On nonparametric stationarity and unit root tests under AO adjustment with possible persistent effects

Julio A. Afonso-Rodríguez
Department of Institutional Economics, Economic Statistics and Econometrics
Scholl of Economics and Business Administration. University of La Laguna
Campus de Guajara. Camino La Hornera s/n
38071. San Cristóbal de La Laguna. Tenerife. Islas Canarias
email: jafonsor@ull.es. Tfno.: 922.317041, Fax: 922.317042

February 2012

Abstract

It is widely known that the empirical size and power profile of many unit root and stationarity tests are not invariant to additive outlying observations in the time series. The common practice is to take into account this observations by introducing dummy variables in the auxiliary regressions of this test procedures. In this paper we study the asymptotic and finite sample properties of a set of nonparametric tests for the null of stationarity against a unit root (Kwiatkowski et.al. (1992), Xiao (2001) and Giraitis et.al. (2003)), and also for the variance-ratio nonparametric test of the opposite hypothesis (Breitung (2002)), when the true location and nature of the outlier can differ from the assumed in the specification of the estimated regression. Under a general framework, that allow for a possible misspecified location and a persistent effect of the true contaminating process, we consider proper assumptions on the outlier magnitudes related to the sample size to get a well defined asymptotic null distribution for these test statistics. We show that the stationarity tests are quite robust to a misspecification in the location and true nature of the outlier, even for a relatively high magnitude of the effect, except in the case of an extremely persistent effect of the perturbation (temporary change) where we find a significant increase in their empirical size. The variance-ratio test statistic behaves quite similarly, although it is more robust than the other test statistics, even under a high degree of persistence.

Key words: Stationarity tests, variance-ratio unit root test, additive outlier, dummy variables, impulse and step functions

JEL classification codes: C12, C22, C58

Subject area: Quantitative Methods in Economics and Business Administration

ON NONPARAMETRIC STATIONARITY AND UNIT ROOT TESTS UNDER AO ADJUSTMENT WITH POSSIBLE PERSISTENT EFFECTS

1. Introduction

Following the contribution of Kwiatkowski et.al. (1992), testing for stationarity against a unit root has become a central part of theoretical and applied time series econometrics. However, due to the lack of robustness of these test procedures under misspecification of the basic data generating process and some problems in their implementation in finite samples, it is usually argue that it can be of interest to test the opposite hypothesis, that is of a unit root, when investigating the dynamic properties of a time series. Section 2 presents the general structure of the components model that allows to build three related nonparametric test statistics for testing the null of stationarity against the alternative of a unit root for the observed process and an additional one, the variance-ratio test statistic, for testing the opposite hypothesis. We review the main stochastic properties of these tests, in terms of their asymptotic null distributions, asymptotic power profiles and the effects of several changes in the nature of the stochastic components of the model. Also, we will review the effects of different types of misspecification in the systematic component of the model, paying special attention to the effects of outlying observations in the sample. To some extent there is a connection between the addition of dummy variables to pick up the effect of an outlier and the existence of a structural break in the systematic component of the model. However, there are some important differences in the treatment of each problem in order to determine the effect of a possible misspecification in location and/or in the number of events of each type. Section 3 presents a modification of the underlying components model that allows to capture the effect of additive outliers and to analyze the consequences of outliers with persistent effects as well as a possible misspecification in the location. Through a convenient normalization of the outlier magnitude, we are able to determine the finite sample and asymptotic effects of different types of contamination in the stochastic properties of all these test statistics.

2. Review of nonparametric tests for stationarity and for a unit root

Along with unit root tests, tests for the null of stationarity have often been used in practice. Among these, there is a set which are based on different fluctuation measures of the scaled partial sum process of OLS residuals obtained in an auxiliary regression resulting from the following generalized local-level model (basic DGP)

$$Y_t = d_t(p,\lambda) + \rho_t + \varepsilon_{0,t} \quad t = 1,2,\ldots,n \tag{2.1}$$

$$\rho_t = \rho_{t-1} + \varepsilon_{1,t} \tag{2.2}$$

In (2.1) and (2.2) it is assumed that the error terms $\varepsilon_{i,t}$ (i = 0, 1) are stationary zero mean sequences that satisfy appropriate conditions to jointly verify a bivariate invariance principle such that

$$\begin{pmatrix} W_{n,0}(r) \\ W_{n,1}(r) \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{j=1}^{t-1} \begin{pmatrix} \varepsilon_{0,j} \\ \varepsilon_{1,j} \end{pmatrix} \Rightarrow \boldsymbol{B}(r) = \boldsymbol{\Omega}^{1/2} \boldsymbol{W}(r) \quad \frac{t-1}{n} \le r < \frac{t}{n}, t = 1,\ldots,n \tag{2.3}$$

with $W_{n,i}(r) = n^{-1/2} S_{[nr],i} = n^{-1/2} \sum_{j=1}^{[nr]} \varepsilon_{i,j}$, and where "⇒" indicates weak convergence of the associated probability measures. $\boldsymbol{B}(r) = (B_0(r), B_1(r))'$ denotes a bivariate Brownian motion process with covariance matrix $\boldsymbol{\Omega}$, $\boldsymbol{W}(r) = (W_0(r), W_1(r))'$ a bivariate standard Brownian motion process (with $W_0(r)$ and $W_1(r)$ mutually independent univariate standard Brownian motion processes), and

