October 29, 1999

# Comparison of Unit Root Tests for Time Series with Level Shifts *

by


Markku Lanne

University of Helsinki

Helmut Lütkepohl

Humboldt University Berlin

and

Pentti Saikkonen

University of Helsinki



### Abstract

Unit root tests are considered for time series which have a level shift at a known point in time. The shift can have a very general nonlinear form and additional deterministic mean and trend terms are allowed for. Prior to the tests the deterministic parts and other nuisance parameters of the data generation process are estimated in a first step. Then the series are adjusted for these terms and unit root tests of the Dickey-Fuller type are applied to the adjusted series. The properties of previously suggested tests of this sort are analyzed and a range of modifications is proposed which take into account estimation errors in the nuisance parameters. An important result is that estimation under the null hypothesis is preferable to estimation under local alternatives. This contrasts with results obtained by other authors for time series without level shifts.



*We are grateful to Ralf Brüggemann for helping with the computations and to Christian Müller for comments. Moreover, we thank the Deutsche Forschungsgemeinschaft, SFB 373, and the European Commission under the Training and Mobility of Researchers Programme (contract No. ERBFMRXCT980213) for financial support. The third author also thanks the Alexander von Humboldt Foundation for financial support under a Humboldt research award.

# 1 Introduction

Modeling structural shifts in time series has become an issue of central importance due to the massive interventions that occur regularly in economic systems. In this context testing for unit roots in the presence of structural shifts has attracted considerable attention in the recent literature (see, e.g., Perron (1989, 1990), Perron & Vogelsang (1992), Banerjee, Lumsdaine & Stock (1992), Zivot & Andrews (1992), Amsler & Lee (1995), Leybourne, Newbold & Vougas (1998), Montañés & Reyes (1998)). In some of the literature the time where the structural change occurs is assumed to be known and in other articles it is assumed unknown. In this study we assume that the break point is known. In practice, such an assumption is often reasonable because the timing of many interventions is known when the analysis is performed. For example, on January 1, 1999, a common currency was introduced in a number of European countries or the German unification is known to have occurred in 1990. These events have had an impact on some economic time series.

We will follow Saikkonen & Lütkepohl (1999) (henceforth S&L) and Lütkepohl, Müller & Saikkonen (1999) (henceforth LMS) and consider models with very general nonlinear deterministic shift functions. These authors propose a number of tests for unit roots based on the idea that the deterministic part is estimated in a first step and is subtracted from the series. Standard unit root tests are then applied to the adjusted series. The purpose of this study is to propose modifications of these tests which are expected to work well in small sample situations and we will perform Monte Carlo comparisons of the properties of the tests. The results lead to useful recommendations for applied work.

The structure of the study is as follows. Two alternative model types from S&L and LMS are presented in Sec. 2 together with the assumptions needed for asymptotic derivations. Estimation of the nuisance parameters within these models is discussed in Sec. 3 and a range of unit root tests is presented in Sec. 4 including the asymptotic distributions of the test statistics. Since some of the tests have distributions under the null hypothesis which are not tabulated, simulated critical values are presented in Sec. 5. Also in that section we present some local power simulations. A small sample comparison of the tests based on a Monte Carlo experiment is reported in Sec. 6 and conclusions are given in Sec. 7. Some proofs are provided in the Appendix.

The following general notation is used throughout. The lag and differencing operators are denoted by $L$ and $\Delta$, respectively. Hence, for a time series variable $y_t$, $Ly_t = y_{t-1}$ and $\Delta y_t = y_t - y_{t-1}$. Convergence in probability and in distribution are denoted by $\xrightarrow{p}$ and $\xrightarrow{d}$, respectively. Independently, identically distributed will be abbreviated as $iid(\cdot,\cdot)$, where the first and second moments are indicated in parentheses. Related to this notation, the normal distribution is written in the usual way as $N(\cdot,\cdot)$. Furthermore, $O(\cdot)$, $o(\cdot)$, $O_p(\cdot)$ and $o_p(\cdot)$ are the usual symbols for the order of convergence and convergence in probability, respectively, of a sequence. The symbol $\lambda_{min}(A)$ is reserved to denote the minimal eigenvalue of a matrix $A$. Moreover, $\|\cdot\|$ denotes the Euclidean norm. The abbreviations sup and inf are used as usual for supremum and infimum, respectively. The $n$-dimensional Euclidean space is signified as $\mathbf{R}^n$. DGP abbreviates data generation process, DF is short for Dickey-Fuller and OLS and GLS are used for ordinary least squares and generalized least squares, respectively. Moreover, AR abbreviates autoregressive or autoregressive process.

# 2 The Models

We consider two different general models for time series with a possible unit root and a level shift. The first one is of the form

$$y_t = \mu_0 + \mu_1 t + f_t(\theta)'\gamma + x_t, \qquad t = 1, 2, \ldots, \tag{2.1a}$$

where the scalars $\mu_0$ and $\mu_1$, the $(m \times 1)$ vector $\theta$ and the $(k \times 1)$ vector $\gamma$ are unknown parameters and $f_t(\theta)$ is a $(k \times 1)$ vector of deterministic sequences depending on the parameters $\theta$. The functional form of $f_t(\theta)$ is assumed to be known. If the sequence represents a level shift the timing of the shift is also known. For example, $f_t(\theta)$ may be thought of as a shift dummy variable which has the value zero before some given break period $T_1$ and the value one from then onwards. In that case, the break date $T_1$ is assumed to be known. Much more general situations are covered by our framework, however. Further discussion may be found in S&L and other examples are considered in Sec. 6.

The quantity $x_t$ represents an unobservable stochastic error term which is assumed to have a finite order AR representation,

$$b(L)(1 - \rho L)x_t = \varepsilon_t, \tag{2.1b}$$

where $\varepsilon_t \sim iid(0, \sigma^2)$ and $b(L) = 1 - b_1 L - \cdots - b_p L^p$ is a polynomial in the lag operator with roots bounded away from the unit circle. More precisely, for some $\epsilon > 0$, $b(L) \neq 0$ for

$|L| \leq 1+\epsilon$. For simplicity, we assume that a suitable number of presample values of the observed series $y_t$ is available. Obviously, if $\rho = 1$ and, hence, the DGP of $x_t$ has a unit root, then the same is true for $y_t$.

The second model is

$$b(L)y_t = \mu_0 + \mu_1 t + f_t(\theta)'\gamma + v_t, \qquad t = 1, 2, \ldots, \tag{2.2a}$$

where the error term $v_t$ is assumed to be an AR process of order 1,

$$v_t = \rho v_{t-1} + \varepsilon_t. \tag{2.2b}$$

As before, $\varepsilon_t \sim iid(0, \sigma^2)$ and $-1 < \rho \leq 1$ with $\rho = 1$ implying a unit root in $y_t$.

The parameters $\mu_0$, $\mu_1$ and $\gamma$ in these models are supposed to be unrestricted. Conditions required for the parameters $\theta$ and the sequence $f_t(\theta)$ are collected in the following set of assumptions which are partly taken from S&L.

**Assumption 1**

**(a)** The parameter space of $\theta$, denoted by $\Theta$, is a compact subset of $\mathbf{R}^m$.

**(b)** For each $t = 1, 2, \ldots, f_t(\theta)$ is a continuously differentiable function in an open set containing the parameter space $\Theta$ and, denoting by $F_t(\theta)$ the vector of all partial derivatives of $f_t(\theta)$,

$$\sup_T \sum_{t=1}^{T} \sup_{\theta \in \Theta} \|\Delta f_t(\theta)\| < \infty \quad \text{and} \quad \sup_T \sum_{t=1}^{T} \sup_{\theta \in \Theta} \|\Delta F_t(\theta)\| < \infty$$

where $f_0(\theta) = 0$ and $F_0(\theta) = 0$.

**(c)** Defining $g_t(\theta) = [1 : f_t(\theta)']'$ for $t = 1, 2, \ldots,$ with $\Delta g_1(\theta) = [1 : f_1(\theta)']'$, and $G_t(\theta) = [f_t(\theta)' : F_t(\theta)']'$ for $t = 1, 2, \ldots,$ there exists a real number $\epsilon > 0$ and an integer $T_*$ such that, for all $T \geq T_*$,

$$\inf_{\theta \in \Theta} \lambda_{min} \left\{ \sum_{t=1}^{T} \Delta g_t(\theta) \Delta g_t(\theta)' \right\} \geq \epsilon, \qquad \inf_{\theta \in \Theta} \lambda_{min} \left\{ \sum_{t=2}^{T} \Delta f_t(\theta) \Delta f_t(\theta)' \right\} \geq \epsilon$$

and

$$\inf_{\theta \in \Theta} \lambda_{min} \left\{ \sum_{t=2}^{T} \Delta G_t(\theta) \Delta G_t(\theta)' \right\} \geq \epsilon.$$

□

As mentioned earlier, some of these conditions are just repeated from S&L. The extensions are conditions for the partial derivatives of $f_t(\theta)$. They are used here to accommodate the modifications of the estimation procedures and unit root tests considered in the following sections. A compact parameter space $\Theta$ and the continuity requirement in Assumption 1(b) are standard assumptions in nonlinear estimation and testing problems. Furthermore, the summability conditions in Assumption 1(b) are now needed for the function $f_t(\theta)$ and its partial derivatives $F_t(\theta)$. They hold in the applications we have in mind, if the parameter space $\Theta$ is defined in a suitable way. Therefore the condition is not critical for our purposes. The conditions in Assumption 1(b) and (c) are formulated for differences of the sequences $f_t(\theta)$, $g_t(\theta)$ and the partial derivatives because our aim is to study unit root tests. Hence, estimation of the parameters $\mu, \theta$ and $\gamma$ is considered under the null hypothesis that the error process contains a unit root. Efficient estimation then requires that the variables are differenced.

To understand Assumption 1(c), assume first that the value of the parameter $\theta$ is known and that the parameters $\mu$ and $\gamma$ are estimated by applying OLS to the differenced models. Then these assumptions guarantee linear independence of the regressors when $T$ is large enough. There is of course no need to include the infimum in the condition of Assumption 1(c) if $\theta$ is known. It is needed, however, when the value of $\theta$ is unknown and has to be estimated. We have to impose an assumption which guarantees that the above mentioned linear independence of regressors holds whatever the value of $\theta$ because consistent estimation of $\theta$ is not possible. This is the purpose of Assumption 1(c).

Consistent estimation of $\theta$ and $\gamma$ is not possible because, by Assumption 1(b), the variation of (the differenced) regressors does not increase as $T \to \infty$. The present formulation of Assumption 1(b) also applies when the sequence $f_t(\theta)$ and hence $g_t(\theta)$ depends on $T$ which may be convenient occasionally. This feature is not made explicit in stating the assumption because it is not needed in the present application of Assumption 1 although it may sometimes be useful to allow the shift function to depend on $T$.

In the terminology of Elliott, Rothenberg & Stock (1996, Condition B), our assumptions imply that, for each value of $\theta$, the sequence $g_t(\theta)$ defines a slowly evolving trend, although our conditions are stronger than those of Elliott et al.. No attempt has been made here to weaken Assumption 1 because it is convenient for our purposes and applies to the models of interest in the following. More discussion of Assumption 1 is given in S&L.

We compare unit root tests within the models (2.1) and (2.2). More precisely, we consider tests of the pair of hypotheses

$$H_0 : \rho = 1 \qquad \text{vs.} \qquad H_1 : |\rho| < 1.$$

The idea is to estimate the nuisance parameters, in particular those related to the deterministic part, first and then remove the deterministic part and perform a test on the adjusted series. In the next section we therefore discuss estimation of the nuisance parameters.

# 3 Alternative Estimators of Nuisance Parameters

## 3.1 Model 2.1

Suppose that the process $x_t$ specified in (2.1$b$) is near integrated so that

$$\rho = \rho_T = 1 + \frac{c}{T}, \tag{3.1}$$

where $c \leq 0$ is a fixed real number. The estimation procedure proposed by S&L employs an empirical counterpart of the parameter $c$. This means that we shall replace $c$ by a chosen value $\bar{c}$ and pretend that $\bar{c} = c$ although we do not assume that this presumption is actually true. The idea is to apply a GLS procedure by first transforming the variables in (2.1) by the filter $1 - \bar{\rho}_T L$ where $\bar{\rho}_T = 1 + \frac{\bar{c}}{T}$ and then applying GLS to the transformed model. The choice of $\bar{c}$ will be discussed later.

For convenience we will use matrix notation and define

$$Y = [y_1 : (y_2 - \bar{\rho}_T y_1) : \cdots : (y_T - \bar{\rho}_T y_{T-1})]', \tag{3.2$a$}$$

$$Z_1 = \begin{bmatrix} 1 & 1 - \bar{\rho}_T & \cdots & 1 - \bar{\rho}_T \\ 1 & (2 - \bar{\rho}_T) & \cdots & (T - \bar{\rho}_T(T-1)) \end{bmatrix}' \tag{3.2$b$}$$

and

$$Z_2(\theta) = [f_1(\theta) : f_2(\theta) - \bar{\rho}_T f_1(\theta) : \cdots : f_T(\theta) - \bar{\rho}_T f_{T-1}(\theta)]'. \tag{3.2$c$}$$

Here, for simplicity, the notation ignores the dependence of the quantities on the chosen value $\bar{c}$. Using this notation, the transformed form of (2.1) can be written as

$$Y = Z(\theta)\phi + U, \tag{3.3}$$

where $Z(\theta) = [Z_1 : Z_2(\theta)]$, $\phi = [\mu_0 : \mu_1 : \gamma']'$ and $U = [u_1 : \cdots : u_T]'$ is an error term such that $u_t = x_t - \bar{\rho}_T x_{t-1} = b(L)^{-1}\varepsilon_t + T^{-1}(c - \bar{c})x_{t-1}$. Our GLS estimation is based on the covariance matrix resulting from $b(L)^{-1}\varepsilon_t$, denoted by $\sigma^2\Sigma(b)$, where $b = [b_1 : \cdots : b_p]'$. The GLS estimators are thus obtained by minimizing the generalized sum of squares function

$$Q_T(\phi, \theta, b) = (Y - Z(\theta)\phi)'\Sigma(b)^{-1}(Y - Z(\theta)\phi). \tag{3.4}$$

The above estimation procedure makes use of the initial value assumption $x_0 = 0$. Although initial values which are independent of the sample size have asymptotically no effect the situation may be different in finite samples. Therefore we shall also consider an alternative approach which is free of this feature. In this approach the first observation in the transformed regression model (3.3) is omitted. Since the regressor corresponding to the level parameter $\mu_0$ will then be asymptotically zero we shall not try to estimate this parameter. Instead, after deleting the first observation from the regression model (3.3) we replace the first column of the regressor matrix by a vector of ones. This means that instead of $\mu_0$ we consider the parameter $\mu_0^* = \mu_0(1 - \bar{\rho}_T)$. Of course, if there is no estimate of $\mu_0$ it is then not possible to obtain a sample analog of the process $x_t$ but only of $x_t + \mu_0$ so that this feature has to be taken into account when unit root tests based on this approach are developed in Sec. 4.

Define

$$Y^* = [(y_2 - \bar{\rho}_T y_1) : \cdots : (y_T - \bar{\rho}_T y_{T-1})]', \tag{3.5a}$$

$$Z_1^* = \begin{bmatrix} 1 & \cdots & 1 \\ (2 - \bar{\rho}_T) & \cdots & (T - \bar{\rho}_T(T-1)) \end{bmatrix}' \tag{3.5b}$$

and

$$Z_2^*(\theta) = [(f_2(\theta) - \bar{\rho}_T f_1(\theta)) : \cdots : (f_T(\theta) - \bar{\rho}_T f_{T-1}(\theta))]'. \tag{3.5c}$$

Instead of (3.3) we have

$$Y^* = Z^*(\theta)\phi^* + U^*, \tag{3.6}$$

where $Z^*(\theta) = [Z_1^* : Z_2^*(\theta)]$, $\phi^* = [\mu_0^* : \mu_1 : \gamma']'$ and $U^* = [u_2 : \cdots : u_T]'$ with $u_t$ as before. Here, for simplicity, the dependence of the parameters $\mu_0^*$ and $\phi^*$ on the sample size has not been indicated. The GLS estimator of the parameter vector $\phi^*$ is obtained by minimizing the function

$$Q_T^*(\phi^*, \theta, b) = (Y^* - Z^*(\theta)\phi^*)'\Sigma^*(b)^{-1}(Y^* - Z^*(\theta)\phi^*), \tag{3.7}$$

where $\Sigma^*(b)$ is a $((T-1)\times(T-1))$ analog of the matrix $\Sigma(b)$. Notice that if $\bar{c}=0$ (or $\bar{\rho}_T=1$) the two columns of the matrix $Z_1^*$ are identical so that one of them can be deleted. In what follows the treatment of this problem is fairly obvious so that it will not be discussed any further. We will also consider the case where no linear trend term is present and, hence, $\mu_1=0$ a priori. In that case the corresponding column is deleted from the regression matrix.

