}M_{13}$  can be shown as

$$\begin{aligned}T^{-1/2}M_{15} &= T^{-1/2}\left(m_{15}^\mu \sum_{t=2}^T \Delta u_t + m_{25}^\mu u_1 + m_{35}^\mu (u_{T_B+1} - u_{T_B})\right) \\ &= T^{-1/2}\left(m_{15}^\mu \sum_{t=2}^T \Delta u_t + m_{25}^\mu u_1 + m_{35}^\mu v_{T_B+1}\right),\end{aligned}\tag{A.15}$$

$$\begin{aligned}T^{-3/2}M_{14} &= T^{-3/2}\left(m_{14}^\mu \sum_{t=2}^T \Delta u_t + m_{24}^\mu \sum_{t=2}^T u_{t-1} + m_{34}^\mu \sum_{t=2}^T t\Delta u_t \right. \\ &\quad \left. + m_{44}^\mu u_1 + m_{54}^\mu u_{T_B} + m_{64}^\mu u_{T_B+1}\right),\end{aligned}\tag{A.16}$$

$$\begin{aligned}T^{-5/2}M_{13} &= T^{-5/2}\left(m_{13}^\mu \sum_{t=2}^T \Delta u_t + m_{23}^\mu \sum_{t=2}^T u_{t-1} + m_{33}^\mu \sum_{t=2}^T t\Delta u_t \right. \\ &\quad \left. + m_{43}^\mu \sum_{t=2}^T tu_{t-1} + m_{53}^\mu u_1 + m_{63}^\mu u_{T_B} + m_{73}^\mu u_{T_B+1}\right).\end{aligned}\tag{A.17}$$

From Lemma 1.1, Lemma 1.2, Remark .1, Herrndorf's weak convergence result, and continuous mapping theorem, we can establish that

$$T^{-1/2}M_{15} \implies \sigma m_{15}^\mu B_c(1), \quad (\text{A.18})$$

$$T^{-3/2}M_{14} \implies \sigma \left( m_{24}^\mu \int_0^1 B_c(r) dr + m_{34}^\mu \left( B_c(1) - \int_0^1 B_c(r) dr \right) \right), \quad (\text{A.19})$$

$$T^{-5/2}M_{13} \implies \sigma m_{43}^\mu \int_0^1 r B_c(r) dr. \quad (\text{A.20})$$

Combining (A.18)-(A.20) in (A.13) first part of Theorem 2.1 is proven by continuous mapping theorem.

Using the OLS normal equations the GLS bias  $(\tilde{\delta} - \delta)$  can be decomposed as

$$(\tilde{\delta} - \delta) = \frac{T^5 M_{25} + T^4 M_{24} + T^3 M_{23} + O_p(T^2)}{T^6 M_T}.$$

We normalize this equation to obtain

$$T^{1/2}(\tilde{\delta} - \delta) = \frac{T^{-1/2}M_{25} + T^{-3/2}M_{24} + T^{-5/2}M_{23} + O_p(T^{-7/2})}{M_T}, \quad (\text{A.21})$$

where

$$\begin{aligned} T^{-1/2}M_{25} &= T^{-1/2} \left( m_{15}^\delta \sum_{t=2}^T \Delta u_t + m_{25}^\delta u_1 + m_{35}^\delta (u_{T_B+1} - u_{T_B}) \right) \\ &= T^{-1/2} \left( m_{15}^\delta \sum_{t=2}^T \Delta u_t + m_{25}^\delta u_1 + m_{35}^\delta v_{T_B+1} \right), \end{aligned} \quad (\text{A.22})$$

$$\begin{aligned} T^{-3/2}M_{24} &= T^{-3/2} \left( m_{14}^\delta \sum_{t=2}^T \Delta u_t + m_{24}^\delta \sum_{t=2}^T u_{t-1} + m_{34}^\delta \sum_{t=2}^T t \Delta u_t + m_{44}^\delta \sum_{t=T_B+2}^T \Delta u_t \right. \\ &\quad \left. + m_{54}^\delta u_1 + m_{64}^\delta u_{T_B+1} + m_{74}^\delta u_{T_B} \right), \end{aligned} \quad (\text{A.23})$$

$$\begin{aligned} T^{-5/2}M_{23} &= T^{-5/2} \left( m_{13}^\delta \sum_{t=2}^T \Delta u_t + m_{23}^\delta \sum_{t=2}^T u_{t-1} + m_{33}^\delta \sum_{t=2}^T t \Delta u_t + m_{43}^\delta \sum_{t=2}^T t u_{t-1} \right. \\ &\quad \left. + m_{53}^\delta \sum_{t=T_B+2}^T \Delta u_t + m_{63}^\delta \sum_{t=T_B+2}^T u_{t-1} + m_{73}^\delta u_1 + m_{83}^\delta u_{T_B+1} + m_{93}^\delta u_{T_B} \right). \end{aligned} \quad (\text{A.24})$$