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_0^2 & \rho_{01}\omega_0\omega_1 \\ & \omega_1^2 \end{pmatrix} \tag{2.4}$$

is the long-run covariance matrix of $\boldsymbol{\varepsilon}_t = (\varepsilon_{0,t}, \varepsilon_{1,t})'$, $\boldsymbol{\Omega} = \lim_{n\to\infty} Var(n^{-1/2} \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t)$, with

$$\omega_i^2 = \lim_{n\to\infty} Var[n^{-1/2} \sum_{t=1}^{n} \varepsilon_{i,t}] = \sigma_i^2 + 2\sum_{h=1}^{\infty} \gamma_i(h) = \sigma_i^2 \cdot (1 + 2\sum_{h=1}^{\infty} \rho_i(h)) \tag{2.5}$$

the long-run variances, where $\sigma_i^2 = \lim_{n\to\infty} (1/n) \sum_{t=1}^{n} E[\varepsilon_{i,t}^2] \ge 0$, i = 0, 1, and $|\rho_{01}| \le 1$. When testing for the null of stationarity, that is $\sigma_1^2 = 0$ (and thus $\omega_1^2 = 0$), it is common to assume that both error processes are mutually independent, so that $\rho_{01} = 0$ in (2.4) and $B_i(r) = \omega_i \cdot W_i(r)$, i = 0, 1. In (2.1), $d_t(p,\lambda)$ is the deterministic kernel which is usually parameterized as a generalized polynomial time-trend function

$$d_t(p,\lambda) = \boldsymbol{x}'_{t,p} \boldsymbol{\beta}_{p,t}(\lambda) \tag{2.6}$$

where $\boldsymbol{x}'_{t,p} = (1, t, \ldots, t^p)$, $p \ge 0$, $\boldsymbol{x}'_{t,0} = 1$ for p = 0, and

$$\boldsymbol{\beta}_{p,t}(\lambda) = \boldsymbol{\beta}_p + h_t(\lambda)\boldsymbol{\alpha}_p = \boldsymbol{\beta}_p(1 - h_t(\lambda)) + h_t(\lambda)\boldsymbol{\theta}_p \tag{2.7}$$

with $\boldsymbol{\alpha}_p = \boldsymbol{\theta}_p - \boldsymbol{\beta}_p$, $0 < \lambda < 1$ and $h_t(\lambda)$ a step function defined as $h_t(\lambda) = I(t > [n\lambda])$, thus incorporating the possibility of a deterministic structural change in the systematic component. The standard case of no structural change results with λ = 0, so that $\boldsymbol{\beta}_{p,t}(0) = \boldsymbol{\theta}_p = \boldsymbol{\beta}_p + \boldsymbol{\alpha}_p$ for all t = 1, ..., n, or with λ = 1 where $\boldsymbol{\beta}_{p,t}(1) = \boldsymbol{\beta}_p$. Furthermore, associated to the given specification of the deterministic component, it is assumed that there exist a diagonal, non-singular and deterministic scaling matrix $\boldsymbol{D}_{p,n}$, $\boldsymbol{D}_{p,n} = diag(d_{0,n}, d_{1,n}, ..., d_{p,n})$, such that $\boldsymbol{D}_{p,n}\boldsymbol{x}_{t,p} = \boldsymbol{x}_p(d_{t,n})$ is a element of the unit interval $(0,1]^{p+1}$. Thus, as long as $n \to \infty$, $\boldsymbol{x}_p(d_{[nr],n}) \to \boldsymbol{x}_p(r) \in [0,1]^{p+1}$ with $r \in$ [0, 1]. Also, calling $\boldsymbol{z}_{p,n}(\frac{[nr]}{n}) = n^{-1}\sum_{j=1}^{[nr]} \boldsymbol{x}_p(\frac{j}{n})$ and $\bar{\boldsymbol{U}}_{[nb]-[na]}(p) = n^{-1}\sum_{j=[na]+1}^{[nb]} \boldsymbol{x}_p(\frac{j}{n})\boldsymbol{x}'_p(\frac{j}{n})$ for $0 \le a < b \le 1$, we then have the following well defined limits

$$\boldsymbol{z}_{p,n}(\tfrac{[nr]}{n}) \to \boldsymbol{z}_p(r) = \int_0^r \boldsymbol{x}_p(s)ds \tag{2.8}$$

$$\bar{\boldsymbol{U}}_{[nb]-[na]}(p) \to \boldsymbol{U}_{b-a}(p) = \int_a^b \boldsymbol{x}_p(s)\boldsymbol{x}'_p(s)ds < \infty \tag{2.9}$$

For the polynomial trend function we then have $\boldsymbol{D}_{p,n} = diag(1, \frac{1}{n}, ..., \frac{1}{n^p})$ and $\boldsymbol{x}_p(r) = (1, r, ..., r^p)'$ which satisfies the above limit results.