Using Assumption 1(c) it is easy to see as in S&L that GLS estimators of the parameters $\phi^*$, $\theta$ and $b$ exist for all $T$ large enough. These GLS estimators are denoted by $\hat{\phi}^*$, $\hat{\theta}^*$ etc. Their asymptotic properties are given in Lemma A.1 in the Appendix.

## 3.2 Model 2.2

Suppose now that the error process $v_t$ specified in (2.2$b$) is near integrated so that $\rho=\rho_T$ is as in (3.1). Then the generating process of $v_t$ can be written as

$$v_t=\rho_T v_{t-1}+\varepsilon_t,\qquad t=1,2,\ldots \tag{3.8}$$

Again the first estimation procedure proposed by LMS employs an empirical counterpart of the parameter $c$ so that we shall replace $c$ by a chosen value $\bar{c}$ and pretend that $\bar{c}=c$. Now, if $\bar{\rho}_T=1+\frac{\bar{c}}{T}$, the idea is to first transform the variables in (2.2$a$) by the filter $1-\bar{\rho}_T L$. Using the matrix notation from (3.2), the transformed form of (2.2$a$) can be written as

$$Y=W(\theta)\beta+\mathcal{E}, \tag{3.9}$$

where $W(\theta)=[Z(\theta):V]$ with $Z(\theta)=[Z_1:Z_2(\theta)]$, as before, and $V$ the $(T\times p)$ matrix containing lagged values of $y_t$ transformed in the same way as the other variables. Furthermore, $\beta=[\mu_0:\mu_1:\gamma':b']'$ and $\mathcal{E}=[e_1:\cdots:e_T]'$ is an error term such that $e_t=v_t-\bar{\rho}_T v_{t-1}=\varepsilon_t+T^{-1}(c-\bar{c})v_{t-1}$. We shall consider a nonlinear OLS estimation of (3.9) by proceeding in the same way as in the case $c=0$, that is, $e_t=\varepsilon_t$ or under the null hypothesis. Our estimators are thus obtained by minimizing the sum of squares function

$$S_T(\theta,\beta)=(Y-W(\theta)\beta)'(Y-W(\theta)\beta). \tag{3.10}$$

The estimator of $\beta$ can be written as

$$\tilde{\beta}=(W(\tilde{\theta})'W(\tilde{\theta}))^{-1}W(\tilde{\theta})'Y, \tag{3.11}$$

where $\tilde{\theta}$ is the value of $\theta$ which minimizes (3.10) jointly with $\tilde{\beta}$.

In the same way as above one may also wish to consider a modification of the above approach to avoid potential adverse finite sample effects of unrealistic initial value assumptions. Thus, we define $Y^*$, $Z_1^*$ and $Z_2^*(\theta)$ as before and the $((T-1) \times p)$ matrix $V^*$ and the $((T-1) \times 1)$ vector $\mathcal{E}^*$ are defined by deleting the first row and first component from $V$ and $\mathcal{E}$, respectively. Instead of (3.9) we now consider

$$Y^* = W^*(\theta)\beta^* + \mathcal{E}^*, \tag{3.12}$$

where $W^*(\theta) = [Z^*(\theta) : V^*]$ with $Z^*(\theta) = [Z_1 : Z_2^*(\theta)]$ and $\beta^* = [\mu_0^* : \mu_1 : \gamma' : b']'$ with $\mu_0^* = \mu_0(1 - \bar{\rho}_T)$. For simplicity the notation again ignores the dependence of the quantities on the chosen value of $\bar{c}$ and on the sample size. Thus, in this approach we do not try to estimate the parameter $\mu_0$. This means that we cannot obtain an empirical counterpart of the process $v_t$ but only of $v_t + \mu_0$. This feature will be taken into account in constructing unit root tests in the next section.

We estimate the parameters $\beta^*$ and $\theta$ in (3.12) by minimizing the obvious analog of the sum of squares function in (3.10). If $\bar{c} = 0$ (or $\bar{\rho}_T = 1$) the two columns of the matrix $Z_1^*$ are identical so that the regression model (3.12) is not of full column rank. Then we shall delete the first column of $Z_1^*$ and accordingly delete $\mu_0^*$ from the parameter vector $\beta^*$. Since the treatment of this special case is fairly obvious it will not be discussed here in more detail. Again we may also impose the restriction $\mu_1 = 0$ if a linear trend term is not needed.

# 4 The Tests

## 4.1 Model 2.1

We consider the following tests. Once the nuisance parameters in (2.1) have been estimated one can form the residual series $\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - f_t(\hat{\theta})'\hat{\gamma}$ $(t = 1, \ldots, T)$ and use it to obtain unit root tests. S&L propose the following procedure.

Consider the auxiliary regression model

$$\hat{x}_t = \rho \hat{x}_{t-1} + u_t^*, \qquad t = 2, \ldots, T. \tag{4.1}$$

In the previous section it was seen that if $\hat{x}_t$ is replaced by $x_t$ the covariance matrix of the error term in (4.1) is $\sigma^2\Sigma^*(b)$. Since the parameter vector $b$ is estimated to obtain $\hat{x}_t$ it seems reasonable to use this estimator also here and base a unit root test on (4.1) with $\rho$ estimated by feasible GLS with weight matrix $\Sigma^*(\hat{b})^{-1}$. We denote the usual $t$-statistic for testing the null hypothesis $\rho = 1$ associated with the feasible GLS estimator of $\rho$ by $\tau_{S\&L}$ because it is the statistic considered by S&L except that these authors use residuals $\hat{x}_t$ for $t = 1, \dots, T$ in (4.1) with initial value $\hat{x}_0 = 0$.

The error term in the auxiliary regression model (4.1) also contains estimation errors caused by replacing the nuisance parameters $\mu_0$, $\mu_1$, $\theta$ and $\gamma$ by their GLS estimators. Being able to allow for the effect of these estimation errors might improve the finite sample properties of the above test and particularly the performance of the asymptotic size approximation. To investigate this, consider the special case where the shift function is a step dummy variable $f_t(\theta) = d_{1t}$ which is zero up to period $T_1 - 1$ and one from period $T_1$ onwards. Suppose that the null hypothesis holds. Then it is straightforward to check that

$$u_t^* = \Delta x_t - (\hat{\mu}_1 - \mu_1) - \Delta d_{1t} (\hat{\gamma} - \gamma), \qquad t = 2, \dots, T.$$

Thus, augmenting the auxiliary regression model (4.1) by an intercept term and the impulse dummy $\Delta d_{1t}$ would result in an error term which, under the null hypothesis, would not depend on the errors caused by estimating the nuisance parameters $\mu_1$ and $\gamma$. It is fairly obvious that the inclusion of the impulse dummy $\Delta d_{1t}$ has no effect on the asymptotic properties of the GLS estimator of the parameter $\rho$ and, consequently, on the limiting distribution of the resulting test. Below we will see that the inclusion of an intercept term results in a different limiting distribution. Therefore, we will consider tests with and without intercept in the following.

If the step dummy $d_{1t}$ is replaced by the general function $f_t(\theta)$ the above modification becomes slightly more complicated. We then have

$$\begin{aligned} u_t^* &= \Delta x_t - (\hat{\mu}_1 - \mu_1) - \Delta f_t(\hat{\theta})'\hat{\gamma} + \Delta f_t(\theta)'\gamma \\ &= \Delta x_t - (\hat{\mu}_1 - \mu_1) - \Delta f_t(\hat{\theta})'(\hat{\gamma} - \gamma) - \left(\Delta f_t(\hat{\theta}) - \Delta f_t(\theta)\right)' \gamma, \qquad t = 2, \dots, T. \end{aligned} \tag{4.2}$$

In the last expression the third term can be handled in the same way as in the previously considered case of a step dummy but the fourth term requires additional considerations. A fairly obvious approach is to assume that the function $f_t(\theta)$ is continuously differentiable in an open set containing the parameter space $\Theta$ and use the Taylor series approximation $\Delta f_t(\hat{\theta}) - \Delta f_t(\theta) \approx \Delta \left(\partial f_t(\hat{\theta})/\partial\theta'\right)(\hat{\theta} - \theta)$. Instead of (4.1) we then consider the auxiliary regression model

$$\hat{x}_t = \rho \hat{x}_{t-1} + \Delta f_t(\hat{\theta})'\pi_1 + \Delta F_t(\hat{\theta})'\pi_2 + u_t^\dagger, \qquad t = 2, \ldots, T, \tag{4.3}$$

where $F_t(\hat{\theta})$ is a $(mk \times 1)$ vector containing the partial derivatives in $\partial f_t(\hat{\theta})/\partial\theta$. Let $\tau_{adj}$ be the usual '$t$-statistic' based on the GLS estimation of the parameters in (4.3) with weight matrix $\Sigma^*(\hat{b})^{-1}$. Here the subscript indicates that the statistic is obtained from the *adj*usted auxiliary regression model.

In these tests we still do not make adjustments for the fact that the $b$ parameters are also estimated. A possible modification that adjusts for the estimation of $b$ may be obtained as follows. Define $w_t = b(L)x_t$ so that $w_t = \rho w_{t-1} + \varepsilon_t$. Thus, if we condition on $y_1, \ldots, y_p$, a version of the test statistic $\tau_{S\&L}$ may be obtained from the auxiliary regression model

$$\hat{w}_t = \rho \hat{w}_{t-1} + error_t, \qquad t = p+1, \ldots, T,$$

where $\hat{w}_t = \hat{b}(L)\hat{x}_t$.

Now, to obtain a modification which takes into account estimation errors in $\hat{b}$, consider the identity

$$\begin{aligned} \hat{w}_t &= w_t + \hat{b}(L)\hat{x}_t - b(L)x_t \\ &= w_t + \hat{b}(L)(\hat{x}_t - x_t) + (\hat{b}(L) - b(L))\hat{x}_t - (\hat{b}(L) - b(L))(\hat{x}_t - x_t), \quad t = p+1, \ldots, T. \end{aligned}$$

Multiplying both sides of this equation by $\rho(L)$ and observing that $\rho(L)w_t = \varepsilon_t$ yields

$$\hat{w}_t = \rho \hat{w}_{t-1} + \rho(L)\hat{b}(L)(\hat{x}_t - x_t) + \sum_{j=1}^{p}(\hat{b}_j - b_j)\rho(L)\hat{x}_{t-j} + r_t, \qquad t = p+2, \ldots, T,$$

where $r_t = \varepsilon_t - (\hat{b}(L) - b(L))\rho(L)(\hat{x}_t - x_t)$ is an error term. Since we try to improve the size performance of the test statistic $\tau_{S\&L}$ we now assume that the null hypothesis holds and replace $\rho(L)$ on the r.h.s. by $\Delta$. Thus, we consider the auxiliary regression model

$$\hat{w}_t = \rho \hat{w}_{t-1} + \hat{b}(L)(\Delta\hat{x}_t - \Delta x_t) + \sum_{j=1}^{p}(\hat{b}_j - b_j)\Delta\hat{x}_{t-j} + r_t, \qquad t = p+2, \ldots, T.$$

Note that estimation errors in $r_t$ are expected to be smaller than those in the second and third terms on the r.h.s. of this equation because, under $H_0$, they are affected through the product $(\hat{b}(L) - b(L))(\Delta\hat{x}_t - \Delta x_t)$ only. To be able to use this auxiliary model we still have to deal with the second term on the r.h.s.. This, however, leads to considerations very similar to those in the previous modifications and expanding the difference $\Delta\hat{x}_t - \Delta x_t$ we get the auxiliary model

$$\hat{w}_t = \rho\hat{w}_{t-1} + [\hat{b}(L)\Delta f_t(\hat{\theta})']\pi_1 + [\hat{b}(L)\Delta F_t(\hat{\theta})']\pi_2 + \sum_{j=1}^{p}\alpha_j\Delta\hat{x}_{t-j} + r_t^{\dagger}, \quad t = p+2,\ldots,T. \quad (4.4)$$

The modified test statistic is obtained as the usual $t$-statistic for the hypothesis $\rho = 1$ based on OLS estimation of this model. It will be denoted by $\tau_{adj}^{+}$.

Because the actual mean of the $\hat{x}_t$ may be nonzero, it may be reasonable to include an intercept term in the previously considered auxiliary regressions. For instance, instead of (4.3) we may consider

$$\hat{x}_t = \nu + \rho\hat{x}_{t-1} + \Delta f_t(\hat{\theta})'\pi_1 + \Delta F_t(\hat{\theta})'\pi_2 + u_t^{+}, \qquad t = 2,\ldots,T. \quad (4.5)$$

The relevant unit root $t$-statistic will be denoted by $\tau_{int}$, where the subscript indicates that an *int*ercept is included in the model. Similarly, if an intercept term is added to (4.4), the resulting unit root test statistic will be denoted by $\tau_{int}^{+}$.

Because $x_t$ is a zero mean process it is not unreasonable to set $x_0 = 0$ in estimating the parameters of the deterministic part. However, in small samples it may be preferable to avoid such an assumption because the actual values of the process may be different from zero. Therefore it may be useful to consider the estimators based on model (3.6) and the series $\hat{x}_t^* = y_t - \hat{\mu}_1^* t - f_t(\hat{\theta}^*)'\gamma^*$ $(t = 2,\ldots,T)$. The theoretical counterpart of $\hat{x}_t^*$ is $x_t^* = x_t + \mu_0$ for which we have $x_t^* = \nu + \rho x_{t-1}^* + u_t^{(0)}$ $(t = 1, 2, \ldots)$, where $\nu = (1-\rho)\mu_0$. Thus, in this approach our unit root tests are based on the auxiliary regression model

$$\hat{x}_t^* = \nu + \rho\hat{x}_{t-1}^* + u_t^{**}, \qquad t = 2,\ldots,T. \quad (4.6)$$

The resulting test statistic based on feasible GLS estimation of this model will be denoted by $\tau_{S\&L}^*$.

It is also possible to include terms to take care of estimation errors and base the unit root test on an auxiliary regression similar to (4.5),

$$\hat{x}_t^* = \nu + \rho\hat{x}_{t-1}^* + \Delta f_t(\hat{\theta}^*)'\pi_1 + \Delta F_t(\hat{\theta}^*)'\pi_2 + u_t^{+*}, \qquad t = 2,\ldots,T. \quad (4.7)$$

The resulting unit root test statistic will be denoted by $\tau_{int}^*$. If in addition we condition on $y_1,\ldots,y_p$ and use a model corresponding to (4.4) with $\hat{w}_t^* = \hat{b}^*(L)\hat{x}_t^*$ and similar modifications for the other terms the resulting test statistic will be denoted by $\tau_{int}^{*+}$ (see also Table 1). Note that $\hat{b}^*$ is the estimator of $b$ obtained from minimizing (3.7).

Moreover, if we have the a priori restriction $\mu_1 = 0$ the estimation procedures in Section 3 and the definitions of $\hat{x}_t$ and $\hat{x}_t^*$ are adjusted accordingly. Since in this case the limiting distributions of the corresponding unit root tests change, we augment the test statistics with a superscript 0 to distinguish them from the statistics which allow for a linear time trend. In other words, the test statistics based on the restriction $\mu_1 = 0$ are denoted as $\tau_{S\&L}^0$, $\tau_{adj}^0$, $\tau_{adj}^{+0}$, $\tau_{int}^0$, $\tau_{int}^{+0}$, $\tau_{S\&L}^{*0}$, $\tau_{int}^{*0}$ and $\tau_{int}^{*+0}$, respectively. The limiting null distributions of all the test statistics are given in the following theorem which is partly proven in the Appendix and partly reviews results from the related literature.