*Remark .2.*  $T^{-\epsilon} \sum_{t=T_B+2}^T \Delta u_t \xrightarrow{a.s.} 0$  for  $\epsilon > 1/2$ ,  $T^{-\epsilon} \sum_{t=T_B+2}^T t \Delta u_t \xrightarrow{a.s.} 0$  for  $\epsilon > 3/2$ ,  $T^{-\epsilon} \sum_{t=T_B+2}^T u_{t-1} \xrightarrow{a.s.} 0$  for  $\epsilon > 3/2$ ,  $T^{-\epsilon} \sum_{t=T_B+2}^T t u_t \xrightarrow{a.s.} 0$  for  $\epsilon > 5/2$  by strong law of McLeish (1975).

A straightforward application of Lemma 1.1, Lemma 1.2, Remark .1, Remark .2, Herrndorf's weak convergence result, and continuous mapping theorem yields

$$T^{-1/2}M_{25} \implies \sigma m_{15}^\delta B_c(1), \quad (\text{A.25})$$

$$T^{-3/2}M_{24} \implies \sigma \left( m_{24}^\delta \int_0^1 B_c(r) dr + m_{34}^\delta \left( B_c(1) - \int_0^1 B_c(r) dr \right) \right), \quad (\text{A.26})$$

$$T^{-5/2}M_{23} \implies \sigma m_{43}^\delta \int_0^1 r B_c(r) dr. \quad (\text{A.27})$$

We substitute the values of  $m_{ij}^\delta$  from the OLS normal equations in (A.25)-(A.27), then, some tedious algebra and the continuous mapping theorem yields the result in second part of Theorem 2.1 directly from (A.21) and (A.25)-(A.27).

From the OLS normal equations, the expression for  $(\tilde{\theta} - \theta)$  can be written as

$$\begin{aligned} (\tilde{\theta} - \theta) &= \frac{T^6 \left( (12 - 12\bar{c} + 4\bar{c}^2)u_{T_{B+1}} + (12 - 12\bar{c} + 4\bar{c}^2)u_{T_B} \right) + O_p(T^5)}{T^6 M_T} \\ &= \frac{T^6 \left( (-12 + 12\bar{c} - 4\bar{c}^2)(u_{T_{B+1}} - u_{T_B}) \right) + O_p(T^5)}{T^6 M_T} \\ &= \frac{T^6 \left( (-12 + 12\bar{c} - 4\bar{c}^2)v_{T_{B+1}} \right) + O_p(T^5)}{T^6 M_T}. \end{aligned} \quad (\text{A.28})$$

From this equation, it is simple to establish that

$$(\tilde{\theta} - \theta) \implies v_{T_{B+1}},$$

where we used (A.12) and continuous mapping theorem.

## Proof of Theorem 2.2

Proof of theorem Theorem 2.2 follows the same steps as the proof of Theorem 2.1. First step is to obtain  $(\tilde{\beta} - \beta) = \left( \sum_{t=1}^T \tilde{z}_t \tilde{z}'_t \right)^{-1} \left( \sum_{t=1}^T \tilde{z}_t \tilde{u}_t \right)$  redefining the shift dummy in (A.8) by  $D_t = \{0 | t \leq T_B, t - T_B | t > T_B\}$ .

**Fact .3.** Given the definition  $D_t = \{0 | t \leq T_B, t - T_B | t > T_B\}$  and noting that under GLS transformation the first observation is not transformed, the elements of the moment matrix  $\sum_{t=1}^T \tilde{z}_t \tilde{z}'_t$  are given by

$$\begin{aligned} \sum_{t=1}^T ((1 - \bar{\rho}L))^2 &= 1 + \frac{(T-1)\bar{c}^2}{T^2}, \\ \sum_{t=1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)t &= 1 - \frac{\bar{c} \left( (1 - \frac{1}{2}\bar{c}) T - 1 + \frac{1}{2}\bar{c} \right)}{T}, \\ \sum_{t=T_B+1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)D_t &= \frac{1}{2} \bar{c} (\lambda - 1) (\bar{c} \lambda + 2 - \bar{c}) + \frac{1}{2} \frac{\bar{c}^2 (\lambda - 1)}{T}, \\ \sum_{t=1}^T ((1 - \bar{\rho}L)t)^2 &= \left( \frac{1}{3} \bar{c}^2 - \bar{c} + 1 \right) T + \bar{c} - \frac{1}{2} \bar{c}^2 + \frac{1}{6} \frac{\bar{c}^2}{T}, \\ \sum_{t=T_B+1}^T (1 - \bar{\rho}L)t(1 - \bar{\rho}L)D_t &= \frac{1}{6} (\lambda - 1) (\bar{c}^2 \lambda^2 + \bar{c}^2 \lambda - 2\bar{c}^2 - 6 + 6\bar{c}) T \\ &\quad + \frac{1}{2} \bar{c} (-2 + \bar{c}) (\lambda - 1) - \frac{1}{6} \frac{\bar{c}^2 (\lambda - 1)}{T}, \end{aligned}$$