With all these requirement, the model (2.1) and (2.2) allows to specify the following auxiliary linear regression that, for a given value of $\lambda \in (0,1)$, will be estimated by OLS

$$Y_t = \boldsymbol{x}'_{t,p}\boldsymbol{\beta}_{p,t}(\lambda) + \eta_t = \boldsymbol{x}'_{t,p}\boldsymbol{\beta}_p + \boldsymbol{x}'_{t,p}h_t(\lambda)\boldsymbol{\alpha}_p + \eta_t \quad t = 1, ..., n \tag{2.10}$$

and where the disturbance term, η_t, is given by the sum of an stationary component $\varepsilon_{0,t}$ and an integrated component ρ_t when $\sigma_1^2 = E[\varepsilon_{1,t}^2] > 0$, that is

$$\eta_t = \varepsilon_{0,t} + \rho_t = \varepsilon_{0,t} + \rho_0 + \sum_{j=1}^{t} \varepsilon_{1,j} \tag{2.11}$$

Within this framework, the hypothesis of stationarity corresponds to $\sigma_1^2 = 0$, so that $\eta_t = \varepsilon_{0,t} + \rho_0 = O_p(1)$, while the hypothesis of a unit root process for Y_t in (2.1) is given by $\sigma_1^2 > 0$, with $\eta_t = O_p(\kappa \cdot n^{1/2})$ and $\kappa^2 = \sigma_1^2 / \sigma_0^2$ when $\sigma_0^2 > 0$, or $\eta_t = O_p(n^{1/2})$ when $\sigma_0^2 = 0$. The initial value ρ_0 can be considered both negligible ($\rho_0 = 0$ or $\rho_0 = o_p(1)$) or not (ρ_0 fixed and finite or $\rho_0 = O_p(1)$), without affecting all the subsequent analysis because of the normalizing by functions of the sample size used in the building of the different test statistics. If the deterministic component includes a constant term and ρ_0 is treated as fixed and finite, then it can be simply added to the intercept. This framework is the standard case and is very different from the one used in Müller (2005) when the stationary component is modelled as a mean reverting process with strong autocorrelation in a local-to-unity asymptotic analysis. All the test statistics that we are going to study use the basic DGP in (2.1) and (2.2), and thus can be obtained using different functionals of the residuals from the OLS fitting of (2.10),

$$\begin{aligned}\hat{\eta}_{t,p}(\lambda) &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,t}(\lambda) = Y_t - \boldsymbol{x}'_{t,p}[\hat{\boldsymbol{\beta}}_{p,n}(\lambda) + h_t(\lambda)\hat{\boldsymbol{\alpha}}_{p,n}(\lambda)] \\ &= Y_t - (1-h_t(\lambda))\boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n}(\lambda) - h_t(\lambda)\boldsymbol{x}'_{t,p}\hat{\boldsymbol{\theta}}_{p,n}(\lambda) \qquad t = 1,\dots,n\end{aligned} \tag{2.12}$$

for a given value of the breakpoint $0 < \lambda < 1$ such that $[n\lambda]$, $n-[n\lambda] \geq p+1$, with $\hat{\boldsymbol{\beta}}_{p,n}(\lambda)$ and $\hat{\boldsymbol{\theta}}_{p,n}(\lambda)$ the OLS estimator of the vector parameter for the first and second subsample, respectively. In the case of no structural break, we have the standard result

$$\hat{\eta}_{t,p} = Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n} = \eta_t - \boldsymbol{x}'_{t,p}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p) \tag{2.13}$$

Under the model specification in (2.10), the residuals can be also written as

$$\begin{aligned}\hat{\eta}_{t,p}(\lambda) &= \eta_t - (1-h_t(\lambda))\boldsymbol{x}'_{t,p}(\hat{\boldsymbol{\beta}}_{p,n}(\lambda) - \boldsymbol{\beta}_p) - h_t(\lambda)\boldsymbol{x}'_{t,p}(\hat{\boldsymbol{\theta}}_{p,n}(\lambda) - \boldsymbol{\theta}_p) \\ &= \eta_t - n^{-v}(1-h_t(\lambda))\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}^{-1}_{[n\lambda]}(p)n^{v-1}\sum_{j=1}^{[n\lambda]}\boldsymbol{x}_p(\tfrac{j}{n})\eta_j \\ &\quad - n^{-v}h_t(\lambda)\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p)n^{v-1}\sum_{j=[n\lambda]+1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})\eta_j\end{aligned} \tag{2.14}$$