**Theorem 1.**

Suppose that Assumption 1 holds and that the matrices $Z(\theta)$ and $Z^*(\theta)$ are of full column rank for all $T \geq k+1$ and all $\theta \in \Theta$. Then,

$$\tau_{S\&L}^0, \tau_{adj}^0, \tau_{adj}^{+0} \xrightarrow{d} \left( \int_0^1 B_c(s)^2 ds \right)^{-1/2} \int_0^1 B_c(s) dB_c(s), \tag{4.8}$$

where $B_c(s) = \int_0^s \exp\{c(s-u)\} dB_0(u)$ with $B_0(u)$ a standard Brownian motion,

$$\tau_{int}^0, \tau_{int}^{+0}, \tau_{S\&L}^{*0}, \tau_{int}^{*0}, \tau_{int}^{*+0} \xrightarrow{d} \left( \int_0^1 \bar{B}_c(s)^2 ds \right)^{-1/2} \int_0^1 \bar{B}_c(s) dB_c(s), \tag{4.9}$$

where $\bar{B}_c(s)$ is the mean-adjusted version of $B_c(s)$,

$$\tau_{S\&L}, \tau_{adj}, \tau_{adj}^{+} \xrightarrow{d} \left( \int_0^1 G_c(s;\bar{c})^2 ds \right)^{-1/2} \int_0^1 G_c(s;\bar{c}) dG_c(s;\bar{c}), \tag{4.10}$$

where $G_c(s;\bar{c}) = B_c(s) - sK_c(\bar{c})$ with

$$K_c(\bar{c}) = h(\bar{c})^{-1} \int_0^1 (1-\bar{c}s) dB_0(s) + h(\bar{c})^{-1}(c-\bar{c}) \int_0^1 (1-\bar{c}s) B_c(s) ds$$

and $h(\bar{c}) = 1-\bar{c}+\bar{c}^2/3$. Here the stochastic integral is a short-hand notation for $\int_0^1 G_c(s;\bar{c}) dB_c(s) - K_c(\bar{c}) \int_0^1 G_c(s;\bar{c}) ds$. Moreover,

$$\tau_{int}, \tau_{int}^{+} \xrightarrow{d} \left( \int_0^1 \bar{G}_c(s;\bar{c})^2 ds \right)^{-1/2} \int_0^1 \bar{G}_c(s;\bar{c}) dG_c(s;\bar{c}), \tag{4.11}$$

where $\bar{G}_c(s;\bar{c})$ is a mean-adjusted version of $G_c(s;\bar{c})$. Furthermore,

$$\tau_{S\&L}^*, \tau_{int}^*, \tau_{int}^{*+} \xrightarrow{d} \left( \int_0^1 \bar{G}_c^*(s;\bar{c})^2 ds \right)^{-1/2} \int_0^1 \bar{G}_c^*(s;\bar{c}) dG_c^*(s;\bar{c}) \tag{4.12}$$

where $G_c^*(s;\bar{c}) = B_c(s) - sK_c^*(\bar{c})$, with $K_c^*(0) = B_c(1)$ and, for $\bar{c} < 0$,

$$K_c^*(\bar{c}) = \frac{12}{\bar{c}} \int_0^1 \left(\frac{1}{2} - s\right) dB_0(s) + \frac{12(c-\bar{c})}{\bar{c}} \int_0^1 \left(\frac{1}{2} - s\right) B_c(s)ds,$$

$\bar{G}_s^*(s;\bar{c})$ is a mean-adjusted version of $G_s^*(s;\bar{c})$ and the stochastic integral is a short-hand notation for $\int_0^1 \bar{G}_c^*(s;\bar{c})dB_c(s) - K_c^*(\bar{c}) \int_0^1 \bar{G}_c^*(s;\bar{c})ds$. □

Notice that for $c = 0$ the null distributions in (4.8) and (4.9) are conventional Dickey-Fuller (DF) distributions for unit root tests in models without deterministic terms and with intercept, respectively. The distribution in (4.10) was given by S&L for the statistic $\tau_{S\&L}$ in the form

$$\frac{1}{2}\left(\int_0^1 G_c(s;\bar{c})^2 ds\right)^{-1/2} (G_c(1;\bar{c})^2 - 1),$$

where

$$G_c(s;\bar{c}) = B_c(s) - s\left(\lambda B_c(1) + 3(1-\lambda)\int_0^1 sB_c(s)ds\right)$$

with $\lambda = (1-\bar{c})/h(\bar{c})$. It can be shown that this limiting distribution is equivalent to the one in (4.10) (see the Appendix). We use the latter version now because it facilitates a comparison with the other limiting distributions given in the theorem.

The limiting null distribution of the test statistics $\tau_{int}$ and $\tau_{int}^+$ are again obtained by setting $c = 0$. It is free of unknown nuisance parameters but depends on the quantity $\bar{c}$. It differs from that of $\tau_{S\&L}$, $\tau_{adj}$ and $\tau_{adj}^+$ in that $G_c(s;\bar{c})$ is replaced by a mean-adjusted version. This difference is due to the intercept term included in the auxiliary regression model (4.5). In this sense $\tau_{int}$ may be called a "mean-adjusted version" of $\tau_{adj}$ etc.

Obviously, the asymptotic distribution of the test statistics $\tau_{S\&L}^*, \tau_{int}^*$ and $\tau_{int}^{*+}$ also differs from the other ones. Instead of $G_c(s;\bar{c})$ in (4.11) we have $G_c^*(s;\bar{c})$ in (4.12). The difference between these two quantities is due to the different limiting distributions of the estimators $\hat{\mu}_1$ and $\hat{\mu}_1^*$. This difference results from a different treatment of the intercept term in the regression models (3.3) and (3.6) and in the special case $\bar{c} = 0$ this difference vanishes.

To the best of our knowledge the asymptotic distributions in (4.11) and (4.12) have not been studied previously so that critical values and suggestions for appropriate values of $\bar{c}$ are not available. Thus, simulations are required to make the test statistics $\tau_{int}$, $\tau_{S\&L}^*$, $\tau_{int}^*$ and their relatives applicable and to study their power properties. Even without such simulations it is clear, however, that in terms of asymptotic local power the test statistics in (4.11) and

(4.12) are inferior to those in (4.10) because they are not asymptotically equivalent to $\tau_{S\&L}$ and the asymptotic local power of $\tau_{S\&L}$ is indistinguishable from optimal (see Elliott et al. (1996)). However, since this result is based on an initial value assumption which may be unrealistic in some cases (see Elliott et al. (1996, pp. 819-820)) the performance of the $\tau_{int}$, $\tau^*_{S\&L}$ and $\tau^*_{int}$ tests may be preferable in some finite sample situations. We will provide critical values, local power results and small sample comparisons for these tests in the following sections. Before that we shall discuss unit root tests in the framework of model (2.2).

## 4.2 Model 2.2

Once the nuisance parameters in (2.2$a$) have been estimated, the residual series $\tilde{v}_t = \tilde{b}(L)y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - f_t(\tilde{\theta})'\tilde{\gamma}$ may be used to obtain unit root tests. There are several possible choices. LMS suggest using DF $t$-tests like, for instance, Elliott et al. (1996). In the following we shall also consider these tests.

Consider the auxiliary regression model

$$\tilde{v}_t = \rho\tilde{v}_{t-1} + e^*_t, \qquad t = 2, \ldots, T. \tag{4.13}$$

If $\tilde{v}_t$ is replaced by $v_t$ the error term in (4.13) becomes $\varepsilon_t$ so that we can use OLS to obtain a test statistic. LMS consider the usual $t$-statistic for testing $\rho = 1$ in (4.13). In the following this statistic will be denoted by $\mathbf{t}_{LMS}$. Note that LMS use the model (4.13) for $t = 1, \ldots, T$ with $\tilde{v}_0 = 0$.

In the same way as in the previous subsection a modification of the test statistic can be considered which takes into account that the error term in the auxiliary regression model (4.13) also contains estimation errors caused by replacing the nuisance parameters $b = [b_1 : \cdots : b_p]'$, $\mu_0$, $\mu_1$, $\theta$ and $\gamma$ by their OLS estimators. As far as the finite sample properties of the above test and particularly the performance of the asymptotic size approximation are concerned it might therefore be worthwhile to try to allow for this feature. To investigate this possibility, suppose the null hypothesis holds and note that, by straightforward calculation, one can readily see that

$$e^*_t = \varepsilon_t - [\tilde{b}(L) - b(L)]\Delta y_t - (\tilde{\mu}_1 - \mu_1) - [\Delta f_t(\tilde{\theta})'\tilde{\gamma} - \Delta f_t(\theta)'\gamma], \quad t = 2, \ldots, T. \tag{4.14}$$

For simplicity, consider first the special case where the function $f_t(\theta)$ is defined by the step dummy $d_{1t}$ so that it is independent of the parameter $\theta$. As is clear from equation

(4.14), the estimation errors caused by using estimators of nuisance parameters can then be allowed for by augmenting the auxiliary regression model (4.13) by the impulse dummy $\Delta d_{1t}$, the lagged differences $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$, and an intercept term. After this the test statistic can be defined on the basis of the OLS estimator of $\rho$ in the same way as before. The inclusion of an impulse dummy in (4.13) will not change the limiting distribution of the resulting unit root test but the inclusion of an intercept term does. We shall consider both modifications. It should be noted, however, that since the mean value of the lagged differences $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$ is generally nonzero the inclusion of these variables as additional regressors in (4.13) will change the limiting distribution of the resulting unit root test. The reason is that these lagged differences are not asymptotically orthogonal to the variable $\tilde{v}_{t-1}$. It turns out, however, that this feature can be allowed for by using the mean-adjusted variables $\Delta y_{t-j} - \tilde{\mu}_*$ $(j = 1, \ldots, p)$ where $\tilde{\mu}_* = \tilde{\mu}_1 / \tilde{b}(1)$.

When the function $f_t(\theta)$ depends on the parameter vector $\theta$ the treatment of the fourth term on the right hand side of (4.14) becomes slightly more complicated than in the foregoing special case. In the same way as in Subsection 4.1 we shall then assume that the function $f_t(\theta)$ is continuously differentiable in an open set containing the parameter space $\Theta$ and use the Taylor series approximation $\Delta f_t(\tilde{\theta}) - \Delta f_t(\theta) \approx \Delta(\partial f_t(\tilde{\theta})/\partial\theta')(\tilde{\theta} - \theta)$. Thus, instead of (4.13) we shall consider the auxiliary regression model

$$\tilde{v}_t = \rho\tilde{v}_{t-1} + \Delta f_t(\tilde{\theta})'\pi_1 + \Delta F_t(\tilde{\theta})'\pi_2 + \tilde{q}_t'\pi_3 + e_t^\dagger, \quad t = 2, \ldots, T, \tag{4.15}$$

where $F_t(\tilde{\theta})$ is a $(mk \times 1)$ vector containing the partial derivatives in the matrix $\partial f_t(\tilde{\theta})/\partial\theta$ and $\tilde{q}_t' = [\Delta y_{t-1} - \tilde{\mu}_* : \cdots : \Delta y_{t-p} - \tilde{\mu}_*]$. Let $\mathbf{t}_{adj}$ be the usual $t$-statistic for the null hypothesis $\rho = 1$ based on the OLS estimator of $\rho$ in (4.15).

Including an intercept term in the auxiliary regression gives

$$\tilde{v}_t = \nu + \rho\tilde{v}_{t-1} + \Delta f_t(\tilde{\theta})'\pi_1 + \Delta F_t(\tilde{\theta})'\pi_2 + \tilde{q}_t'\pi_3 + e_t^\dagger, \quad t = 2, \ldots, T, \tag{4.16}$$

and the relevant $t$-statistic will be denoted by $\mathbf{t}_{int}$.

Using the estimators $\tilde{b}^*$, $\tilde{\mu}_1^*$, $\tilde{\gamma}^*$ and $\tilde{\theta}^*$ we can form the series $\tilde{v}_t^* = \tilde{b}^*(L)y_t - \tilde{\mu}_1^* t - f_t(\tilde{\theta}^*)'\tilde{\gamma}$ $(t = 2, \ldots, T)$. Its theoretical counterpart is $v_t^* = v_t + \mu_0$ for which we have $v_t^* = \nu + \rho v_{t-1}^* + \varepsilon_t$ where $\nu = (1 - \rho)\mu_0$. Thus, in this approach our unit root tests are based on the auxiliary regression model

$$\tilde{v}_t^* = \nu + \rho\tilde{v}_{t-1}^* + e_t^{**}, \quad t = 2, \ldots, T. \tag{4.17}$$

Our test statistic, denoted by $\mathbf{t}^*_{LMS}$, is the $t$-statistic for the null hypothesis $\rho = 1$ in (4.17) based on OLS estimation.

It is also possible to include terms to take care of estimation errors and base the unit root test on an auxiliary regression similar to (4.16),

$$\tilde{v}^*_t = \nu + \rho\tilde{v}^*_{t-1} + \Delta f_t(\tilde{\theta}^*)'\pi_1 + \Delta F_t(\tilde{\theta}^*)'\pi_2 + \tilde{q}^{*\prime}_t\pi_3 + e^{*\dagger}_t, \quad t = 2, \ldots, T, \tag{4.18}$$

The resulting unit root test statistic will be denoted by $\mathbf{t}^*_{int}$.

Moreover, if we have the a priori restriction $\mu_1 = 0$ the estimation procedures in Section 3 and the definitions of $\tilde{v}_t$ and $\tilde{v}^*_t$ are adjusted accordingly. Since in this case the limiting distributions of the corresponding unit root tests change, we augment the test statistics with a superscript 0 to distinguish them from the statistics which allow for a linear time trend. In other words, the test statistics are denoted by $\mathbf{t}^0_{LMS}$, $\mathbf{t}^0_{adj}$, $\mathbf{t}^0_{int}$, $\mathbf{t}^{*0}_{LMS}$ and $\mathbf{t}^{*0}_{int}$, respectively. The limiting null distributions of all the test statistics are given in the following theorem which is also partly proven in the Appendix.

**Theorem 2.**
Suppose that Assumption 1 holds and that the matrices $Z(\theta)$ and $Z^*(\theta)$ are of full column rank for all $T \geq k+1$ and all $\theta \in \Theta$. Then, using the notation of Theorem 1,

$$\mathbf{t}^0_{LMS}, \mathbf{t}^0_{adj} \xrightarrow{d} \left(\int_0^1 B_c(s)^2 ds\right)^{-1/2} \int_0^1 B_c(s) dB_c(s), \tag{4.19}$$

$$\mathbf{t}^0_{int}, \mathbf{t}^{*0}_{LMS}, \mathbf{t}^{*0}_{int} \xrightarrow{d} \left(\int_0^1 \bar{B}_c(s)^2 ds\right)^{-1/2} \int_0^1 \bar{B}_c(s) dB_c(s), \tag{4.20}$$

$$\mathbf{t}_{LMS}, \mathbf{t}_{adj} \xrightarrow{d} \left(\int_0^1 G_c(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 G_c(s;\bar{c}) dG_c(s;\bar{c}), \tag{4.21}$$

$$\mathbf{t}_{int} \xrightarrow{d} \left(\int_0^1 \bar{G}_c(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 \bar{G}_c(s;\bar{c}) dG_c(s;\bar{c}) \tag{4.22}$$

and

$$\mathbf{t}^*_{LMS}, \mathbf{t}^*_{int} \xrightarrow{d} \left(\int_0^1 \bar{G}^*_c(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 \bar{G}^*_c(s;\bar{c}) dG^*_c(s;\bar{c}). \tag{4.23}$$

□

Thus, the $\mathbf{t}$ statistics have the same asymptotic distributions as the corresponding $\tau$ statistics in the previous subsection. For both models, alternative approaches such as point optimal tests are possible in the present context. For instance, if the auxiliary regression model (4.13) is used as a starting point these tests would be based on the statistics $\hat{\sigma}^2(1)$ and $\hat{\sigma}^2(\bar{\rho}_T)$ defined by replacing $\tilde{\rho}$ in the variance estimator by unity and $\bar{\rho}_T$, respectively. According to the simulation results of Elliott et al. (1996) the overall properties of their DF $t$-statistic appeared somewhat better than those of the point optimal tests. Therefore we use the DF test versions. It may also be worth noting that seasonal dummies may be included without affecting the limiting distributions of our tests.