$$\begin{aligned} \sum_{t=T_B+1}^T ((1 - \bar{\rho}L)D_t)^2 &= -\frac{1}{3}(\lambda - 1) (\bar{c}^2 \lambda^2 + 3\bar{c}\lambda - 2\bar{c}^2 \lambda + 3 + \bar{c}^2 - 3\bar{c}) T \\ &\quad - \frac{1}{2}\bar{c}(\lambda - 1)(\bar{c}\lambda + 2 - \bar{c}) - \frac{1}{6}\frac{\bar{c}^2(\lambda - 1)}{T}. \end{aligned}$$

**Fact .4.** Noting that under GLS transformation the first observation is not transformed, we directly calculate the elements of the vector  $\sum_{t=1}^T \tilde{\mathbf{z}}_t \tilde{u}_t$ , which are given by

$$\begin{aligned} \sum_{t=1}^T (1 - \bar{\rho}L)(1 - \bar{\rho}L)u_t &= u_1 - \frac{\bar{c} \sum_{t=2}^T \Delta u_{t-1}}{T} + \frac{\bar{c}^2 \sum_{t=2}^T u_{t-1}}{T^2}, \\ \sum_{t=1}^T (1 - \bar{\rho}L)t(1 - \bar{\rho}L)u_t &= u_1 + \sum_{t=2}^T \Delta u_{t-1} \\ &\quad + \frac{\bar{c} (\sum_{t=2}^T \Delta u_{t-1} - \sum_{t=2}^T t \Delta u_{t-1} - \sum_{t=2}^T u_{t-1})}{T} \\ &\quad + \frac{\bar{c}^2 (\sum_{t=2}^T t u_{t-1} - \sum_{t=2}^T u_{t-1})}{T^2}, \\ \sum_{t=T_B+1}^T (1 - \bar{\rho}L)D_t(1 - \bar{\rho}L)u_t &= u_{T_B+1} - \left(1 + \frac{c}{T}\right) u_{T_B} + (1 + \bar{c}\lambda) \sum_{t=T_B+2}^T \Delta u_{t-1} \\ &\quad + \frac{-\bar{c} \sum_{t=T_B+2}^T t \Delta u_{t-1} - \bar{c}(1 + \bar{c}\lambda) \sum_{t=T_B+2}^T u_{t-1}}{T} \\ &\quad + \frac{\bar{c} \sum_{t=T_B+2}^T \Delta u_{t-1}}{T} + \frac{\bar{c}^2 \sum_{t=T_B+2}^T t u_{t-1} - \bar{c}^2 \sum_{t=T_B+2}^T u_{t-1}}{T^2}. \end{aligned}$$

Using the results from Fact .3 and Fact .4 we obtain an expression for each element of  $(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$ . These expressions can be written as

$$(\tilde{\mu} - \mu) = \frac{T^6 H_{\mu 7} + T^5 H_{\mu 6} + T^4 H_{\mu 5} + T^3 H_{\mu 4} + O_p(T^2)}{T^6 H_T}, \quad (\text{A.29a})$$

$$(\tilde{\delta} - \delta) = \frac{T^6 H_{\delta 6} + T^5 H_{\delta 5} + T^4 H_{\delta 4} + O_p(T^3)}{T^7 H_T}, \quad (\text{A.29b})$$

$$(\tilde{\theta} - \theta) = \frac{T^6 H_{\theta 6} + T^5 H_{\theta 5} + T^4 H_{\theta 4} + O_p(T^3)}{T^7 H_T}, \quad (\text{A.29c})$$

where  $H_{ij}$  are polynomials in  $\bar{c}$  and  $\lambda$  involving various sums of  $u_t$ ,  $u_{t-1}$ ,  $\Delta u_t$ , etc, and  $H_T$  is given by

$$H_T = \sum_{j=1}^k m_{Tj} T^{j-k}, \quad k = 6, 7.$$

The  $M_T$  is a sixth degree polynomial in  $T$  involving various powers of  $\bar{\lambda}$  and  $\bar{c}$ . The full expression for  $M_T$  is very lengthy and omitted. The expression for  $H_T$  can be written as

$$\begin{aligned} \lim_{T \rightarrow \infty} H_T &= 2\bar{c}^4 \lambda^5 + 2\bar{c}^4 \lambda^4 + (-10\bar{c}^4 - 24\bar{c}^2 + 24\bar{c}^3) \lambda^3 \\ &\quad + (6\bar{c}^4 + 24\bar{c}^2 - 24\bar{c}^3) \lambda^2 + (24\bar{c}^2 - 72\bar{c} + 72) \lambda \\ &= H. \end{aligned} \quad (\text{A.30})$$