thus reflecting the stochastic properties of the sequence of error terms η_t, $t = 1, \dots, n$. Since from the work of Busetti and Harvey (2002, 2003) is well established the asymptotic distributions and the way to proceed in practice with a structural break in the systematic component, we use this specification in what follows. The case of no break occurring in the sample can be obtain simply as a particular case, as in standard regression problems. Kurozumi (2002) determines the limiting distribution of the one-sided LM test statistic (Kwiatkowski, et.al. (1992)) for the null of stationarity against a unit root alternative with different particular parameterizations of the structural break in the leading cases p = 0, 1, and under the sequence of local alternatives given by $\sigma_1^2 = \sigma_0^2 \cdot (\frac{\theta}{n})^2$, with θ a constant, also including the absence of structural break as a particular case. The model with a sudden and instantaneous change in the structure of the polynomial trend function is termed by Kurozumi (2002) as the additive outlier (AO) model, in the sense that when the structural change occurred the shock affects the observations only at one time. Also considers the case where the structural change disturbs the variables with lagged effects, which is termed the innovational outlier (IO) model with $\boldsymbol{\beta}_{p,t}(\lambda) = \boldsymbol{\beta}_p + \psi_m(L)h_t(\lambda)\boldsymbol{\alpha}_p$ in (2.7), $\psi_m(L) = 1 + \psi_1 L + \dots + \psi_m L^m$ a mth order lag polynomial. With this we have that $\psi_m(L)h_t(\lambda) = \tau_0 h_t(\lambda) + \sum_{i=1}^{m} I_{t-i}(\lambda)\tau_i$, $I_{t-i}(\lambda) = I(t = \lambda + i)$, $\tau_0 = \psi_m(1)$, and $\tau_i = -\sum_{j=i}^{m}\psi_j$, so that the additional regressors resulting from the interaction with the indicator functions are asymptotically negligible, and thus the limiting distributions remains unaltered. From the analysis in the next section of the paper it will be clear that our approach deals with the strict notion of an outlier in the sample space of observations for the dependent variable, not in the regressors space as in Kurozumi (2002), thus making the structure of our analysis, the main results and conclusions completely different.

In what follows we will introduce the test statistics subject to analysis in the present study, three of them for testing the null hypothesis of stationarity around the deterministic component (trend stationarity) against the alternative of a unit root process (difference stationarity) and an additional one for testing the opposite hypothesis. The four test statistics have in common their nonparametric nature and the use of the OLS residuals from the estimation of the auxiliary regression in (2.10), but each one is based on a particular measure of the fluctuation of these trying to capture the degree of persistence of a shock in the non-systematic part of the DGP.

Given the sequence of OLS residuals $\hat{\eta}_{t,p}(\lambda)$ in (2.12)-(2.14) and the partial sum process $\hat{S}_{t,p}(\lambda)=\sum_{j=1}^{t}\hat{\eta}_{j,p}(\lambda)$, t = 1, ..., n, we consider the following nonparametric univariate tests statistics

$$\hat{M}_{n,p}^{(1)}(m_n,\lambda)=[n\cdot\hat{\omega}_n^2(m_n)]^{-1}\sum_{t=1}^{n}(\tfrac{1}{\sqrt{n}}\hat{S}_{t,p}(\lambda))^2, \qquad (2.15)$$

$$\hat{M}_{n,p}^{(2)}(m_n,\lambda)=[n\cdot\hat{\omega}_n^2(m_n)]^{-1}\left\{\sum_{t=1}^{n}(\tfrac{1}{\sqrt{n}}\hat{S}_{t,p}(\lambda))^2-n^{-1}\left(\sum_{t=1}^{n}\tfrac{1}{\sqrt{n}}\hat{S}_{t,p}(\lambda)\right)^2\right\} \qquad (2.16)$$

or, alternatively, $\hat{M}_{n,p}^{(2)}(m_n,\lambda)=\hat{M}_{n,p}^{(1)}(m_n,\lambda)-\frac{1}{n^2\cdot\hat{\omega}_n^2(m_n)}(\sum_{t=1}^{n}\frac{1}{\sqrt{n}}\hat{S}_{t,p}(\lambda))^2$, and

$$\hat{M}_{n,p}^{(3)}(m_n,\lambda)=\hat{\omega}_n^{-1}(m_n)\cdot\max_{t=1,\ldots,n}\left|\tfrac{1}{\sqrt{n}}\hat{S}_{t,p}(\lambda)-\tfrac{t}{n}(\tfrac{1}{\sqrt{n}}\hat{S}_{n,p}(\lambda))\right| \qquad (2.17)$$

as the basis for developing three of the most used tests of the null hypothesis of stationarity against the alternative of a linear unit root procedure. These are the KPSS test by Kwiatkowski et.al. (1992), i = 1, the V/S test by Giraitis et.al. (2003), i = 2, and the KS test by Xiao (2001), i = 3. The KPSS test statistic in (2.15) is the one-sided LM test for the null hypothesis of stationarity, H_0: $\sigma_1^2=0$, against the alternative of a unit root, H_1: $\sigma_1^2>0$, and can also be interpreted as the LBI test statistic under the additional assumption of normality. In (2.16) $\sum_{t=1}^{n}\hat{S}_{t,p}(1)=0$, so that $\hat{M}_{n,p}^{(1)}(m_n,1)=\hat{M}_{n,p}^{(2)}(m_n,1)$ when regression contains a linear trend (p=1), and $\hat{S}_{n,p}(\lambda)=0$ in (2.17) if regression contains a constant term. Notice that while the KPSS test statistic uses de Cramér-von