# 5 Critical Values and Local Power Simulations

All the tests considered in the previous section are summarized in Table 1 for the case where no a priori restriction is available for $\mu_1$. In order to investigate the null distributions and local power of the test statistics we have generated time series

$$x_t = \rho_T x_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \quad x_0 = 0, \quad \rho_T = 1 + c/T, \quad \varepsilon_t \sim iid\, N(0, 1). \qquad (5.1)$$

Thus, $p = 0$ so that there is no additional dynamics. Moreover, there is no deterministic part and we can use the generated series to investigate the tests with and without the restriction $\mu_1 = 0$. For this purpose we use again $\bar{\rho}_T = 1 + \bar{c}/T$ and consider the following $\hat{x}_t$ series:

- $\hat{x}_t^{(0)} = x_t - \hat{\mu}_0$ $(t = 1, \ldots, T)$, where $\hat{\mu}_0$ is obtained from a regression $(1 - \bar{\rho}_T L)x_t = \mu_0 z_{0t} + error_t$ $(t = 1, \ldots, T)$ with

$$z_{0t} = \begin{cases} 1, & t = 1, \\ 1 - \bar{\rho}_T, & t = 2, \ldots, T, \end{cases}$$

- $\hat{x}_t^{(1)} = x_t - \hat{\mu}_0 - \hat{\mu}_1 t$ $(t = 1, \ldots, T)$, where $\hat{\mu}_0$ and $\hat{\mu}_1$ are obtained from a regression $(1 - \bar{\rho}_T L)x_t = \mu_0 z_{0t} + \mu_1(t - \bar{\rho}_T(t-1)) + error_t$ $(t = 1, \ldots, T)$.

Moreover, the $\hat{x}_t^*$ series are obtained as:

- $\hat{x}_t^{*(0)} = x_t$ $(t = 1, \ldots, T)$

- $\hat{x}_t^{*(1)} = x_t - \hat{\mu}_1 t$ $(t = 1, \ldots, T)$, where $\hat{\mu}_1$ is obtained from a regression $(1 - \bar{\rho}_T L)x_t = \nu_0 + \mu_1(t - \bar{\rho}_T(t-1)) + error_t$ $(t = 2, \ldots, T)$.

The series $\hat{x}_t^{(i)}$ $(i = 0, 1)$ are used to compute $t$-statistics for the null hypothesis $\rho = 1$ based on the regression model (4.1), the series $\hat{x}_t^{(1)}$ is also used to compute the $t$-statistic for $\rho = 1$ in $\hat{x}_t^{(1)} = \nu + \rho\hat{x}_{t-1}^{(1)} + u_t^*$ and the series $\hat{x}_t^{*(i)}$ $(i = 0, 1)$ are used for the same purpose in conjunction with model (4.6). For large sample size $T$ and $c = 0$ (i.e., $\rho_T = 1$) we get realizations of the null distributions corresponding to (4.8) - (4.12) and, hence, of (4.19) - (4.23) in this way.

**Table 1.** Summary of Tests

| Test statistic | Underlying auxiliary regression |
|---|---|
| | Asymptotic distribution $\left(\int_0^1 G_c(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 G_c(s;\bar{c}) dG_c(s;\bar{c})$ |
| $\tau_{S\&L}$ | $\hat{x}_t = \rho\hat{x}_{t-1} + u_t^*$ |
| $\tau_{adj}$ | $\hat{x}_t = \rho\hat{x}_{t-1} + \Delta f_t(\hat{\theta})'\pi_1 + \Delta F_t(\hat{\theta})'\pi_2 + u_t^\dagger$ |
| $\tau_{adj}^+$ | $\hat{w}_t = \rho\hat{w}_{t-1} + [\hat{b}(L)\Delta f_t(\hat{\theta})']\pi_1 + [\hat{b}(L)\Delta F_t(\hat{\theta})']\pi_2 + \sum_{j=1}^p \alpha_j \Delta\hat{x}_{t-j} + r_t^\dagger$ |
| $\mathbf{t}_{LMS}$ | $\tilde{v}_t = \rho\tilde{v}_{t-1} + e_t^*$ |
| $\mathbf{t}_{adj}$ | $\tilde{v}_t = \rho\tilde{v}_{t-1} + \Delta f_t(\tilde{\theta})'\pi_1 + \Delta F_t(\tilde{\theta})'\pi_2 + \tilde{q}_t'\pi_3 + e_t^\dagger$ |
| | Asymptotic distribution $\left(\int_0^1 \bar{G}_c(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 \bar{G}_c(s;\bar{c}) dG_c(s;\bar{c})$ |
| $\tau_{int}$ | $\hat{x}_t = \nu + \rho\hat{x}_{t-1} + \Delta f_t(\hat{\theta})'\pi_1 + \Delta F_t(\hat{\theta})'\pi_2 + u_t^+$ |
| $\tau_{int}^+$ | $\hat{w}_t = \nu + \rho\hat{w}_{t-1} + [\hat{b}(L)\Delta f_t(\hat{\theta})']\pi_1 + [\hat{b}(L)\Delta F_t(\hat{\theta})']\pi_2 + \sum_{j=1}^p \alpha_j \Delta\hat{x}_{t-j} + r_t^+$ |
| $\mathbf{t}_{int}$ | $\tilde{v}_t = \nu + \rho\tilde{v}_{t-1} + \Delta f_t(\tilde{\theta})'\pi_1 + \Delta F_t(\tilde{\theta})'\pi_2 + \tilde{q}_t'\pi_3 + e_t^\dagger$ |
| | Asymptotic distribution $\left(\int_0^1 \bar{G}_c^*(s;\bar{c})^2 ds\right)^{-1/2} \int_0^1 \bar{G}_c^*(s;\bar{c}) dG_c^*(s;\bar{c})$ |
| $\tau_{S\&L}^*$ | $\hat{x}_t^* = \nu + \rho\hat{x}_{t-1}^* + u_t^{**}$ |
| $\tau_{int}^*$ | $\hat{x}_t^* = \nu + \rho\hat{x}_{t-1}^* + \Delta f_t(\hat{\theta}^*)'\pi_1 + \Delta F_t(\hat{\theta}^*)'\pi_2 + u_t^{+*}$ |
| $\tau_{int}^{*+}$ | $\hat{w}_t^* = \nu + \rho\hat{w}_{t-1}^* + [\hat{b}^*(L)\Delta f_t(\hat{\theta}^*)']\pi_1 + [\hat{b}^*(L)\Delta F_t(\hat{\theta}^*)']\pi_2 + \sum_{j=1}^p \alpha_j \Delta\hat{x}_{t-j}^* + r_t^*$ |
| $\mathbf{t}_{LMS}^*$ | $\tilde{v}_t^* = \nu + \rho\tilde{v}_{t-1}^* + e_t^{**}$ |
| $\mathbf{t}_{int}^*$ | $\tilde{v}_t^* = \nu + \rho\tilde{v}_{t-1}^* + \Delta f_t(\tilde{\theta}^*)'\pi_1 + \Delta F_t(\tilde{\theta}^*)'\pi_2 + \tilde{q}_t^{*\prime}\pi_3 + e_t^{*\dagger}$ |

Since we do not know which $\bar{c}$ value results in optimal local power of the tests with asymptotic distributions (4.11)/(4.22) and (4.12)/(4.23) we first investigate that issue. To this end we have generated critical values for a 5% significance level based on 10 000 drawings with sample size $T = 500$ using $c = 0$ and then we have simulated the local power curves in

Figures 1 and 2. In Figure 1 it is seen that the local power associated with the distribution in (4.11)/(4.22) is almost invariant to the value of $\bar{c}$. Hence, $\bar{c} = 0$ may just as well be used. In other words, the deterministic terms may be estimated under the null rather than local alternatives in order to get optimal local power for $\tau_{int}$, $\tau_{int}^{+}$ and $\mathbf{t}_{int}$. The same also holds for $\tau_{S\&L}^{*}$, $\tau_{int}^{*}$, $\tau_{int}^{*+}$, $\mathbf{t}_{LMS}^{*}$ and $\mathbf{t}_{int}^{*}$ as is seen in Figure 2. In that figure it is also obvious that for the latter statistics the value of $\bar{c}$ matters. However, optimal local power is achieved for $\bar{c} = 0$, at least for $c = 0, -3, \ldots, -30$.

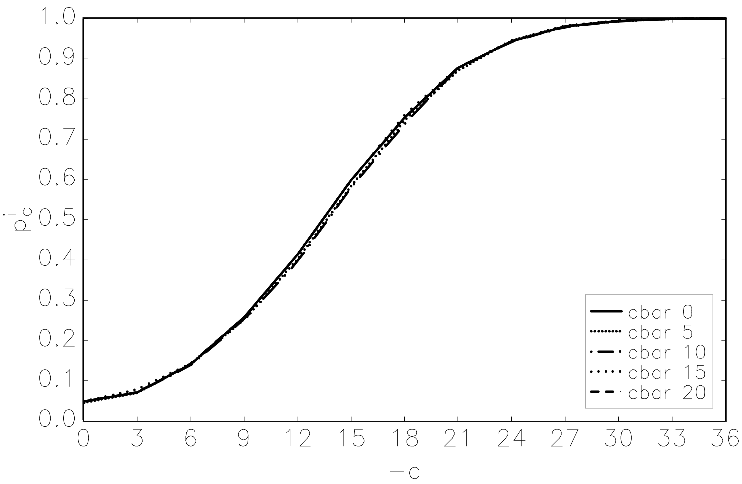


**Figure 1.** Local power associated with (4.11)/(4.22) ($\tau_{int}, \tau_{int}^{+}, \mathbf{t}_{int}$) ($-\bar{c} = 0, 5, 10, 15, 20$).

Some quantiles obtained from 10 000 drawings for different sample sizes and different values of $\bar{c}$ are given in Table 2. In the second and second last panel of the table quantiles are given for nonzero $\bar{c}$ values. They are seen to vary quite a bit with the sample size. In fact, they roughly decline in absolute value with growing $T$. For (4.8)/(4.19) the critical values correspond to the critical values of a DF $t$-test without any deterministic components in the DGP for large $T$ (see, e.g., Fuller (1976, Table 8.5.2)). For smaller sample sizes, however, they differ substantially from the asymptotic quantiles because in generating these null distributions we use an estimator for $\mu_0$ which is obtained under local alternatives. In this case we have used a transformation based on $\bar{\rho}_T = 1 + \bar{c}/T$ with $\bar{c} = -7$ because this value was recommended by Elliott et al. (1996) for processes without deterministic trend component ($\mu_1 = 0$). Elliott et al. show that this choice results in tests with optimal local power properties. Clearly, if the asymptotic critical values (see $T = 1000$ in the table) were

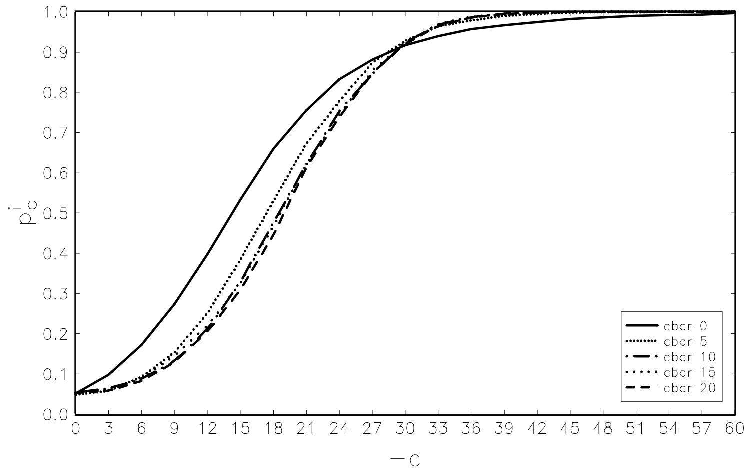


**Figure 2.** Local power associated with (4.12)/(4.23) ($\tau^*_{S\&L}, \tau^*_{int}, \tau^{*+}_{int}, \mathbf{t}^*_{LMS}, \mathbf{t}^*_{int}$) ($-\bar{c} = 0, 5, 10, 15, 20$).

used when the actual sample size is $T = 50$, say, the test would reject considerably more often than indicated by the significance level chosen. For example, the critical value for a 5% level test for $T = 1000$ is $-1.96$ which roughly corresponds to the 10% quantile of the distribution for $T = 50$. Thus, substantial small sample distortions of the size of the tests must be expected given that the present results are simulated under ideal conditions which are not likely to be satisfied in practice. Hence, in practice, additional sources for distortions may be present. The critical values for $\bar{c} = 0$ are less sensitive to the sample size which may be useful in applied work. In the third panel of the table, for all sample sizes, the quantiles are seen to be close to the corresponding quantiles of the DF distributions for DGPs with constant term (see again Table 8.5.2 of Fuller (1976)). Similarly, the simulated quantiles in the fifth panel ((4.10)/(4.21), $\bar{c} = -13.5$) are very close to those in Table I.C of Elliott et al. (1996) for all sample sizes given in that table.

We will now consider the local power properties resulting from the five distributions in Theorems 1 and 2 with $\bar{c} = -7$ for (4.8)/(4.19), $\bar{c} = -13.5$ for (4.10)/(4.21) and $\bar{c} = 0$ for the remaining distributions. As mentioned in Section 4 one would expect the tests based on the $\hat{x}_t^{(i)}$ series to have better power than those based on the $\hat{x}_t^{*(i)}$ because the former make assumptions regarding the initial values and, hence, in this respect they are based on tighter conditions than the latter tests. Viewed in a different way, the specific initial value assumptions are not important asymptotically whereas the tests based on the $\hat{x}_t^{*(i)}$ may suffer from the inclusion of an intercept term even asymptotically. Of course, tests based on the assumption $\mu_1 = 0$ are expected to have more power than the corresponding tests which do not use that a priori restriction. To explore these issues we have generated $x_t$ series according to the mechanism (5.1) with different values of $c$ and $T = 500$. Comparing the resulting test values to the 5% critical values in Table 2 gives the empirical local power of the tests. The corresponding local power curves are plotted in Figure 3. They are again based on 10 000 replications of the simulation experiment.