Normalizing each equation in (A.29) we obtain

$$T^{1/2}(\tilde{\mu} - \mu) = \frac{H_{\mu 7} + T^{-1/2}H_{\mu 6} + T^{-3/2}H_{\mu 5} + T^{-5/2}H_{\mu 4} + O_p(T^{-7/2})}{H_T}, \quad (\text{A.31a})$$

$$T^{1/2}(\tilde{\delta} - \delta) = \frac{T^{-1/2}H_{\delta 6} + T^{-3/2}H_{\delta 5} + T^{-3/2}H_{\delta 4} + O_p(T^{-7/2})}{H_T}, \quad (\text{A.31b})$$

$$T^{1/2}(\tilde{\theta} - \theta) = \frac{T^{-1/2}H_{\theta 6} + T^{-3/2}H_{\theta 5} + T^{-5/2}H_{\theta 4} + O_p(T^{-7/2})}{H_T}. \quad (\text{A.31c})$$

Further, using  $H_{\mu 7} = Hu_1$  we obtain

$$T^{1/2}(\tilde{\mu} - \mu) = \frac{T^{-1/2}H_{\mu 6} + T^{-3/2}H_{\mu 5} + T^{-5/2}H_{\mu 4} + T^{-11/2}O_p(T^2)}{H_T} + \frac{(H - H_T)u_1}{T^{-1/2}H_T}. \quad (\text{A.31a}')$$

The OLS normal equations can be use to show that

$$\begin{aligned} T^{-1/2}H_{j6} &= T^{-1/2} \left( h_{16}^j \sum_{t=2}^T \Delta u_{t-1} + h_{26}^j \sum_{t=T_B+2}^T \Delta u_{t-1} + h_{36}^j u_1 + h_{46}^j (u_{T_B+1} - u_{T_B}) \right), \\ T^{-3/2}H_{j5} &= T^{-3/2} \left( h_{15}^j \sum_{t=2}^T \Delta u_{t-1} + h_{25}^j \sum_{t=2}^T u_{t-1} + h_{35}^j \sum_{t=2}^T t \Delta u_{t-1} + h_{45}^j \sum_{t=T_B+2}^T t \Delta u_{t-1} \right. \\ &\quad \left. + h_{55}^j \sum_{t=T_B+2}^T \Delta u_{t-1} + h_{65}^j \sum_{t=T_B+2}^T u_{t-1} + h_{75}^j u_1 + h_{85}^j u_{T_B} + h_{95}^j u_{T_B+1} \right), \\ T^{-5/2}H_{j4} &= T^{-5/2} \left( h_{14}^j \sum_{t=2}^T \Delta u_{t-1} + h_{24}^j \sum_{t=2}^T u_{t-1} + h_{34}^j \sum_{t=2}^T t \Delta u_{t-1} + h_{44}^j \sum_{t=2}^T t u_{t-1} \right. \\ &\quad \left. + h_{54}^j \sum_{t=T_B+2}^T t u_{t-1} + h_{64}^j \sum_{t=T_B+2}^T t \Delta u_{t-1} + h_{74}^j \sum_{t=T_B+2}^T \Delta u_{t-1} + h_{84}^j \sum_{t=T_B+2}^T u_{t-1} \right. \\ &\quad \left. + h_{94}^j u_1 + h_{10,4}^j u_{T_B} + h_{11,4}^j u_{T_B+1} \right). \end{aligned}$$

where  $j = \mu, \delta, \theta$ .

Using the results from Lemma 1.1, Lemma 1.2, Remark .1, and Remark .2, Herrndorf's weak convergence result, and continuous mapping theorem we can establish that

$$H_{j6} = \sigma(h_{16}^j B_c(1) + h_{26}^j B_c(\lambda)), \quad (\text{A.32})$$

$$\begin{aligned} H_{j5} &= \sigma \left( h_{25}^j \int_0^1 B_c(r) dr + h_{35}^j \left( B_c(1) - \int_0^1 B_c(r) dr \right) \right. \\ &\quad \left. + h_{45}^j \left( (B_c(1) - \lambda B_c(\lambda)) - \left( \int_0^1 B_c(r) dr - \int_0^\lambda B_c(r) dr \right) \right) \right. \\ &\quad \left. + h_{65}^j \int_0^\lambda B_c(r) dr \right), \quad (\text{A.33}) \end{aligned}$$

$$H_{j4} = \sigma \left( h_{54}^j \int_0^1 r B_c(r) dr + h_{64}^j \left( \int_0^1 r B_c(r) dr - \int_0^\lambda r B_c(r) dr \right) \right). \quad (\text{A.34})$$

To complete the proof we obtain the values of  $h_{jk}^j$  from the OLS bias terms, then, after some tedious algebra result follows from applying continuous mapping theorem on (A.31).