Mises measure of fluctuation for the series, the Xiao test statistic is based on the Kolmogorov-Smirnov measure of fluctuation. All three test statistics are based on a nonparametric correction for weak dependence in the error sequences $\varepsilon_{i,t}$, i = 0, 1, through a long-run variance kernel estimator of $\omega_\eta^2 = \lim_{n\to\infty} Var[n^{-1/2}\sum_{t=1}^{n}\eta_t]$ given by

$$\hat{\omega}_n^2(m_n) = \sum_{j=-(n-1)}^{n-1} w(j,m_n)\hat{\gamma}_n(j) = \hat{\gamma}_n(0) + 2\sum_{j=1}^{n-1} w(j,m_n)\hat{\gamma}_n(j) \tag{2.18}$$

where $\hat{\gamma}_n(j) = \hat{\gamma}_n(-j) = n^{-1}\sum_{t=j+1}^{n}\hat{\eta}_{t,p}(\lambda)\hat{\eta}_{t-j,p}(\lambda)$ is the jth order residual autocovariance, $w(j, m_n)$ is a weighting or kernel function and m_n is the bandwidth or lag truncation parameter, usually determined as a fixed proportion of the sample size (deterministic bandwidth) or, alternatively, as the outcome of a data-dependent or automatic rule (stochastic bandwidth) that must satisfy the conditions $m_n^{-1} = o_p(1)$ and $m_n = O_p(n^{1/2-a}) = o_p(n^{1/2})$ for some $0<a<1/2$. Among all existing proposals for choosing a particular combination of a symmetric kernel function with finite support and bounded variation and a stochastic bandwidth, the one that appears to provide better results in finite samples is the Bartlett window, $w(j, m_n) = 1-j/(m_n+1)$ for j = 1, ..., m_n, with a bounded version of the automatic bandwidth determination rule proposed in Andrews (1991) (see, e.g., Kurozumi (2002), Hobijn, et.al. (2004), Carrion-i-Silvestre and Sansó (2006), Jönsson, K. (2006), and Xiao and Lima (2007) for more details on this issue). Under the null of stationarity, where $\eta_t = \varepsilon_{0,t}$, and the given conditions on the bandwidth parameter $\hat{\omega}_n^2(m_n)$ is a consistent estimator of the long-run variance of $\varepsilon_{0,t}$, that is $\hat{\omega}_n^2(m_n) \to^p \omega_0^2$ as $n \to\infty$. Among others in the vast literature on consistent estimation of long-run variance matrices, de Jong and Davidson (2000) proof the consistency of kernel estimators of long-run variances and covariances even under rates of convergence of estimates different from the usual $O_p(n^{-1/2})$ and very general possible characterizations of the weak dependence for the stationary sequence $\varepsilon_{0,t}$. To complement and robustify the results of these stationarity tests, we consider the normalized variance-ratio unit root test by Breitung (2002), given by

$$\bar{\rho}_{n,p}(\lambda) = n^{-1}\hat{\rho}_{n,p}(\lambda) = \left\{\tfrac{1}{n}\sum_{t=1}^{n}(\tfrac{1}{\sqrt{n}}\hat{\eta}_{t,p}(\lambda))^2\right\}^{-1}\tfrac{1}{n}\sum_{t=1}^{n}(\tfrac{1}{n\sqrt{n}}\hat{S}_{t,p}(\lambda))^2 \tag{2.19}$$

From (2.15) we have the following relation with the KPSS test statistic,

$$\bar{\rho}_{n,p}(\lambda) = \hat{M}_{n,p}^{(1)}(m_n,\lambda)\tfrac{1}{n}\hat{\gamma}_n^{-1}(0)\hat{\omega}_n^2(m_n) \tag{2.20}$$

where $\hat{\gamma}_n^{-1}(0)\hat{\omega}_n^2(m_n) \to^p \gamma^{-1}(0)\omega_0^2 = 1 + 2\sum_{j=1}^{\infty}\rho_0(j)$ under stationarity, so that the test statistic $\bar{\rho}_{n,p}(\lambda)$ approaches zero and this is a left tailed test that rejects for small values of (2.19).

Following our own formulation, under the hypothesis of stationarity, that is H_0: $\sigma_1^2 = 0$, the asymptotic distribution of the scaled partial sum of OLS residuals from (2.10) is given by

$$n^{-1/2}\hat{S}_{[nr],p}(\lambda) \Rightarrow \omega_0 \cdot B_p(r,\lambda) \tag{2.21}$$

where

$$B_p(r,\lambda) = \begin{cases} W_0(r) - C_p(r,\lambda) & \text{for } r \le \lambda \\ W_0(r) - (C_p(\lambda,\lambda) + D_p(r,\lambda)) & \text{for } r > \lambda \end{cases} \tag{2.22}$$

is a two-piecewise $(p+1)th$-level standard Brownian Bridge process, with

$$C_p(r,\lambda) = \int_0^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_\lambda^{-1}(p)\int_0^\lambda \boldsymbol{x}_p(s)dW_0(s), \tag{2.23}$$

$$D_p(r,\lambda) = \int_\lambda^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_{1-\lambda}^{-1}(p)\int_\lambda^1 \boldsymbol{x}_p(s)dW_0(s), \tag{2.24}$$