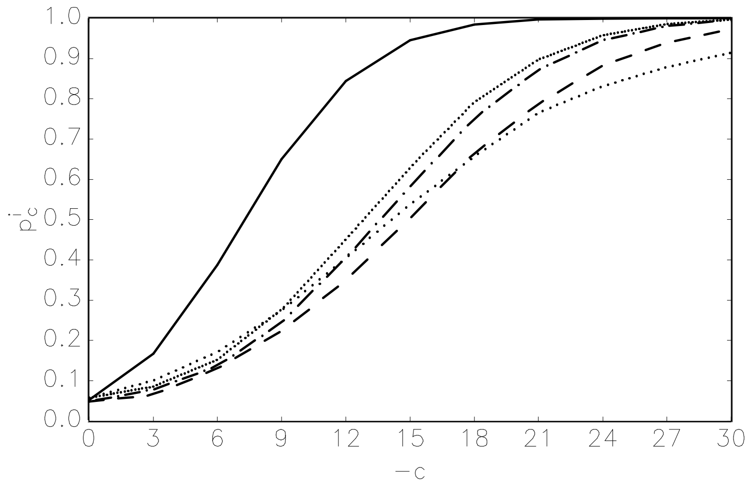


**Figure 3.** Local power of tests ($T = 500$).
(— (4.8)/(4.19) ($\tau^0_{S\&L}, \tau^0_{adj}, \tau^{+0}_{adj}, \mathbf{t}^0_{LMS}, \mathbf{t}^0_{adj}$),
· − · (4.9)/(4.20) ($\tau^0_{int}, \tau^{+0}_{int}, \tau^{*0}_{S\&L}, \tau^{*0}_{int}, \tau^{*+0}_{int}, \mathbf{t}^0_{int}, \mathbf{t}^{*0}_{LMS}, \mathbf{t}^{*0}_{int}$),
······ (4.10)/(4.21) ($\tau_{S\&L}, \tau_{adj}, \tau^+_{adj}, \mathbf{t}_{LMS}, \mathbf{t}_{adj}$), − − − (4.11)/(4.22) ($\tau_{int}, \tau^+_{int}, \mathbf{t}_{int}$),
· · · (4.12)/(4.23) ($\tau^*_{S\&L}, \tau^*_{int}, \tau^{*+}_{int}, \mathbf{t}^*_{LMS}, \mathbf{t}^*_{adj}$))

The results in the figure are as expected. The tests which use the restriction $\mu_1 = 0$ are relatively more powerful than the corresponding ones which do not take the restriction into account. Moreover, tests which include an intercept term in the auxiliary regression (based on $\hat{x}_t^{*(i)}$) tend to be less powerful than the corresponding tests based on the initial value assumption (based on the $\hat{x}_t^{(i)}$). Except for $\tau^0_{S\&L}$, $\tau^0_{adj}$, $\mathbf{t}^0_{LMS}$ and $\mathbf{t}^0_{adj}$, the differences in local power are in fact not very substantial. In other words, if a linear trend term cannot be excluded a priori, the price in terms of local power for not making the initial value assumption is not very high. On the other hand, substantial gains in local power are possible if $\mu_1 = 0$ can be assumed. In this case avoiding the initial value assumption has a quite high price. Generally tests which include an intercept term in the underlying regression model have reduced local power. Of course, local power is a concept based on asymptotic considerations. In small samples the situation may be quite different, in particular, if the initial value assumption is not satisfied for a time series of interest. Therefore we will explore the small sample properties of the different variants of the tests in the next section.

# 6 Small Sample Comparison

We have performed some simulations to investigate the performance of the tests in small samples based on the following two processes:

$$y_t = d_{1t} + x_t, \qquad (1 - b_1 L)(1 - \rho L)x_t = \varepsilon_t, \qquad t = 1, \dots, T, \tag{6.1}$$

and

$$(1 - b_1 L)y_t = d_{1t} + v_t, \qquad v_t = \rho v_{t-1} + \varepsilon_t, \qquad t = 1, \dots, T, \tag{6.2}$$

with $\varepsilon_t \sim iid\, N(0,1)$, $\rho = 1, 0.9, 0.8$, $T = 100, 200$. In some of the simulations we also generated 100 presample values which were discarded except that presample values were used in the estimations underlying model (2.2). Furthermore, we use $T_1/T = 0.5$, that is, the break point is half way through the sample. Preliminary simulations indicated that the location of the break point is not critical for the results as long as it is not very close to the beginning or the end of the sample. Therefore placing it in the middle does not imply a loss of generality for the situations we have in mind. The first process (6.1) is in line with the model (2.1) with an abrupt shift at time $T_1$ so that the $\tau$ tests are the appropriate tests whereas in general the model underlying the $\mathbf{t}$ tests can only approximate the DGP (6.1). Thus applying this test as well should give some indication of the flexibility of the framework and of the consequences of using the 'wrong' model. In contrast, the DGP (6.2) is a special case of (2.2) and generates a smooth shift in the deterministic term. For this process the $\mathbf{t}$ tests are appropriate whereas the $\tau$ tests are approximations only. To capture the smooth transition from one regime to another the $\tau$ tests may be combined with a smooth shift function. For both types of tests we use the shift functions given in Table 3 for both processes. The last two shift functions allow for smooth deterministic shifts. All three shift functions can be shown to satisfy Assumption 1. For some of the tests the derivatives of the shift functions are needed. They are also given in Table 3. Since $f_t^{(1)}$ does not depend on $\theta$ the derivative $F_t^{(1)}$ is zero. Hence, no extra terms $\Delta F_t^{(1)}(\theta)$ appear in the auxiliary regressions for $\tau_{adj}$, $\tau_{adj}^{+}$, $\mathbf{t}_{adj}$, $\tau_{int}$, $\tau_{int}^{+}$, $\mathbf{t}_{int}$, $\tau_{int}^{*}$, $\tau_{int}^{*+}$ and $\mathbf{t}_{int}^{*}$. In the simulations we use a range of $0 < \theta < 2$ for $f_t^{(2)}(\theta)$ and $0 < \theta < 0.8$ for $f_t^{(3)}(\theta)$ in estimating the parameters of the deterministic term. Although there is no linear trend term in the DGPs we allow for such a term in computing some of the test statistics.

Relative rejection frequencies from 1000 replications of the experiment are given in Tables 4 - 9. In Tables 4 and 5 actual sizes are given for tests for which estimation of the deterministic part is done under local alternatives ($\bar{c} = -7$ for the $\tau_{S\&L}^{0}$, $\tau_{adj}^{0}$, $\tau_{adj}^{+0}$, $\mathbf{t}_{LMS}^{0}$ and $\mathbf{t}_{adj}^{0}$ tests and $\bar{c} = -13.5$ for $\tau_{S\&L}$, $\tau_{adj}$, $\tau_{adj}^{+}$, $\mathbf{t}_{LMS}$ and $\mathbf{t}_{adj}$). The nominal significance level is 5% in all cases. Obviously, all tests reject too often in some situations. Note that asymptotic critical values are used so that some overrejection was to be expected on the basis of the discussion related to Table 2. For some cases unexpectedly large rejection frequencies are observed, however. For example, for $T = 100$ it is seen in Table 4 that $\mathbf{t}_{LMS}^{0}$ rejects in more than 50% of the cases for both DGPs if $b_1 = 0.8$ and the shift function $f_t^{(2)}$ is used in the test. Even if $T = 200$, the empirical size is unacceptable in this case, namely more than 30%. Some tests do reasonably well in specific situations. For example, $\tau_{adj}^{+0}$ and $\tau_{adj}^{+}$ produce rejection frequencies close to 5% when the correct shift function $f_t^{(1)}$ is used and the same is true for most of the tests when $T = 200$. Moreover, for most designs $\tau_{S\&L}$ rejects in less than 10% of the replications for $T = 200$ and is thereby best in this respect. Still, none of the tests performs satisfactorily for all shift functions and designs, in particular, for $T = 100$. Therefore the overall message from Tables 4 and 5 is clear: Using nonzero values of $\bar{c}$, that is, estimating under local alternatives, bears the risk of substantially distorted sizes of the tests. Thus, these tests cannot be recommended with the nonzero $\bar{c}$ values considered here. Consequently, there is no point in exploring their small sample power for these $\bar{c}$ values. Hence, in the following we focus on the tests with $\bar{c} = 0$, that is, estimation of the nuisance parameters is done under the null hypothesis.

Power results are given in Tables 6 - 9 for selected tests only. We will first comment on Tables 6 and 7 where the initial values used in the simulations are randomized by simulating 100 presample values as described previously. The results in Tables 6 and 7 show that for

$\bar{c} = 0$ the test sizes are much better in line with the nominal 5% (see $\rho = 1$) at least for those tests presented in the tables. In fact, for $\bar{c} = 0$ some tests tend to be conservative in specific situations and in some cases very much so (see, e.g., $\tau_{int}$ in combination with $f_t^{(1)}$). Most of the tests which are not shown in the tables tend to be generally conservative and therefore do not have much small sample power. In the tables we only show the results for those tests which performed overall best in terms of small sample power within their respective groups, the groups being $\tau^0$ tests ($\tau$ tests without linear trend term), $\tau$ tests (with linear trend), $\mathbf{t}^0$ tests and $\mathbf{t}$ tests. We are only presenting the best tests in the tables to avoid covering up the most important findings by the large volume of results for all the tests and simulation designs. It may be worth noting, however, that some of the other tests were nearly as good as the tests shown in the tables whereas some other tests performed very poorly indeed. Thus, some of the other tests are not very useful for applied work whereas some other ones are almost as good as those presented in the tables.

In the following, we consider only $\tau^0_{adj}$, $\tau^{+0}_{adj}$, $\mathbf{t}^0_{int}$, $\tau_{int}$, $\tau^+_{int}$ and $\mathbf{t}_{int}$. In the group of $\tau^0$ tests which exclude the deterministic trend term, $\tau^0_{adj}$ and $\tau^{+0}_{adj}$ were generally best in terms of power, each having advantages in some situations. Among the $\mathbf{t}^0$ tests, $\mathbf{t}^0_{int}$ was overall clearly best with highest power most of the time and close to the maximum in the other situations. Note also that its empirical size is usually relatively close to the nominal 5%. In all cases with $\rho = 1$ its relative rejection frequency is around 10% or less.

In the group of $\tau$ tests which allow for a linear trend term, $\tau_{int}$ and $\tau^+_{int}$ dominate the other tests. Again there is no clear winner among the two tests. Whereas $\tau_{int}$ is preferable in conjunction with shift function $f_t^{(3)}$, $\tau^+_{int}$ clearly dominates for $f_t^{(1)}$. Note, however, that both tests perform poorly for $b_1 = 0.8$ and $T = 100$. Finally, $\mathbf{t}_{int}$ is overall the best $\mathbf{t}$ test allowing for a trend. Its power is usually very close to that of $\mathbf{t}^*_{int}$, though. In fact, the two tests often produce identical rejection frequencies. Therefore, we present results for just one of them. For $T = 100$, both $\mathbf{t}_{int}$ and $\mathbf{t}^*_{int}$ reject a bit too often if they are used in conjunction with $f_t^{(2)}$ and $f_t^{(3)}$. This may not be too surprising given that using these shift functions for the presently considered DGP means that we are fitting a misspecified model. The tests are doing quite well if the correct shift function $f_t^{(1)}$ is used.

The following further conclusions emerge from Tables 6 and 7. In line with the local power results, excluding a linear trend term from the models when such a restriction is correct results in substantially better power. Although there are power differences between the best tests which allow for a linear trend, there is no clear winner. In other words, each of the tests is advantageous in some situations. On the one hand, $\mathbf{t}_{int}$ has often more power than $\tau_{int}$ and $\tau_{int}^{+}$ and, on the other hand, $\mathbf{t}_{int}$ tends to reject a bit too often. The same is true for tests excluding a linear trend term.

It is also apparent that it is not essential to use a test designed for a particular model when that model is in fact the true DGP. In other words, the performance of the tests is similar for the alternative DGPs (6.1) and (6.2). This may not be very surprising given that the two models are in some sense quite close. To see this multiply both sides of (2.1$a$) by $b(L)$ which yields

$$b(L)y_t = \nu_0 + \nu_1 t + f_t(\theta)'\lambda + b_*(L)\Delta f_t(\theta)'\gamma + v_t, \qquad t = p+1, \dots, T,$$

where $\nu_0$ and $\nu_1$ are functions of $\mu_0$, $\mu_1$ and the coefficients in $b(L)$, $\lambda = b(1)\gamma$ and $b_*(L)$ is obtained from the identity $b(L) = b(1) + b_*(L)\Delta$. Moreover, $v_t$ is as in (2.2$b$). This shows that if we condition on $y_1, \dots, y_p$ in model (2.1) we obtain a model of the form (2.2) except that the additional regressors $\Delta f_t(\theta), \dots, \Delta f_{t-p+1}(\theta)$ are included and nonlinear parameter restrictions are involved. By Assumption 1(b) the variables $\Delta f_t(\theta)$ are "asymptotically negligible."

The results in Tables 6 and 7 show that the performance of the tests depends more strongly on the shift functions than on the type of DGP. Furthermore, changing $b_1$ from 0.5 to 0.8 has a substantial effect. It implies a sizable decline in power in most cases. Again, this behaviour of the tests may not be too surprising because for $b_1$ close to 1 the processes have two roots close to unity and therefore are difficult to distinguish from unit root processes. Finally, the performance of all the tests improves markedly if $T$ is increased from 100 to 200.

It is noteworthy that the tests based on $\hat{x}_t^*$, $\hat{w}_t^*$ or $\hat{v}_t^*$, that is, the tests avoiding specific initial value assumptions in estimating the nuisance parameters, do not appear in the top group in Tables 6 and 7. This result is in line with the local power results. On the other hand, the initial value assumption which is used in deriving some of the tests is violated in the presently considered cases. Therefore we have explored the impact of the initial values by controlling them in some of our simulations. In Table 8, results for $T = 100$ and zero initial values are provided. Clearly, the power of the tests tends to be larger than in the corresponding entries in Table 6, especially for those tests which do not allow for a linear trend. Thus, using the initial values which are assumed in some theoretical derivations of the previous sections helps to improve power even in samples of size $T = 100$ although they have no impact asymptotically.

Since $T = 100$ is obviously too small to ensure the validity of asymptotic properties, it is also not surprising that the power tends to be smaller if unusually large initial values are considered. In Table 9 we show results for the situation where the initial values are all set to 5. For $\rho = 0.9$ and 0.8, the standard deviations of the $y_t$ generated by (6.1) range from about 3 to almost 10, depending on $b_1$. Hence, initial values of 5 may be regarded as moderate or large compared to the randomly chosen values in Tables 6 and 7. Using identical values for $y_{-1}$ and $y_0$ may be reasonable given the large correlation in the $y_t$. In Table 9 the power of the tests tends to be lower than in the corresponding Table 6. In some situations the power decline is particularly strong for tests that do not include an intercept term in the test regression ($\tau^0_{adj}, \tau^{+0}_{adj}$). A similar problem was also observed for some of the other tests based on regressions without an intercept and for which results are not shown in the tables. It may also be worth noting that the relative performance of the tests changed if zero initial values instead of random initial values are used. In that case some tests without intercept term in the test regression did a little better than in the nonzero initial value case. Thus, in particular if unusual initial values are suspected, using one of the tests with intercept term in the test regression is advisable. Alternatively one may remove the first values of a time series under consideration if they appear to be unusual.

The results in Tables 6, 8 and 9 also show that the tests are generally not very reliable if time series with $T = 100$ observations are under consideration. Moreover, the performance of the tests tends to be inferior if one of the misspecified and more complicated shift functions $f_t^{(2)}$ or $f_t^{(3)}$ is used.

# 7 Conclusions

Standard unit root tests are known to have reduced power if they are applied to time series with structural shifts. Therefore we have considered unit root tests that explicitly allow for a level shift of a very general possibly nonlinear form at a known point in time. We have argued that knowing the timing of the shift is quite common in practice whereas the precise form of the shift is usually unknown. Therefore, allowing for general and flexible shift functions is important. In this study we have focussed on models where the shift is regarded as part of the deterministic component of the DGP. Building on proposals by S&L and LMS it is suggested to estimate the deterministic part in a first step by a GLS procedure which may proceed under local alternatives or under the unit root null hypothesis. The original series is adjusted in a second step by subtracting the estimated deterministic part. Then DF type tests are applied to the adjusted series. A number of modifications of previously proposed tests of this sort are considered. In particular, tests are proposed that take into account estimation errors in the nuisance parameters and tests which do not assume specific initial values of the DGP. Local power and small sample properties of the tests are obtained.

The following general results emerge from our study. Some of the suggested modifications work clearly better in small samples than the original tests proposed by S&L and LMS in that they have superior size and power properties. Although local power gains are possible for some of the tests if the nuisance parameters are estimated under local alternatives rather than under the null hypothesis, substantial size distortions may result in small samples in the former case. Therefore we recommend estimating the nuisance parameters under the null hypothesis.

Initial values are found to have an impact on the small sample power of the tests. It turns out that including an intercept term in the test regression is important to guard against undesirable effects of large initial values. In practice, it may be worth discarding unusual values at the very beginning of a time series under consideration to avoid a loss in power due to untypical initial values.

If a deterministic linear time trend can be excluded on a priori grounds, it is recommended to perform tests in models without a linear trend term because excluding it may result in sizable power gains. Finally, using test versions with the best power properties is of particular importance in the present context because in some situations the tests do not perform very well for samples of size as large as $T = 100$.