## Proof of Lemma 2.1

Note that  $y_t^{\tau\mu}$  and  $y_t^{\tau\delta}$  can be written as

$$\begin{aligned} y_t^{\tau\mu} &= u_t - (\tilde{\mu} - \mu) - (\tilde{\delta} - \delta)t - (\tilde{\theta} - \theta)D_t, & D_t &= \{0|t \leq T_B, 1|t > T_B\}, \\ y_t^{\tau\delta} &= u_t - (\tilde{\mu} - \mu) - (\tilde{\delta} - \delta)t - (\tilde{\theta} - \theta)D_t, & D_t &= \{0|t \leq T_B, t - T_B|t > T_B\}. \end{aligned}$$

Mapping  $y_t^{\tau\mu}$  and  $y_t^{\tau\delta}$  into the fixed interval  $r \in [0, 1]$ , we obtain

$$\begin{aligned} T^{-1/2}y_{[Tr]}^{\tau\mu} &= T^{-1/2}u_{[Tr]} - T^{-1/2}(\tilde{\mu} - \mu) \\ &\quad - \left(\frac{[Tr]}{T}\right)T^{-1/2}(\tilde{\delta} - \delta) - \lambda\mathbf{1}(t > T_B)T^{-1/2}(\tilde{\theta} - \theta), \end{aligned} \tag{A.35}$$

$$\begin{aligned} T^{-1/2}y_{[Tr]}^{\tau\delta} &= T^{-1/2}u_{[Tr]} - T^{-1/2}(\tilde{\mu} - \mu) - \left(\frac{[Tr]}{T}\right)T^{-1/2}(\tilde{\delta} - \delta) \\ &\quad - \left(\frac{[(T - T_B)r]}{T}\right)\mathbf{1}(t > T_B)T^{-1/2}(\tilde{\theta} - \theta). \end{aligned} \tag{A.36}$$

*Remark .3.* For the level shift model:  $T^{-1/2}(\tilde{\mu} - \mu) \xrightarrow{p} 0$ ,  $T^{-1/2}(\tilde{\theta} - \theta) \xrightarrow{p} 0$ .

*Remark .4.* For the trend shift model:  $T^{-1/2}(\tilde{\mu} - \mu) \xrightarrow{p} 0$ .

Then, the result in Lemma 2.1 follows from Remark .3, Remark .4, Lemma 1.1, Theorem 2.1, Theorem 2.2, Herrndorf's weak convergence result, and continuous mapping theorem.

## Proof of Theorem 2.3

Proof of Theorem 2.3 requires consistent estimators of  $\sigma_v^2$  and  $\sigma^2$ . Let these estimators be denoted by  $\hat{\sigma}_v^2$  and  $\hat{\sigma}^2$ .

As pointed out by Phillips (1987a),

$$T^{-1} \sum_{t=1}^T \hat{v}_t^2 \xrightarrow{a.s.} \sigma_v^2$$

by strong law of McLeish (1975). Thus, a consistent estimator of  $\sigma_v^2$  is readily available.

A consistent estimator of  $\sigma^2$  can be found by replacing the condition (b) of Assumption 1.1 by the more stronger condition  $\sup_t \mathbf{E} |v_t|^{2\beta} < \infty$  for some  $\beta > 2$ . From Theorem 4.2 of Phillips (1987a)

$$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^T \hat{v}_t^2 + 2T^{-1} \sum_{j=1}^{\ell} \sum_{t=j+1}^T \hat{v}_t \hat{v}_{t-j} \tag{A.37}$$

is a consistent estimator of  $\sigma^2$  if  $\ell = o(T^{1/4})$ . Although, this estimator is consistent it is not constrained to be nonnegative. To obtain a nonnegative estimate one can replace (A.37) by

$$\hat{\sigma}^2(\omega) = T^{-1} \sum_{t=1}^T \hat{v}_t^2 + 2T^{-1} \sum_{j=1}^{\ell} \omega_\ell(j) \sum_{t=j+1}^T \hat{v}_t \hat{v}_{t-j}, \tag{A.38}$$

where a simple choice for  $\omega_\ell(j)$  is  $\omega_\ell(j) = 1 - j/(\ell + 1)$  suggested by Newey and West (1987).