$\boldsymbol{Q}_\lambda(p) = \boldsymbol{U}_\lambda(p)$ and $\boldsymbol{Q}_{1-\lambda}(p) = \boldsymbol{U}_{1-\lambda}(p)$ from (2.9). For example, with demeaned data, $p = 0$ and $\boldsymbol{x}_p(s) = 1$, $B_0(r, \lambda)$ is given by

$$B_0(r,\lambda) = \begin{cases} W_0(r) - \frac{r}{\lambda}W_0(\lambda) & \text{for } r \le \lambda \\ [W_0(r) - W_0(\lambda)] - \frac{r-\lambda}{1-\lambda}[W_0(1) - W_0(\lambda)] & \text{for } r > \lambda \end{cases}$$

as can be seen in Busetti and Harvey (2001). Now, using (2.22)-(2.24) and by the CMP, the asymptotic null distribution of the test statistics $\hat{M}_{n,p}^{(i)}(m_n,\lambda)$, i = 1,2,3, are

$$\hat{M}_{n,p}^{(1)}(m_n,\lambda) \Rightarrow \int_0^1 B_p(s,\lambda)^2 ds$$

$$\hat{M}_{n,p}^{(2)}(m_n,\lambda) \Rightarrow \int_0^1 B_p(s,\lambda)^2 ds - \left(\int_0^1 B_p(s,\lambda)ds\right)^2$$

and $\hat{M}_{n,p}^{(3)}(m_n,\lambda) \Rightarrow \sup_{0\le r\le 1} |B_p(r,\lambda) - r\cdot B_p(1,\lambda)|$. For other particular common specifications of the deterministic component, the exact expressions of $B_p(r,\lambda)$ can be found in Busetti and Harvey (2001) as well as the upper tail percentage points for different values of λ of the asymptotic null distribution of the KPSS test statistic using (2.22)-(2.24). For $\lambda = 1$ we usual asymptotic distribution for $n^{-1/2}\hat{S}_{[nr],p}(1)$ is given by

$$B_p(r,1) = B_p(r) = W_0(r) - \int_0^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_1^{-1}(p)\int_0^1 \boldsymbol{x}_p(s)dW_0(s) \tag{2.25}$$

which is a $(p+1)th$-level standard Brownian Bridge process. In the case of no deterministic component in the DGP, $\boldsymbol{x}_{t,p} = 0$, then $B_p(r) = W_0(r)$, and the simulated asymptotic critical values for the one-sided KPSS test can be found in Hobijn et.al. (2004). The following proposition establish the asymptotic distribution of the normalized variance-ratio unit root test in (2.19) with a break in the polynomial trend.

Proposition 1. *Under the usual set of assumptions on the systematic component and the error terms in (2.1)-(2.2), with a structural break at a given location* $0 < \lambda < 1$*, and with* $\sigma_1^2 > 0$*, so that the process* Y_t *has a unit root, then:*

(*i*) $$n^{-1/2}\hat{\eta}_{[nr],p}(\lambda) \Rightarrow \omega_1 \cdot V_p(r,\lambda) \quad 0 \le r \le 1 \tag{2.26}$$

(*ii*) $$\bar{\rho}_{n,p}(\lambda) \Rightarrow \int_0^1 \left(\int_0^r V_p(s,\lambda)ds \right)^2 dr \Big/ \int_0^1 V_p(s,\lambda)^2 ds \tag{2.27}$$

where

$$V_p(r,\lambda) = \begin{cases} W_1(r) - \boldsymbol{x}_p'(r)\boldsymbol{Q}_\lambda^{-1}(p)\int_0^\lambda \boldsymbol{x}_p(s)W_1(s)ds & \text{for } r \le \lambda \\ W_1(r) - \boldsymbol{x}_p'(r)\boldsymbol{Q}_{1-\lambda}^{-1}(p)\int_\lambda^1 \boldsymbol{x}_p(s)W_1(s)ds & \text{for } r > \lambda \end{cases} \tag{2.28}$$

Proof. See Appendix A.

Remark 1. In absence of a structural change, $\lambda = 1$, we have the standard asymptotic distribution of the normalized variance-ratio test by Breitung (2002), given by

$$V_p(r) = V_p(r,1) = W_1(r) - \boldsymbol{x}_p'(r)\boldsymbol{Q}_1^{-1}(p)\int_0^1 \boldsymbol{x}_p(s)W_1(s)ds \tag{2.29}$$