Although we have focussed on a single shift in a time series, the tests can in principle be extended to allow for more than one shift. Of course, the small sample behaviour may be different in this case and needs to be explored in the future if applied researchers wish to use the tests in this more general context. In future research it may also be of interest to consider the situation where the timing of the shift is unknown and has to be determined from the data. Moreover, a comparison with other unit root tests which allow for structural shifts may be worthwhile. We leave these issues for future investigations.

# Appendix. Proofs

## A.1 Proof of Theorem 1

In the proof of Theorem 1 we focus on the limiting distributions of test statistics for models where $\mu_1$ is not known to be zero a priori. The case where the restriction $\mu_1 = 0$ is imposed follows by making straightforward modifications to these proofs. We begin with the result in (4.10).

The limiting distribution of $\tau_{S\&L}$ is derived in S&L. In that article it is given in a slightly different form, however. To see that the present form is equivalent it may be worth noting that $(A.21)$ of S&L may be written alternatively as

$$\begin{aligned} &T^{-1}\hat{X}_{-1}'\Sigma(\hat{b})^{-1}(\hat{X}-\hat{X}_{-1}) \\ &\quad = T^{-1}\textstyle\sum_{t=p}^{T}[\hat{b}(L)\hat{x}_{t-1}][\hat{b}(L)\Delta\hat{x}_t] + o_p(1) \\ &\quad = T^{-1}\textstyle\sum_{t=p}^{T}[b(1)\{x_{t-1}-(\hat{\mu}_1-\mu_1)(t-1)\}][b(L)\Delta x_t - b(1)(\hat{\mu}_1-\mu_1)] + o_p(1) \\ &\quad \xrightarrow{d} \sigma^2\textstyle\int_0^1 G_c(s;\bar{c})dB_c(s) - \sigma^2 K_c(\bar{c})\int_0^1 G_c(s;\bar{c})ds, \end{aligned} \qquad (A.1)$$

where the last relation follows from well-known limit theorems by noting that the limiting distribution of $\hat{\mu}$ given in (3.12) of S&L can be written alternatively as $\omega K_c(\bar{c})$, where $\omega = \sigma/b(1)$,

$$K_c(\bar{c}) = h(\bar{c})^{-1}\int_0^1 (1-\bar{c}s)dB_0(s) + h(\bar{c})^{-1}(c-\bar{c})\int_0^1 (1-\bar{c}s)B_c(s)ds \qquad (A.2)$$

and $h(\bar{c}) = 1 - \bar{c} + \bar{c}^2/3$. From the representation in $(A.1)$ the limiting distribution in (4.10) follows as in the proof of the asymptotic distribution of the test statistic in S&L. Thus, to prove (4.10), it remains to show that $\tau_{adj}$ and $\tau_{adj}^{+}$ have the same limiting distribution as $\tau_{S\&L}$.

Using

$$T^{-1/2}\hat{x}_{[Ts]} \xrightarrow{d} \omega G_c(s;\bar{c}) \qquad (A.3)$$

(see $(A.18)$ of S&L) and the fact that $f_t(\theta)$ satisfies Assumption 1(b) it can be seen that

$$\left\| T^{-1}\sum_{t=1}^{T}\hat{x}_{t-1}\Delta f_t(\hat{\theta})\right\| \leq T^{-1}\max_{1\leq t\leq T}|\hat{x}_t|\sum_{t=1}^{T}\sup_{\theta\in\Theta}\|\Delta f_t(\theta)\| = O_p\left(T^{-1/2}\right)$$

and that a similar result also holds with $\Delta f_t(\hat{\theta})$ replaced by $\Delta F_t(\hat{\theta})$. Using these facts and arguments similar to those in the proof of Lemma 1 of S&L it can be shown that the appropriately standardized moment matrix in the GLS estimation of (4.3) is asymptotically block diagonal and also positive definite. Since it is further straightforward to show that

$$\sum_{t=1}^{T} \Delta f_t(\hat{\theta}) u_t^{\dagger} = O_p(1)$$

and similarly with $\Delta f_t(\hat{\theta})$ replaced by $\Delta F_t(\hat{\theta})$ it follows that the limiting distribution of the GLS estimator of $\rho$ in (4.3) and hence that of its $t$-ratio is the same as in the case of the auxiliary regression model (4.1). We have thus shown that (4.10) holds for the test statistic $\tau_{adj}$.

As for test statistic $\tau_{adj}^{+}$, note first that the arguments used for $\tau_{adj}$ above and those in the proof of Theorem 1 of S&L show that the appropriately standardized moment matrix in the auxiliary regression model used to obtain the test statistic $\tau_{adj}^{+}$ is asymptotically positive definite and also block diagonal between $\hat{w}_{t-1}$ and the other regressors. Using the expression of the error term in this auxiliary regression model it is further straightforward to show that $\tau_{adj}^{+}$ has the same limiting distribution as $\tau_{S\&L}$ and $\tau_{adj}$. Thus, (4.10) is proven.

Since the test statistics $\tau_{int}$ and $\tau_{int}^{+}$ are obtained by augmenting the auxiliary regression models used to obtain test statistics $\tau_{adj}$ and $\tau_{adj}^{+}$, respectively, by an intercept term, (4.11) can be proven by extending the arguments used above in a standard manner.

Before we prove (4.12) we establish a useful intermediate result regarding the properties of some of the estimators described in Sec. 3. The following lemma complements results presented in Lemma 1 of S&L.

**Lemma A.1.**

Suppose that Assumption 1 holds. Suppose further that the matrix $Z^*(\theta)$ is of full column rank for all $T \geq k+1$ and all $\theta \in \Theta$. Then,

$$\hat{\theta}^* = \theta + O_p(1), \qquad (A.4)$$

$$\hat{\gamma}^* = \gamma + O_p(1), \qquad (A.5)$$

$$\hat{b}^* \xrightarrow{p} b \qquad (A.6)$$

and

$$T^{1/2}(\hat{\mu}_1^* - \mu_1) \xrightarrow{d} \begin{cases} \omega K_c^*(\bar{c}), & \text{if } \bar{c} < 0 \\ \omega B_c(1), & \text{if } \bar{c} = 0 \end{cases} \tag{A.7}$$

where $\omega = \sigma/b(1)$ and

$$K_c^*(\bar{c}) = \frac{12}{\bar{c}} \int_0^1 \left(\frac{1}{2} - s\right) dB_0(s) + \frac{12(c - \bar{c})}{\bar{c}} \int_0^1 \left(\frac{1}{2} - s\right) B_c(s) ds.$$

□

**Remark:** For simplicity we have assumed a full rank condition similar to that in Lemma 1 of S&L. Lemma A.1 shows that, except for $\mu_0$, GLS estimators of the other parameters have similar properties as in Lemma 1 of S&L. The limiting distribution of the GLS estimator of $\mu_1$ depends on the chosen value of $\bar{c}$ in a discontinuous way, though. If $\bar{c} = 0$ the limiting distribution is the same as in Lemma 1 of S&L but if $\bar{c} < 0$ a different result is obtained (see $(A.2)$). Next denote by $Z_{12}$ the second column of $Z_1$ and note the following two facts: (a) $T^{-1} Z_{12}' Z_{12} \to h(\bar{c})$ while the corresponding limit obtained from the mean-adjusted version of $Z_{12}$ is $\bar{c}^2/12$ and (b) $\int_0^1 (1 - \bar{c}s) ds = 1 - \bar{c}/2$ so that the mean-adjusted version of $1 - \bar{c}s$ is $\bar{c}(\frac{1}{2} - s)$. Thus, from $(A.2)$ and the definition of $K_c^*(\bar{c})$ it can be seen that in the present situation the inclusion of a constant term in the regression model affects the result so that the first limiting distribution in $(A.7)$ may be viewed as a 'mean adjusted' version of that of the limiting distribution of $\hat{\mu}_1$ (denoted $\hat{\mu}$) in Lemma 1 of S&L.

**Proof of Lemma A.1.** If $\bar{c} = 0$ the result is fairly obvious because the considered regression model differs from that used in Lemma 1 of S&L only in that the first observation is omitted.

Now suppose that $\bar{c} < 0$. Arguments similar to those used for $(A.1)$ - $(A.3)$ in S&L then readily show that $T^{-1} Z_1^{*\prime} Z_1^*$ converges to a positive definite limit and that $T^{-1/2} Z_1^{*\prime} Z_2^*(\theta) = O(T^{-1/2})$ uniformly in $\theta$. When these results are available it is straightforward to proceed in the same way as in the proof of $(A.8)$ of S&L and establish $(A.5)$ and also that the GLS estimators of $\mu_1$ and $\mu_0^*$ are consistent of order $O_p(T^{-1/2})$. The next step is to prove the consistency of $\hat{b}^*$ but, making use of the above mentioned results, this can be done in the same way as the corresponding step in the proof of Lemma 1 of S&L. In the same way as in that proof one can also show that instead of the representation of $T^{1/2}(\hat{\mu}_1 - \mu_1)$ given there we now have

$$\begin{bmatrix} T^{1/2}(\hat\mu_0^* - \mu_0^*) \\ T^{1/2}(\hat\mu_1^* - \mu_1) \end{bmatrix} = (T^{-1}Z_1^{*\prime}\Sigma^*(\hat b^*)^{-1}Z_1^*)^{-1}T^{-1/2}Z_1^{*\prime}\Sigma^*(\hat b^*)^{-1}U^* + o_p(1)$$

and, furthermore,

$$T^{1/2}(\hat\mu_1^* - \mu_1) = (T^{-1}\bar Z_{12}'\bar Z_{12})^{-1}T^{-1/2}\bar Z_{12}'U + o_p(1),$$

where $\bar Z_{12}$ is a mean-adjusted version of $Z_{12}$ with a typical component $\bar c(\frac{1}{2} - \frac{t}{T}) + \frac{\bar c}{2T}$. By simple calculation, $T^{-1}\bar Z_{12}'\bar Z_{12} = \frac{\bar c^2}{12} + o(1)$. The first result in $(A.7)$ follows from this, (3.3) of S&L and well-known limit theorems. This completes the proof of Lemma A.1. □

Now we can turn to the proof of the result in (4.12). We derive the limit distribution of $\tau_{S\&L}^*$ and just note that the result for the other statistics follows with similar arguments as those used to prove (4.10) and (4.11). Define

$$\hat X^* = [\hat x_2^* : \cdots : \hat x_T^*]'$$

and

$$\hat Q^* = \begin{bmatrix} 1 & \cdots & 1 \\ \hat x_1^* & \cdots & \hat x_{T-1}^* \end{bmatrix}.$$

Then the GLS estimators obtained from (4.6) satisfy

$$\begin{bmatrix} \hat\nu^* - \nu \\ \hat\rho^* - 1 \end{bmatrix} = \left(\hat Q^{*\prime}\Sigma^*(\hat b^*)^{-1}\hat Q^*\right)^{-1}\hat Q^{*\prime}\Sigma^*(\hat b^*)^{-1}\Delta\hat X^*. \qquad (A.8)$$

The proof of the theorem essentially means deriving the limiting distribution of $T(\hat\rho^* - 1)$ and showing that $\hat\nu^* - \nu = O_p(T^{-1/2})$. Details are similar to those in the proof of Theorem 1 of S&L and in the above proof given for (4.10). First note that

$$\hat x_t^* = x_t^* - (\hat\mu_1^* - \mu_1)t - f_t(\hat\theta^*)'\hat\gamma^* + f_t(\theta)'\gamma. \qquad (A.9)$$

In the same way as in the case of $(A.17)$ of S&L we can conclude from this, $T^{-1/2}x_{[sT]} \xrightarrow{d} \omega B_c(s)$ (see (3.3) of S&L) and Lemma A.1 that

$$T^{-\frac{1}{2}}\hat x_{[Ts]}^* \xrightarrow{d} \omega G_c^*(s;\bar c). \qquad (A.10)$$

It is also straightforward to see that $(A.19)$ of S&L holds with $\Delta\hat x_t$ replaced by $\Delta\hat x_t^*$ and that $\hat Q^*$ and $\Delta\hat X^*$ in $(A.8)$ can be replaced by analogs defined in terms of $x_t^* - (\hat\mu_1^* - \mu_1)t$

and $\Delta x_t^* - (\hat{\mu}_1^* - \mu_1) = \Delta x_t - (\hat{\mu}_1^* - \mu_1)$. As far as asymptotic distributions are concerned, the effect of the quantities $f_t(\hat{\theta}^*)'\hat{\gamma}^*$ and $f_t(\theta)'\gamma$ on $x_t^*$ in $(A.9)$ can thus be ignored. Using these facts, $(A.10)$, well-known limit theorems and arguments similar to those in $(A.20)$ of S&L and $(A.1)$ above, one can derive the limiting distribution of the test statistic $\tau_{S\&L}^*$ in a straightforward fashion. Details are omitted.

## A.2 Proof of Theorem 2

We focus again on the case where $\mu_1$ is not zero a priori. The result for $\mathbf{t}_{LMS}$ in (4.21) can be obtained from LMS so that we consider $\mathbf{t}_{adj}$.

We shall first study the appropriately standardized moment matrix in the OLS estimation of the parameters in (4.15). By Lemma A.1 of LMS, we have $\tilde{\mu}_* = \mu_* + O_p(T^{-1/2})$ which, in conjunction with $(A.10)$ and arguments similar to those in $(A.14)$ of LMS, can be used to show that

$$T^{-1}\sum_{t=1}^{T}\tilde{q}_t\tilde{q}_t' \xrightarrow{p} \sigma^2\Sigma(b).$$

The same arguments and the representation given for $\tilde{v}_t$ in the proof of Theorem 1 of LMS yield

$$T^{-3/2}\sum_{t=1}^{T}\tilde{v}_{t-1}\tilde{q}_t' = o_p(1).$$

Finally, when the assumptions made for $f_t(\theta)$ and $F_t(\theta)$ are also used, we get

$$T^{-1}\sum_{t=1}^{T}\tilde{v}_{t-1}\left[\Delta f_t(\tilde{\theta})' : \Delta F_t(\tilde{\theta})'\right] = o_p(1)$$

and

$$T^{-1/2}\sum_{t=1}^{T}\tilde{q}_t\left[\Delta f_t(\tilde{\theta})' : \Delta F_t(\tilde{\theta})'\right] = o_p(1).$$

Thus, the appropriately standardized moment matrix between the three regressors $\tilde{v}_{t-1}$, $\left[\Delta f_t(\tilde{\theta})' : \Delta F_t(\tilde{\theta})'\right]'$ and $\tilde{q}_t$ is asymptotically block diagonal. It is also asymptotically positive definite, as can be seen by using the assumptions and arguments similar to those in the proof of Lemma A.1 of LMS.