Now, we can establish the asymptotic distribution of the test statistics  $\Gamma_1^*$  and  $\Gamma_1^*$ . Note that for the level shift model,

$$T(\hat{\rho} - \rho) = \left( \frac{1}{T^2} \sum_{t=1}^T y_{t-1}^{\tau_\mu^2} \right)^{-1} \left( \frac{1}{T} \sum_{t=1}^T y_{t-1}^{\tau_\mu} v_t \right), \quad (\text{A.39})$$

$$t_\rho = \left( \frac{\hat{\sigma}_v^2}{T^2} \sum_{t=1}^T y_{t-1}^{\tau_\mu^2} \right)^{-1/2} \left( \frac{1}{T} \sum_{t=1}^T y_{t-1}^{\tau_\mu} v_t \right), \quad (\text{A.40})$$

and

$$T(\hat{\rho} - \rho) = \left( \frac{1}{T^2} \sum_{t=1}^T y_{t-1}^{\tau_\delta^2} \right)^{-1} \left( \frac{1}{T} \sum_{t=1}^T y_{t-1}^{\tau_\delta} v_t \right), \quad (\text{A.41})$$

$$t_\rho = \left( \frac{\hat{\sigma}_v^2}{T^2} \sum_{t=1}^T y_{t-1}^{\tau_\delta^2} \right)^{-1/2} \left( \frac{1}{T} \sum_{t=1}^T y_{t-1}^{\tau_\delta} v_t \right) \quad (\text{A.42})$$

for the trend shift model. It can readily be shown that

$$T^{-1} \sum_{t=1}^T y_{t-1}^{\tau_\mu} v_t \implies \int_0^1 B_c^{\tau_\mu}(r) dB(r) + \frac{(\sigma^2 - \sigma_v^2)}{2}, \quad (\text{A.43})$$

$$T^{-2} \sum_{t=1}^T y_{t-1}^{\tau_\mu^2} \implies \int_0^1 B_c^{\tau_\mu}(r)^2 dr \quad (\text{A.44})$$

for the level shift model, and

$$T^{-1} \sum_{t=1}^T y_{t-1}^{\tau_\delta} v_t \implies \int_0^1 B_c^{\tau_\delta}(r) dB(r) + \frac{(\sigma^2 - \sigma_v^2)}{2}, \quad (\text{A.45})$$

$$T^{-2} \sum_{t=1}^T y_{t-1}^{\tau_\delta^2} \implies \int_0^1 B_c^{\tau_\delta}(r)^2 dr \quad (\text{A.46})$$

for the trend shift model. Then, the proof follows from (A.39)-(A.46) and continuous mapping theorem.

### Proof of Theorem 3.1

To prove Theorem 3.1 we only need to show that the detrended process  $y_t^{\tau_\mu}$  under the innovational outlier model has the same limit as the one obtained from the additive outlier model. Note that the detrended process  $y_t^{\tau_\mu}$  for the innovational outlier model can be written as

$$y_t^{\tau_\mu} = u_t - (\tilde{\mu} - \mu) - (\tilde{\delta} - \delta)t - (\tilde{\alpha} - \alpha)' \mathbf{c}_t,$$

where  $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)'$  is a  $(k+1)$ -dimensional parameter vector and  $\mathbf{c}_t = (D_t, \dots, D_{t-k})'$  is a vector containing  $k + 1$  level shift dummies  $D_{t-i} = \{0 | t \leq (T_B + i), 1 | t > (T_B + i)\}$ ,

$i = 0, 1, \dots, k$ . Mapping  $y_t^{\tau_\mu}$  into the fixed interval  $r \in [0, 1]$ , we obtain

$$\begin{aligned} T^{-1/2}y_{[Tr]}^{\tau_\mu} &= T^{-1/2}u_{[Tr]} - T^{-1/2}(\tilde{\mu} - \mu) \\ &\quad - \left(\frac{[Tr]}{T}\right)T^{-1/2}(\tilde{\delta} - \delta) - \lambda\mathbf{1}(t > T_B)T^{-1/2}(\tilde{\alpha} - \alpha). \end{aligned} \quad (\text{A.47})$$

Since  $T^{-1/2}(\tilde{\alpha} - \alpha) \xrightarrow{P} 0$  the expression in (A.47) converges to  $B_c^{\tau_\mu}(r)$  where  $B_c^{\tau_\mu}(r)$  is defined in (2.26), thus, the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by (2.36) and (2.37) which directly follows from the Lemma 2.1 and the Theorem 2.3.

### Proof of Theorem 3.2

As in the case of the level shift model, showing that the detrended process  $y_t^{\tau_\delta}$  under the innovational outlier model has the same limit as the one obtained from the additive outlier model is sufficient to prove the Theorem 3.2. Let  $d_t$  be as defined in (3.41) with  $\mathbf{c}_t$  defined by  $\mathbf{c}_t = (D_t^\delta, D_t^\mu, D_{t-1}^\mu, \dots, D_{t-k+1}^\mu)'$ . Now partition  $\mathbf{c}_t$  and  $\alpha$  in the following manner:  $\mathbf{c}_t = (D_{1,t}, \mathbf{D}'_{2,t})'$  and  $\alpha = (\alpha_\delta, \alpha'_\mu)'$  where  $D_{1,t} = D_t^\delta$ ,  $\mathbf{D}'_{2,t} = (D_t^\mu, D_{t-1}^\mu, \dots, D_{t-k+1}^\mu)'$ ,  $\alpha_\delta = \alpha_0$ , and  $\alpha'_\mu = (\alpha_1, \alpha_2, \dots, \alpha_k)'$ . Now, the detrended process  $y_t^{\tau_\delta}$  for the innovational outlier model can be written as

$$y_t^{\tau_\delta} = u_t - (\tilde{\mu} - \mu) - (\tilde{\delta} - \delta)t - (\tilde{\alpha}_\delta - \alpha_\delta)D_{1,t} - (\tilde{\alpha}_\mu - \alpha_\mu)'\mathbf{D}_{2,t}.$$