Remark 2. For $p = 0$ (demeaned data), $\boldsymbol{x}_p(s) = 1$, $V_0(r,\lambda)$ is a two-level demeaned Brownian motion process, that is

$$\begin{aligned} V_0(r,\lambda) &= W_1(r) - \frac{1}{\lambda}\int_0^\lambda W_1(s)ds \cdot I(r \le \lambda) - \frac{1}{1-\lambda}\int_\lambda^1 W_1(s)ds \cdot I(r > \lambda) \\ &= W_1(r)(1 - \tfrac{1}{\lambda(1-\lambda)}) \\ &+ \tfrac{1}{\lambda}(W_1(r) - \int_0^\lambda W_1(s)ds \cdot I(r \le \lambda)) + \tfrac{1}{1-\lambda}(W_1(r) - \int_\lambda^1 W_1(s)ds \cdot I(r > \lambda)) \end{aligned} \tag{2.30}$$

with $V_0(r,1) = V_0(r) = W_1(r) - \int_0^1 W_1(s)ds$ for $\lambda = 1$. For $p = 1$ (demeaned and detrended data), $\boldsymbol{x}_p(s) = (1,s)'$, and $V_1(r,\lambda)$ is a two-level demeaned and detrended Brownian motion process with $V_1(r) = W_1(r) - (4-6r)\int_0^1 W_1(s)ds - (12r-6)\int_0^1 sW_1(s)ds$ for $\lambda = 1$.

Remark 3. Under the alternative of stationarity, that is with $\sigma_1^2 = 0$, we have that $\sum_{t=1}^n (\frac{1}{\sqrt{n}}\hat{\eta}_{[nr],p}(\lambda))^2 \to^p \sigma_0^2 = Var[\varepsilon_{0,t}]$, and $(n\sqrt{n})^{-1}\hat{S}_{[nr],p}(\lambda) = O_p(n^{-1})$, so that

$\overline{\rho}_{n,p}(\lambda) = O_p(n^{-1})$ and thus the test procedure is consistent because it converge to zero. The asymptotic distribution of the variance-ratio statistic $\hat{\rho}_{n,p}(\lambda)$ under the stationary alternative is given by

$$\hat{\rho}_{n,p}(\lambda) = \left\{ n^{-1} \sum_{t=1}^{n} \hat{\eta}_{t,p}^2(\lambda) \right\}^{-1} n^{-1} \sum_{t=1}^{n} (\tfrac{1}{\sqrt{n}} \hat{S}_{t,p}(\lambda))^2 \Rightarrow (\omega_0^2 / \sigma_0^2) \int_0^1 B_p(s,\lambda)^2 ds \quad (2.31)$$

with $B_p(s,\lambda)$ as in (2.22)-(2.24), so that it is scaled version of the asymptotic distribution of the KPSS test statistic under stationarity by the factor $\omega_0^2 \cdot \sigma_0^{-2}$ which is one when the error terms $\varepsilon_{0,t}$ are iid.

For the stationarity tests, under the alternative of a linear unit root process $\sigma_1^2 > 0$, we have that $\hat{\eta}_{t,p}(\lambda) = O_p(n^{1/2})$, $n^{-2} \sum_{t=1}^{n} \hat{S}_{t,p}^2(\lambda) = O_p(n^2)$ and $\hat{\omega}_n^2(m_n) = O_p(m_n \cdot n) K_n$, with $K_n = m_n^{-1} \cdot \sum_{j=-(n-1)}^{n-1} w(j, m_n) = O(1)$ and $\hat{\gamma}_n(j) = n \cdot O_p(1)$. Thus, all the three test statistics $\hat{M}_{n,p}^{(i)}(m_n, \lambda)$ diverge to infinity at rates $O_p(n/m_n)$ for i = 1, 2, and $O_p((n/m_n)^{1/2})$ for i = 3. Furthermore, as a generalization of the results in Kwiatkowski et.al. (1992), we have the following asymptotic distribution under the alternative of a unit root for i = 1

$$(m_n / n) \cdot \hat{M}_{n,p}^{(1)}(m_n, \lambda) \to K^{-1} \int_0^1 \left(\int_0^r V_p(s,\lambda) ds \right)^2 dr \Big/ \int_0^1 V_p(r,\lambda)^2 dr \quad (2.32)$$

and similarly for i = 2 and 3, where $K = \int_{-1}^{1} w(s) ds$ and $V_p(r,\lambda)$ is the two-level pth-order corrected Brownian process as in Proposition 1. With the Bartlett kernel, K = 1, so that the asymptotic distribution of the scaled KPSS test statistic is the same as the asymptotic distribution of the variance-ratio unit root test statistic under the unit root hypothesis. Under the sequence of local alternatives to the null of stationarity $\omega_{1,n}^2 = \omega_1^2 \cdot n^{-2}$, with $\gamma_{1,n}(j) = \gamma_1(j) \cdot n^{-2}$ for all j = 0, 1, ..., then

$$n^{-1/2} \sum_{t=1}^{[nr]} \eta_t = O_p(1) \Rightarrow \omega_0 (W_0(r) + c \int_0^r W_1(s) ds) = \omega_0 \cdot W_c(r)$$

where $c = \frac{\omega_1}{\omega_0}$. This implies that $n^{-1/2} \hat{S}_{[nr],p}(\lambda) \Rightarrow \omega_0 \cdot B_{c,p}(r,\lambda)$, where $B_{c,p}(r,\lambda)$ is as $B_p(r,\lambda)$ with $W_0(r)$ replaced by $W_c(r)$. Müller (2005) consider the stationary error component modelled as a strongly autocorrelated mean reverting process in the context of a local-to-unity asymptotic analysis as $\varepsilon_{0,t} = \phi_n \varepsilon_{0,t-1} + z_t$, where $\phi_n = 1 - \gamma \cdot n^{-1}$, with $\gamma \geq 0$, and the zero-mean sequence z_t is covariance-stationary with finite autocovariances