We shall next consider the error term $e_t^\dagger$ in (4.15) and show how it is related to the error term $e_t^*$ in (4.13). First, recall that $\rho_T(L) = 1 - \rho_T L$ and observe that, for $t \geq 2$,

$$\begin{aligned} e_t^* &= \varepsilon_t + \left(\tilde{b}(L) - b(L)\right)\left(\Delta y_t - \tfrac{c}{T}y_{t-1}\right) + \tfrac{c}{T}\left(\tilde{\mu}_0 - \mu_0\right) \\ &\quad - \left(\tilde{\mu}_1 - \mu_1\right)\left(1 - \tfrac{c(t-1)}{T}\right) - \left(\rho_T(L)f_t(\tilde{\theta})'\tilde{\gamma} - \rho_T(L)f_t(\theta)'\gamma\right). \end{aligned} \tag{A.11}$$

Here we have used the definitions of $v_t$ and $\tilde{v}_t$ and the identities $\rho_T(L)y_t = \Delta y_t - \frac{c}{T}y_{t-1}$ and $\rho_T(L)t = 1 - \frac{c(t-1)}{T}$. Identifying the parameter vector $\pi_3$ in (4.15) with $-(\tilde{b} - b)$ shows that the inclusion of the regressors $\Delta y_{t-j} - \tilde{\mu}_*$ $(j = 1, \ldots, p)$ in (4.13) changes the second term on the r.h.s. of $(A.11)$ to $(\tilde{b}(L) - b(L))(\tilde{\mu}_* - \frac{c}{T}y_{t-1})$. It is also easy to see that, as far as the limiting distribution of the test statistics $\mathbf{t}_{LMS}$ and $\mathbf{t}_{adj}$ is concerned, the contribution of the third and fifth terms on the r.h.s. of $(A.11)$ is negligible. In the same way one can conclude that, from our point of view, terms which are added to the error term $e_t^\dagger$ by including the regressors $\Delta f_t(\tilde{\theta})$ and $\Delta F_t(\tilde{\theta})$ are asymptotically negligible. Thus, we can conclude that for our purposes the error term $e_t^\dagger$ can be treated by using the approximation

$$e_t^\dagger \approx \varepsilon_t - (\tilde{b}(L) - b(L)) \left(\tilde{\mu}_* - \frac{c}{T}y_{t-1}\right) - (\tilde{\mu}_1 - \mu_1) \left(1 - \frac{c(t-1)}{T}\right). \tag{A.12}$$

Using Lemma A.1 and equation $(A.10)$ of LMS in conjunction with the representation given for $\tilde{v}_t$ in the proof of Theorem 1 of the same paper it can be shown that $e_t^\dagger$ can be further approximated by replacing $\tilde{\mu}_*$ on the r.h.s. of $(A.12)$ first by $\mu_*$ and then by $\Delta y_t$. Since the third and fifth terms on the r.h.s. of $(A.11)$ can be ignored we have thus demonstrated that the above approximation of $e_t^\dagger$ becomes $e_t^\dagger \approx e_t^*$ and, since the appropriately standardized moment matrix in the OLS estimation of (4.15) is asymptotically block diagonal, it follows that the OLS estimators obtained for the parameter $\rho$ from (4.13) and (4.15) are asymptotically equivalent. Since it is straightforward to show that the same is true for the related error variance estimators the limiting distribution of $\mathbf{t}_{adj}$ follows.

Using the definition of the test statistic $\mathbf{t}_{int}$ and the above arguments it is straightforward to prove (4.22). In order to prove (4.23) we first need the following analog of Lemma A.1.

**Lemma A.2.**

Suppose that the assumptions of Theorem 2 hold. Then,

$$\tilde{b}^* \xrightarrow{p} b \tag{A.13}$$

$$\tilde{\theta}^* = \theta + O_p(1) \tag{A.14}$$

$$\tilde{\gamma}^* = \gamma + O_p(1) \tag{A.15}$$

and

$$T^{1/2}(\tilde{\mu}_1^* - \tilde{b}^*(1)\mu_1/b(1)) \xrightarrow{d} \sigma K_c^*(\bar{c}), \tag{A.16}$$

where $K_c^*(\bar{c})$ is as in Lemma A.1 with $K_c^*(0) = B_c(1)$.

**Proof:** The proof can be obtained by using the arguments in the proof of Lemma A.1 of LMS and those in the proof of Lemma A.1. Details are straightforward and therefore omitted. □

Once the result of Lemma A.2 is available, the limiting distribution of the test statistic $\mathbf{t}_{LMS}^*$ can be obtained by following the arguments in the proof of Theorem 1 of LMS and in deriving the limiting distribution of the test statistic $\tau_{S\&L}^*$. Similar arguments combined with those used to prove the asymptotic distribution of the test statistic $\mathbf{t}_{adj}$ show that $\tau_{int}^*$ has the same limiting distribution as $\mathbf{t}_{LMS}^*$.

# References


Amsler, C. & J. Lee (1995), An LM test for a unit root in the presence of a structural change, *Econometric Theory*, 11, 359 - 368.

Banerjee, A., R.L. Lumsdaine & J.H. Stock (1992), Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business & Economic Statistics*, 10, 271 - 287.

Elliott, G., T.J. Rothenberg & J.H. Stock (1996), Efficient tests for an autoregressive unit root, *Econometrica*, 64, 813 - 836.

Fuller, W.A. (1976), *Introduction to Statistical Time Series*, New York: Wiley.

Leybourne, S., P. Newbold & D. Vougas (1998), Unit roots and smooth transitions, *Journal of Time Series Analysis*, 19, 83 - 97.

Lütkepohl, H., C. Müller & P. Saikkonen (1999), Unit root tests for time series with a structural break when the break point is known, Discussion Paper, SFB 373, Humboldt-Universität zu Berlin.

Montañés, A. & M. Reyes (1998), Effect of a shift in the trend function on Dickey-Fuller unit root tests, *Econometric Theory*, 14, 355 - 363.

Perron, P. (1989), The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361 - 1401.

Perron, P. (1990), Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics*, 8, 153 - 162.

Perron, P. & T.J. Vogelsang (1992), Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business & Economic Statistics*, 10, 301 - 320.

Saikkonen, P. & H. Lütkepohl (1999), Testing for unit roots in time series with level shifts, Discussion Paper, SFB 373, Humboldt-Universität zu Berlin.

Zivot, E. & D.W.K. Andrews (1992), Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis, *Journal of Business & Economic Statistics*, 10, 251 - 270.

**Table 2.** Simulated Quantiles of Null Distributions of Test Statistics Based on 10 000 Replications

| Distribution | $T$ | $\alpha_{0.01}$ | $\alpha_{0.025}$ | $\alpha_{0.05}$ | $\alpha_{0.1}$ |
|---|---|---|---|---|---|
| (4.8)/(4.19) ($\bar{c} = 0$) | 50 | -2.65 | -2.26 | -1.97 | -1.63 |
| | 100 | -2.61 | -2.25 | -1.96 | -1.62 |
| | 200 | -2.64 | -2.26 | -1.94 | -1.62 |
| | 500 | -2.60 | -2.25 | -1.95 | -1.62 |
| | 1000 | -2.55 | -2.24 | -1.96 | -1.61 |
| (4.8)/(4.19) ($\bar{c} = -7$) | 50 | -2.93 | -2.56 | -2.28 | -1.98 |
| | 100 | -2.73 | -2.41 | -2.15 | -1.83 |
| | 200 | -2.68 | -2.34 | -2.05 | -1.73 |
| | 500 | -2.64 | -2.30 | -2.00 | -1.67 |
| | 1000 | -2.56 | -2.22 | -1.96 | -1.63 |
| (4.9)/(4.20) ($\bar{c} = 0$) | 50 | -3.64 | -3.28 | -2.99 | -2.67 |
| | 100 | -3.58 | -3.22 | -2.94 | -2.62 |
| | 200 | -3.58 | -3.22 | -2.93 | -2.62 |
| | 500 | -3.47 | -3.17 | -2.90 | -2.62 |
| | 1000 | -3.48 | -3.15 | -2.88 | -2.58 |
| (4.10)/(4.21) ($\bar{c} = 0$) | 50 | -3.34 | -2.96 | -2.65 | -2.37 |
| | 100 | -3.23 | -2.90 | -2.61 | -2.33 |
| | 200 | -3.17 | -2.91 | -2.64 | -2.33 |
| | 500 | -3.22 | -2.92 | -2.64 | -2.35 |
| | 1000 | -3.18 | -2.86 | -2.62 | -2.33 |
| (4.10)/(4.21) ($\bar{c} = -13.5$) | 50 | -3.83 | -3.48 | -3.21 | -2.91 |
| | 100 | -3.62 | -3.30 | -3.03 | -2.74 |
| | 200 | -3.51 | -3.24 | -2.96 | -2.66 |
| | 500 | -3.43 | -3.09 | -2.84 | -2.57 |
| | 1000 | -3.40 | -3.11 | -2.85 | -2.57 |
| (4.11)/(4.22) (4.12)/(4.23) ($\bar{c} = 0$) | 50 | -3.81 | -3.45 | -3.15 | -2.86 |
| | 100 | -3.73 | -3.38 | -3.11 | -2.80 |
| | 200 | -3.64 | -3.32 | -3.06 | -2.77 |
| | 500 | -3.62 | -3.32 | -3.08 | -2.79 |
| | 1000 | -3.55 | -3.28 | -3.03 | -2.76 |

**Table 3.** Shift Functions and Their Derivatives

| Shift function $f_t(\theta)$ | Derivatives $F_t(\theta)$ |
|---|---|
| $f_t^{(1)}(\theta) = d_{1t} = \begin{cases} 0, & t < T_1 \\ 1, & t \geq T_1 \end{cases}$ | $F_t^{(1)}(\theta) = 0$ |
| $f_t^{(2)}(\theta) = \begin{cases} 0, & t < T_1 \\ 1 - \exp\{-\theta(t - T_1)\}, & t \geq T_1 \end{cases}$ | $F_t^{(2)}(\theta) = \begin{cases} 0, & t < T_1 \\ (t - T_1)\exp\{-\theta(t - T_1)\}, & t \geq T_1 \end{cases}$ |
| $f_t^{(3)}(\theta) = \begin{bmatrix} \dfrac{d_{1,t}}{1 - \theta L} \\ \dfrac{d_{1,t-1}}{1 - \theta L} \end{bmatrix}$ | $F_t^{(3)}(\theta) = \begin{bmatrix} \dfrac{d_{1,t-1}}{(1 - \theta L)^2} \\ \dfrac{d_{1,t-2}}{(1 - \theta L)^2} \end{bmatrix}$ |

**Table 4.** Empirical Sizes of Tests, $T = 100$, $T_1 = 50$, $\bar{c} = -7/-13.5$, Nominal Significance Level 5%, Random Initial Values

| Shift function | DGP | $b_1$ | $\tau^0_{S\&L}$ | $\tau^0_{adj}$ | $\tau^{+0}_{adj}$ | $\tau_{S\&L}$ | $\tau_{adj}$ | $\tau^+_{adj}$ | $\mathbf{t}^0_{LMS}$ | $\mathbf{t}^0_{adj}$ | $\mathbf{t}_{LMS}$ | $\mathbf{t}_{adj}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_t^{(1)}$ | (6.1) | 0.5 | 0.077 | 0.076 | 0.069 | 0.085 | 0.087 | 0.071 | 0.244 | 0.104 | 0.243 | 0.123 |
| | | 0.8 | 0.164 | 0.165 | 0.064 | 0.072 | 0.073 | 0.063 | 0.460 | 0.143 | 0.261 | 0.145 |
| | (6.2) | 0.5 | 0.077 | 0.077 | 0.070 | 0.081 | 0.083 | 0.068 | 0.241 | 0.106 | 0.260 | 0.128 |
| | | 0.8 | 0.168 | 0.168 | 0.064 | 0.070 | 0.071 | 0.062 | 0.472 | 0.158 | 0.272 | 0.160 |
| $f_t^{(2)}$ | (6.1) | 0.5 | 0.196 | 0.207 | 0.307 | 0.191 | 0.240 | 0.343 | 0.377 | 0.225 | 0.358 | 0.249 |
| | | 0.8 | 0.258 | 0.247 | 0.367 | 0.140 | 0.144 | 0.226 | 0.566 | 0.276 | 0.396 | 0.271 |
| | (6.2) | 0.5 | 0.201 | 0.217 | 0.317 | 0.184 | 0.249 | 0.354 | 0.373 | 0.229 | 0.358 | 0.248 |
| | | 0.8 | 0.249 | 0.276 | 0.409 | 0.122 | 0.170 | 0.270 | 0.568 | 0.279 | 0.390 | 0.264 |
| $f_t^{(3)}$ | (6.1) | 0.5 | 0.193 | 0.269 | 0.224 | 0.158 | 0.360 | 0.262 | 0.387 | 0.225 | 0.327 | 0.217 |
| | | 0.8 | 0.206 | 0.533 | 0.227 | 0.080 | 0.501 | 0.160 | 0.564 | 0.278 | 0.357 | 0.269 |
| | (6.2) | 0.5 | 0.197 | 0.272 | 0.231 | 0.153 | 0.351 | 0.258 | 0.397 | 0.233 | 0.337 | 0.212 |
| | | 0.8 | 0.215 | 0.526 | 0.235 | 0.080 | 0.510 | 0.158 | 0.586 | 0.302 | 0.376 | 0.280 |

**Table 5.** Empirical Sizes of Tests, $T = 200$, $T_1 = 100$, $\bar{c} = -7/-13.5$, Nominal Significance Level 5%, Random Initial Values

| Shift function | DGP | $b_1$ | $\tau^0_{S\&L}$ | $\tau^0_{adj}$ | $\tau^{+0}_{adj}$ | $\tau_{S\&L}$ | $\tau_{adj}$ | $\tau^+_{adj}$ | $\mathbf{t}^0_{LMS}$ | $\mathbf{t}^0_{adj}$ | $\mathbf{t}_{LMS}$ | $\mathbf{t}_{adj}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f^{(1)}_t$ | (6.1) | 0.5 | 0.059 | 0.059 | 0.054 | 0.060 | 0.061 | 0.053 | 0.082 | 0.070 | 0.082 | 0.064 |
| | | 0.8 | 0.057 | 0.057 | 0.042 | 0.057 | 0.057 | 0.051 | 0.214 | 0.069 | 0.152 | 0.069 |
| | (6.2) | 0.5 | 0.057 | 0.057 | 0.052 | 0.059 | 0.061 | 0.053 | 0.084 | 0.071 | 0.084 | 0.064 |
| | | 0.8 | 0.052 | 0.052 | 0.041 | 0.057 | 0.058 | 0.050 | 0.217 | 0.069 | 0.167 | 0.073 |
| $f^{(2)}_t$ | (6.1) | 0.5 | 0.134 | 0.141 | 0.154 | 0.107 | 0.129 | 0.145 | 0.176 | 0.144 | 0.122 | 0.127 |
| | | 0.8 | 0.083 | 0.089 | 0.127 | 0.088 | 0.100 | 0.129 | 0.348 | 0.147 | 0.217 | 0.124 |
| | (6.2) | 0.5 | 0.133 | 0.141 | 0.158 | 0.111 | 0.136 | 0.155 | 0.172 | 0.141 | 0.122 | 0.128 |
| | | 0.8 | 0.086 | 0.098 | 0.144 | 0.087 | 0.123 | 0.167 | 0.339 | 0.140 | 0.227 | 0.129 |
| $f^{(3)}_t$ | (6.1) | 0.5 | 0.131 | 0.156 | 0.145 | 0.093 | 0.135 | 0.123 | 0.184 | 0.148 | 0.127 | 0.121 |
| | | 0.8 | 0.090 | 0.226 | 0.101 | 0.059 | 0.312 | 0.090 | 0.355 | 0.143 | 0.203 | 0.111 |
| | (6.2) | 0.5 | 0.129 | 0.159 | 0.149 | 0.095 | 0.133 | 0.120 | 0.181 | 0.151 | 0.130 | 0.125 |
| | | 0.8 | 0.088 | 0.234 | 0.107 | 0.063 | 0.322 | 0.097 | 0.367 | 0.153 | 0.210 | 0.119 |