Now, by mapping  $y_t^{\tau_\delta}$  into the fixed interval  $r \in [0, 1]$ , we obtain

$$\begin{aligned} T^{-1/2}y_{[Tr]}^{\tau_\delta} &= T^{-1/2}u_{[Tr]} - T^{-1/2}(\tilde{\mu} - \mu) - \left(\frac{[Tr]}{T}\right)T^{-1/2}(\tilde{\delta} - \delta) \\ &\quad - \left(\frac{[(T - T_B)r]}{T}\right)\mathbf{1}(t > T_B)T^{-1/2}(\tilde{\alpha}_\delta - \alpha_\delta) \\ &\quad - \lambda\mathbf{1}(t > T_B)T^{-1/2}(\tilde{\alpha}_\mu - \alpha_\mu). \end{aligned} \quad (\text{A.48})$$

The expression in (A.48) converges to  $B_c^{\tau_\delta}(r)$  where  $B_c^{\tau_\delta}(r)$  is defined in (2.27), since  $T^{-1/2}(\tilde{\alpha}_\mu - \alpha_\mu) \xrightarrow{P} 0$ . Therefore, it directly follows from the Lemma 2.1 and the Theorem 2.3 that the asymptotic distributions of the test statistics  $\Gamma_1^*$  and  $\Gamma_2^*$  are given by (2.38) and (2.39).

### Proof of Theorem 3.3

The proof of the Theorem 3.4 is the same as the proof of the Theorem 3.1 redefining  $\mathbf{c}_t$  by  $\mathbf{c}_t = (D_{T_{B_1}}, \dots, D_{T_{B_q}})'$  where  $D_{T_{B_k}} = \{0|t \leq T_{B_k}, 1|t > T_{B_k}\}$  for  $k = 1, 2, \dots, q$ .

### Proof of Theorem 3.4

The proof of the Theorem 3.4 is the same as the proof of the Theorem 3.2 redefining  $\mathbf{c}_t$  by  $\mathbf{c}_t = (D_{T_{B_1}}, \dots, D_{T_{B_q}})'$  where  $D_{T_{B_k}} = \{0|t \leq T_{B_k}, t - T_{B_k}|t > T_{B_k}\}$  for  $k = 1, 2, \dots, q$ .

### Proof of Theorem 3.5

Since the finite dimensional the asymptotic distributions of the test statistics are already established by Theorem 2.3, we only need to show that

$$\begin{aligned}\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( c + \left( \int_0^1 B_c^{\tau\mu}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{-1} \right), \\ \inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( \left( \left( \int_0^1 B_c^{\tau\mu}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{-1} + c \right) \right. \\ &\quad \left. \times \left( \int_0^1 B_c^{\tau\mu}(r, \lambda)^2 dr \right)^{1/2} \right).\end{aligned}$$

for the level shift model, and

$$\begin{aligned}\inf_{\lambda \in \Lambda} \Gamma_1^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( c + \left( \int_0^1 B_c^{\tau\delta}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\delta}(r, \lambda)^2 dr \right)^{-1} \right), \\ \inf_{\lambda \in \Lambda} \Gamma_2^*(\lambda) &\implies \inf_{\lambda \in \Lambda} \left( \left( \left( \int_0^1 B_c^{\tau\delta}(r, \lambda) dB(r) \right) \left( \int_0^1 B_c^{\tau\delta}(r, \lambda)^2 dr \right)^{-1} + c \right) \right. \\ &\quad \left. \times \left( \int_0^1 B_c^{\tau\delta}(r, \lambda)^2 dr \right)^{1/2} \right),\end{aligned}$$

for the trend shift model. To prove this result we appeal to Theorem 1 of Zivot and Andrews (1992). The details of the proof can be found in Appendix A of Zivot and Andrews (1992).