**Table 6.** Relative Rejection Frequencies of Tests, $T = 100$, $T_1 = 50$, $\bar{c} = 0$, Nominal Significance Level 5%, Random Initial Values

| Shift function | Test | DGP (6.1), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.1), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f^{(1)}_t$ | $\tau^0_{adj}$ | 0.039 | 0.291 | 0.535 | 0.016 | 0.156 | 0.315 | 0.040 | 0.285 | 0.527 | 0.020 | 0.140 | 0.275 |
| | $\tau^{+0}_{adj}$ | 0.063 | 0.353 | 0.590 | 0.050 | 0.292 | 0.436 | 0.061 | 0.343 | 0.575 | 0.053 | 0.287 | 0.382 |
| | $\mathbf{t}^0_{int}$ | 0.054 | 0.292 | 0.561 | 0.067 | 0.234 | 0.354 | 0.060 | 0.289 | 0.577 | 0.065 | 0.227 | 0.345 |
| | $\tau_{int}$ | 0.020 | 0.090 | 0.302 | 0.000 | 0.006 | 0.034 | 0.022 | 0.091 | 0.305 | 0.001 | 0.004 | 0.029 |
| | $\tau^+_{int}$ | 0.080 | 0.233 | 0.526 | 0.065 | 0.167 | 0.286 | 0.075 | 0.216 | 0.499 | 0.064 | 0.149 | 0.262 |
| | $\mathbf{t}_{int}$ | 0.081 | 0.217 | 0.455 | 0.077 | 0.159 | 0.268 | 0.079 | 0.216 | 0.468 | 0.079 | 0.161 | 0.269 |
| $f^{(2)}_t$ | $\tau^0_{adj}$ | 0.063 | 0.247 | 0.486 | 0.036 | 0.145 | 0.263 | 0.064 | 0.259 | 0.486 | 0.046 | 0.158 | 0.281 |
| | $\tau^{+0}_{adj}$ | 0.069 | 0.253 | 0.496 | 0.042 | 0.166 | 0.281 | 0.072 | 0.266 | 0.491 | 0.049 | 0.176 | 0.299 |
| | $\mathbf{t}^0_{int}$ | 0.100 | 0.286 | 0.547 | 0.099 | 0.238 | 0.361 | 0.095 | 0.306 | 0.535 | 0.088 | 0.227 | 0.366 |
| | $\tau_{int}$ | 0.055 | 0.142 | 0.348 | 0.018 | 0.034 | 0.059 | 0.051 | 0.157 | 0.358 | 0.022 | 0.040 | 0.085 |
| | $\tau^+_{int}$ | 0.059 | 0.150 | 0.362 | 0.031 | 0.048 | 0.080 | 0.059 | 0.160 | 0.371 | 0.026 | 0.051 | 0.116 |
| | $\mathbf{t}_{int}$ | 0.135 | 0.288 | 0.509 | 0.141 | 0.237 | 0.330 | 0.134 | 0.290 | 0.505 | 0.134 | 0.244 | 0.355 |
| $f^{(3)}_t$ | $\tau^0_{adj}$ | 0.064 | 0.266 | 0.417 | 0.079 | 0.223 | 0.302 | 0.060 | 0.268 | 0.426 | 0.082 | 0.217 | 0.293 |
| | $\tau^{+0}_{adj}$ | 0.059 | 0.249 | 0.404 | 0.037 | 0.144 | 0.249 | 0.056 | 0.252 | 0.418 | 0.036 | 0.140 | 0.243 |
| | $\mathbf{t}^0_{int}$ | 0.110 | 0.259 | 0.435 | 0.105 | 0.208 | 0.293 | 0.108 | 0.268 | 0.445 | 0.101 | 0.217 | 0.304 |
| | $\tau_{int}$ | 0.060 | 0.141 | 0.322 | 0.074 | 0.086 | 0.133 | 0.062 | 0.146 | 0.325 | 0.072 | 0.091 | 0.134 |
| | $\tau^+_{int}$ | 0.048 | 0.120 | 0.314 | 0.016 | 0.028 | 0.064 | 0.052 | 0.129 | 0.317 | 0.014 | 0.029 | 0.068 |
| | $\mathbf{t}_{int}$ | 0.134 | 0.278 | 0.468 | 0.140 | 0.213 | 0.314 | 0.135 | 0.279 | 0.474 | 0.129 | 0.229 | 0.322 |

**Table 7.** Relative Rejection Frequencies of Tests, $T = 200$, $T_1 = 100$, $\bar{c} = 0$, Nominal Significance Level 5%, Random Initial Values

| Shift function | Test | DGP (6.1), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.1), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_t^{(1)}$ | $\tau_{adj}^0$ | 0.035 | 0.653 | 0.867 | 0.017 | 0.494 | 0.726 | 0.033 | 0.633 | 0.850 | 0.017 | 0.472 | 0.650 |
| | $\tau_{adj}^{+0}$ | 0.050 | 0.693 | 0.878 | 0.041 | 0.591 | 0.774 | 0.044 | 0.678 | 0.869 | 0.040 | 0.556 | 0.706 |
| | $\mathbf{t}_{int}^0$ | 0.055 | 0.698 | 0.966 | 0.054 | 0.495 | 0.799 | 0.053 | 0.703 | 0.964 | 0.058 | 0.519 | 0.781 |
| | $\tau_{int}$ | 0.028 | 0.469 | 0.907 | 0.009 | 0.130 | 0.513 | 0.024 | 0.466 | 0.899 | 0.007 | 0.125 | 0.476 |
| | $\tau_{int}^+$ | 0.050 | 0.617 | 0.943 | 0.064 | 0.434 | 0.753 | 0.053 | 0.613 | 0.933 | 0.061 | 0.415 | 0.732 |
| | $\mathbf{t}_{int}$ | 0.059 | 0.556 | 0.910 | 0.058 | 0.387 | 0.675 | 0.061 | 0.568 | 0.913 | 0.062 | 0.394 | 0.687 |
| $f_t^{(2)}$ | $\tau_{adj}^0$ | 0.051 | 0.610 | 0.819 | 0.028 | 0.472 | 0.684 | 0.053 | 0.608 | 0.810 | 0.028 | 0.464 | 0.656 |
| | $\tau_{adj}^{+0}$ | 0.051 | 0.614 | 0.824 | 0.033 | 0.482 | 0.691 | 0.053 | 0.614 | 0.812 | 0.029 | 0.474 | 0.664 |
| | $\mathbf{t}_{int}^0$ | 0.071 | 0.629 | 0.904 | 0.068 | 0.474 | 0.724 | 0.073 | 0.631 | 0.907 | 0.067 | 0.476 | 0.708 |
| | $\tau_{int}$ | 0.046 | 0.489 | 0.844 | 0.011 | 0.142 | 0.509 | 0.047 | 0.491 | 0.896 | 0.011 | 0.161 | 0.511 |
| | $\tau_{int}^+$ | 0.044 | 0.492 | 0.891 | 0.020 | 0.174 | 0.531 | 0.043 | 0.497 | 0.889 | 0.020 | 0.195 | 0.534 |
| | $\mathbf{t}_{int}$ | 0.071 | 0.558 | 0.877 | 0.075 | 0.398 | 0.682 | 0.074 | 0.559 | 0.867 | 0.077 | 0.393 | 0.658 |
| $f_t^{(3)}$ | $\tau_{adj}^0$ | 0.049 | 0.571 | 0.747 | 0.051 | 0.507 | 0.656 | 0.045 | 0.553 | 0.742 | 0.050 | 0.492 | 0.637 |
| | $\tau_{adj}^{+0}$ | 0.046 | 0.563 | 0.741 | 0.028 | 0.463 | 0.635 | 0.043 | 0.542 | 0.737 | 0.026 | 0.447 | 0.619 |
| | $\mathbf{t}_{int}^0$ | 0.068 | 0.534 | 0.809 | 0.074 | 0.385 | 0.625 | 0.065 | 0.536 | 0.802 | 0.067 | 0.389 | 0.596 |
| | $\tau_{int}$ | 0.051 | 0.484 | 0.854 | 0.075 | 0.241 | 0.512 | 0.042 | 0.488 | 0.861 | 0.075 | 0.237 | 0.529 |
| | $\tau_{int}^+$ | 0.036 | 0.479 | 0.856 | 0.012 | 0.167 | 0.486 | 0.034 | 0.473 | 0.861 | 0.013 | 0.170 | 0.497 |
| | $\mathbf{t}_{int}$ | 0.075 | 0.545 | 0.862 | 0.075 | 0.389 | 0.634 | 0.079 | 0.552 | 0.861 | 0.082 | 0.399 | 0.645 |

**Table 8.** Relative Rejection Frequencies of Tests, $T = 100$, $T_1 = 50$, $\bar{c} = 0$, Nominal Significance Level 5%, Zero Initial Values

| Shift function | Test | DGP (6.1), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.1), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.5$ $\rho = 1$ | 0.9 | 0.8 | DGP (6.2), $b_1 = 0.8$ $\rho = 1$ | 0.9 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_t^{(1)}$ | $\tau_{adj}^0$ | 0.049 | 0.512 | 0.845 | 0.015 | 0.259 | 0.584 | 0.044 | 0.446 | 0.730 | 0.017 | 0.210 | 0.404 |
| | $\tau_{adj}^{+0}$ | 0.067 | 0.617 | 0.889 | 0.066 | 0.487 | 0.751 | 0.069 | 0.560 | 0.785 | 0.067 | 0.427 | 0.580 |
| | $\mathbf{t}_{int}^0$ | 0.047 | 0.257 | 0.596 | 0.046 | 0.197 | 0.327 | 0.050 | 0.271 | 0.594 | 0.053 | 0.208 | 0.329 |
| | $\tau_{int}$ | 0.021 | 0.094 | 0.325 | 0.001 | 0.005 | 0.043 | 0.019 | 0.091 | 0.317 | 0.002 | 0.004 | 0.032 |
| | $\tau_{int}^+$ | 0.072 | 0.238 | 0.571 | 0.072 | 0.181 | 0.321 | 0.070 | 0.229 | 0.559 | 0.070 | 0.190 | 0.317 |
| | $\mathbf{t}_{int}$ | 0.067 | 0.245 | 0.583 | 0.073 | 0.171 | 0.310 | 0.067 | 0.244 | 0.561 | 0.067 | 0.170 | 0.319 |
| $f_t^{(2)}$ | $\tau_{adj}^0$ | 0.055 | 0.459 | 0.782 | 0.024 | 0.248 | 0.500 | 0.057 | 0.414 | 0.668 | 0.032 | 0.246 | 0.435 |
| | $\tau_{adj}^{+0}$ | 0.059 | 0.476 | 0.786 | 0.024 | 0.269 | 0.517 | 0.064 | 0.427 | 0.681 | 0.034 | 0.271 | 0.451 |
| | $\mathbf{t}_{int}^0$ | 0.083 | 0.288 | 0.570 | 0.085 | 0.219 | 0.336 | 0.088 | 0.296 | 0.560 | 0.095 | 0.203 | 0.344 |
| | $\tau_{int}$ | 0.041 | 0.150 | 0.366 | 0.009 | 0.033 | 0.068 | 0.046 | 0.142 | 0.349 | 0.011 | 0.037 | 0.088 |
| | $\tau_{int}^+$ | 0.052 | 0.157 | 0.391 | 0.017 | 0.038 | 0.083 | 0.053 | 0.155 | 0.373 | 0.021 | 0.045 | 0.115 |
| | $\mathbf{t}_{int}$ | 0.136 | 0.340 | 0.588 | 0.130 | 0.260 | 0.379 | 0.129 | 0.316 | 0.562 | 0.126 | 0.251 | 0.370 |
| $f_t^{(3)}$ | $\tau_{adj}^0$ | 0.058 | 0.420 | 0.630 | 0.049 | 0.319 | 0.496 | 0.061 | 0.402 | 0.559 | 0.055 | 0.312 | 0.463 |
| | $\tau_{adj}^{+0}$ | 0.052 | 0.394 | 0.624 | 0.023 | 0.245 | 0.457 | 0.054 | 0.375 | 0.550 | 0.027 | 0.237 | 0.401 |
| | $\mathbf{t}_{int}^0$ | 0.112 | 0.213 | 0.400 | 0.120 | 0.174 | 0.250 | 0.106 | 0.230 | 0.409 | 0.115 | 0.170 | 0.257 |
| | $\tau_{int}$ | 0.054 | 0.157 | 0.379 | 0.046 | 0.076 | 0.133 | 0.054 | 0.153 | 0.360 | 0.051 | 0.094 | 0.135 |
| | $\tau_{int}^+$ | 0.043 | 0.140 | 0.365 | 0.009 | 0.026 | 0.072 | 0.043 | 0.132 | 0.348 | 0.009 | 0.029 | 0.078 |
| | $\mathbf{t}_{int}$ | 0.126 | 0.276 | 0.467 | 0.132 | 0.211 | 0.308 | 0.128 | 0.272 | 0.475 | 0.136 | 0.224 | 0.318 |

**Table 9.** Relative Rejection Frequencies of Tests, $T = 100$, $T_1 = 50$, $\bar{c} = 0$, Nominal Significance Level 5%, Initial Values 5

| Shift function | Test | DGP (6.1), $b_1 = 0.5$ | | | DGP (6.1), $b_1 = 0.8$ | | | DGP (6.2), $b_1 = 0.5$ | | | DGP (6.2), $b_1 = 0.8$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\rho = 1$ | 0.9 | 0.8 | $\rho = 1$ | 0.9 | 0.8 | $\rho = 1$ | 0.9 | 0.8 | $\rho = 1$ | 0.9 | 0.8 |
| $f_t^{(1)}$ | $\tau_{adj}^0$ | 0.049 | 0.126 | 0.104 | 0.015 | 0.180 | 0.234 | 0.042 | 0.044 | 0.053 | 0.016 | 0.166 | 0.151 |
| | $\tau_{adj}^{+0}$ | 0.067 | 0.168 | 0.147 | 0.066 | 0.372 | 0.405 | 0.064 | 0.066 | 0.070 | 0.046 | 0.349 | 0.272 |
| | $\mathbf{t}_{int}^0$ | 0.064 | 0.261 | 0.593 | 0.066 | 0.183 | 0.311 | 0.067 | 0.329 | 0.659 | 0.062 | 0.243 | 0.410 |
| | $\tau_{int}$ | 0.021 | 0.084 | 0.265 | 0.001 | 0.004 | 0.034 | 0.021 | 0.060 | 0.155 | 0.001 | 0.002 | 0.017 |
| | $\tau_{int}^+$ | 0.072 | 0.214 | 0.429 | 0.072 | 0.179 | 0.293 | 0.068 | 0.189 | 0.292 | 0.062 | 0.222 | 0.298 |
| | $\mathbf{t}_{int}$ | 0.070 | 0.208 | 0.421 | 0.077 | 0.169 | 0.263 | 0.076 | 0.166 | 0.311 | 0.080 | 0.133 | 0.177 |
| $f_t^{(2)}$ | $\tau_{adj}^0$ | 0.055 | 0.162 | 0.153 | 0.024 | 0.190 | 0.260 | 0.062 | 0.041 | 0.059 | 0.037 | 0.164 | 0.105 |
| | $\tau_{adj}^{+0}$ | 0.059 | 0.165 | 0.168 | 0.024 | 0.203 | 0.270 | 0.069 | 0.045 | 0.063 | 0.055 | 0.194 | 0.120 |
| | $\mathbf{t}_{int}^0$ | 0.094 | 0.292 | 0.580 | 0.095 | 0.204 | 0.330 | 0.095 | 0.355 | 0.665 | 0.104 | 0.270 | 0.413 |
| | $\tau_{int}$ | 0.041 | 0.117 | 0.231 | 0.009 | 0.032 | 0.066 | 0.047 | 0.096 | 0.166 | 0.011 | 0.032 | 0.060 |
| | $\tau_{int}^+$ | 0.052 | 0.126 | 0.226 | 0.017 | 0.042 | 0.082 | 0.051 | 0.116 | 0.160 | 0.025 | 0.106 | 0.140 |
| | $\mathbf{t}_{int}$ | 0.122 | 0.243 | 0.373 | 0.133 | 0.222 | 0.300 | 0.116 | 0.190 | 0.272 | 0.146 | 0.181 | 0.197 |
| $f_t^{(3)}$ | $\tau_{adj}^0$ | 0.058 | 0.182 | 0.170 | 0.048 | 0.256 | 0.296 | 0.072 | 0.053 | 0.050 | 0.185 | 0.279 | 0.137 |
| | $\tau_{adj}^{+0}$ | 0.052 | 0.161 | 0.157 | 0.023 | 0.189 | 0.245 | 0.056 | 0.047 | 0.045 | 0.027 | 0.142 | 0.087 |
| | $\mathbf{t}_{int}^0$ | 0.112 | 0.268 | 0.506 | 0.120 | 0.185 | 0.288 | 0.111 | 0.333 | 0.585 | 0.125 | 0.231 | 0.362 |
| | $\tau_{int}$ | 0.053 | 0.137 | 0.262 | 0.046 | 0.081 | 0.121 | 0.057 | 0.144 | 0.216 | 0.173 | 0.283 | 0.264 |
| | $\tau_{int}^+$ | 0.043 | 0.124 | 0.248 | 0.009 | 0.031 | 0.071 | 0.045 | 0.111 | 0.199 | 0.012 | 0.055 | 0.121 |
| | $\mathbf{t}_{int}$ | 0.126 | 0.268 | 0.435 | 0.132 | 0.216 | 0.298 | 0.117 | 0.253 | 0.391 | 0.139 | 0.199 | 0.260 |