## References

- Bahargava, A.**, “On the Theory of Testing for Unit Roots in Observed Time Series,” *Review of Economic Studies*, 1986, *53*, 369–384.
- Balcilar, Mehmet**, “Efficient and Near Efficient Unit Root Tests in Models with Structural Change.” PhD dissertation, Wayne State University 1996.
- Banerjee, Anindya, Robin L. Lumsdaine, and James H. Stock**, “Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence,” *Journal of Business & Economic Statistics*, 1992, *10*, 271–287.
- Bobkoski, M. J.**, “Hypothesis Testing in Nonstationary Time Series.” PhD dissertation, Department of Statistics, University of Wisconsin 1983.
- Cavanagh, C.**, “Roots Local to Unity,” 1985. manuscript, Department of Economics, Harvard University.
- Chan, N. H. and C. Z. Wei**, “Asymptotic Inference for Nearly Nonstationary AR(1) Processes,” *Annals of Statistics*, 1987, *15*, 1050–1063.
- Christiano, Lawrence J.**, “Searching for Break in GNP,” *Journal of Business & Economic Statistics*, 1992, *10*, 237–249.
- Dickey, D. A.**, “Estimation and Hypothesis Testing for Nonstationary Time Series in Finite Samples.” PhD dissertation, Iowa State University 1976.
- and **W. A. Fuller**, “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association*, 1979, *74*, 427–431.
- and —, “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root,” *Econometrica*, 1981, *49*, 1057–1072.
- Diebold, Francis X. and Marc Nerlove**, “Unit Roots in Economic Time Series: A Selective Survey,” in Thomas B. Fomby and George F. Rhodes, eds., *Advances in Econometrics*, Vol. 8 1990, pp. 3–69.
- Dufour, Jean-Marie and Maxwell King**, “Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR(1) Errors,” *Journal of Econometrics*, 1991, *47*, 115–143.
- Elliot, Graham, Thomas J. Rothenberg, and James H. Stock**, “Efficient Tests for an Autoregressive Unit Root,” 1992. NBER Technical Working Paper No. 130.
- Fuller, Wayne A.**, *Introduction to Statistical Time Series*, John Wiley & Sons, Inc., 1976.
- McLeish, D. L.**, “A Maximal Inequality and Dependent Strong Laws,” *Annals of Probability*, 1975, *3*, 829–839.
- Nabeya, Seiji and Katsuto Tanaka**, “Limiting Power of Unit-Root Tests in Time Series Regressions,” *Journal of Econometrics*, 1990, *46*, 247–271.
- Nelson, Charles N. and Charles I. Plosser**, “Trends and Random Walks in Macroeconomic Time Series,” *Journal of Monetary Economics*, 1982, *50*, 136–162.
- Newey, W. K. and K. D. West**, “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 1987, *55*, 703–708.

- Noriega-Muro, Antonio E.**, *Nonstationarity and Structural Breaks in Economic Time series*, Avebury, 1993.
- Ouliaris, S., J. Y. Park, and P. C. B. Phillips**, "Testing for a Unit Root in the Presence of a Maintained Trend," in B. Raj, ed., *Advances in Econometrics and Modelling*, Amsterdam: Kluwer, 1989, pp. 7–28.
- Park, J. Y. and B. Choi**, "A new approach to Testing for Unit Roots," 1988. (CAE Working Paper #88-23, Cornell University, Ithaca, New York).
- Park, Joon Y. and Jaewhan Sung**, "Testing for Unit Roots in Models with Structural Change," *Econometric Theory*, 1994, 10, 917–936.
- Perron, Pierre**, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 1989, 57, 1361–1401.
- , "Testing for a Unit Root in Time Series With a Changing Mean," *Journal of Business & Economic Statistics*, 1990a, 8, 153–162.
- , "Further Evidence on Breaking Trend Functions in Macroeconomic Variables," 1990b. Research Memorandum No. 350, Economic Research Program, Princeton University.
- and **Timoty J. Vogelsang**, "Nonstationarity and Level Shifts With an Application to Purchasing Power Parity," *Journal of Business & Economic Statistics*, 1992, 10, 301–320.
- Phillips, Peter C. B.**, "Time Series Regression with a Unit Root," *Econometrica*, 1987a, 55, 277–301.
- , "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 1987b, 74, 535–547.
- , "Regression Theory for Near-integrated Time Series," *Econometrica*, 1988, 56, 1021–1043.
- and **Pierre Perron**, "Testing for a Unit Root in Time Series Regression," *Biometrika*, 1988, 75, 335–346.
- Said, Said E. and David A. Dickey**, "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, 1984, 71, 599–607.
- Saikkonen, Pentti and Ritva Luukkonen**, "Point Optimal Tests for Testing the Order of Differencing in ARIMA Models," *Econometric Theory*, 1993, 9, 343–362.
- Sargan, J. D. and Alok Bahargava**, "Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk," *Econometrica*, 1983, 51, 153–174.
- Schwert, G. W.**, "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business & Economic Statistics*, 1989, 7, 147–160.
- Stock, J. H. and M. W. Watson**, "Testing for Common Trends," *Journal of the American Statistical Association*, 1988, 83, 1097–1107.
- Stock, James H.**, "Deciding between I(1) and I(0)," *Journal of Econometrics*, 1994, 63, 105–131.
- Zivot, Eric and Donald W. K. Andrews**, "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business & Economic Statistics*, 1992, 10, 251–270.