$\gamma_z(j)<\infty$, so that $\omega_z^2=\sum_{j=-\infty}^{\infty}\gamma_z(j)<\infty$ and $n^{-1/2}\sum_{j=1}^{[nr]}z_j \Rightarrow \omega_z W_0(r)$. For $\gamma > 0$, $|\phi_n| <$ 1 and the initial value $\varepsilon_{0,0}=\sum_{s=0}^{\infty}\phi_n^s z_{-s}$ lead jointly to the stationarity of the series $\varepsilon_{0,t}$ and will have a considerably effect on the asymptotic distributions considered below (see Elliot (1999), Müller and Elliot (2003) and Müller (2005)). In particular, we then have that the weak limit of $n^{-1/2}\eta_{[nr]}$ is a mixture of two continuous time processes $n^{-1/2}\eta_{[nr]} \Rightarrow \omega_z(\tilde{M}(r)+\frac{\omega_1}{\omega_z}W_1(r))$, where $\tilde{M}(r)=M(r)+\zeta(2\gamma)^{-1/2}$, ζ is a standard normal variable independent of $W_0(r)$, $M(r)=W_0(r)$ for $\gamma = 0$, and $M(r)=\zeta(e^{-\gamma r}-1)(2\gamma)^{-1/2}+J_{-\gamma}(r)$ for $\gamma > 0$, with $J_{-\gamma}(r)$ an Ornstein-Uhlenbeck process given by $J_{-\gamma}(r)=\int_0^r e^{-\gamma(r-s)}dW_0(s)=W_0(r)-\gamma\int_0^r e^{-\gamma(r-s)}W_0(s)ds$. For $\gamma>0$, the component $\tilde{M}(r)$ of the limit process is a stationary continuous time process. Then, with $n^{-2}\hat{\omega}_n^2(m_n)=o_p(1)$ for $m_n = o_p(n)$, $\frac{1}{n}\sum_{t=1,n}(\frac{1}{n\sqrt{n}}\hat{S}_{t,p}(\lambda))^2 \Rightarrow \omega_z^2\cdot\int_0^1\left(\int_0^r M_p(s,\lambda)ds\right)^2 dr$ for any $\gamma = n\cdot(1-\phi_n) \geq 0$ and $\sigma_1^2=0$, with $M_p(s,\lambda)$ as $V_p(s,\lambda)$ where $W_1(s)$ is replaced by $M(s)$, then the stationarity test based on $\hat{M}_{n,p}^{(1)}(m_n,\lambda)$ reject the null hypothesis of stationarity with probability one under local-to-unity asymptotics (Müller (2005), Proposition 1, p.201).

There are many other studies about the behavior and stochastic properties of these test statistics under different assumptions about the nature of the stochastic components in (2.1)-(2.2), but in what follows we will be concerned with the effects of a misspecification in the systematic component in (2.1). To analyze the possible distortions of these specification errors, we will use a kind of local-to-the correct specification approach by letting the implied parameters depend on the sample size through an additional parameter that allows us to control their asymptotic effects.

As in linear regression analysis, there are many possible situations where a misspecification in the systematic component will affect the behavior and properties of a particular test statistic. We are going to consider two groups of this type of error: wrong identification of the order in the polynomial trend, omission of a structural break in the polynomial trend or wrong location of the break-point in the sample, and second: omission of outliers or influential observations, that is the main topic of this paper. The other cases are considered for comparison purposes.

Thus, for example, in the general case of a structural break at a known given position λ, we could consider the situation where the true order of the polynomial trend function p_0 differs from the specified one, p. When $p_0>p$, [1] then we have $\boldsymbol{x}'_{t,p_0}\boldsymbol{\beta}_{p_0,t}(\lambda)=\boldsymbol{x}'_{t,p}\boldsymbol{\beta}_{p,t}(\lambda)+\boldsymbol{x}'_{t,p_0-p}\boldsymbol{\beta}_{p_0-p,t}(\lambda)$, so that the error term in (2.10) is now $\xi_t=\eta_t+\boldsymbol{x}'_{t,p_0-p}\boldsymbol{\beta}_{p_0-p,t}(\lambda)$, and thus the t-th scaled OLS residual will be given by

$$\begin{aligned} n^{-1/2}\cdot\hat{\xi}_{t,p}(\lambda) &= n^{-1/2}\cdot\hat{\eta}_{t,p}(\lambda) \\ &+(1-h_t(\lambda))n^{-1}\left(\boldsymbol{x}'_{p_0-p}(\tfrac{t}{n})-\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}^{-1}_{[n\lambda]}(p)\tfrac{1}{n}\sum_{j=1}^{[n\lambda]}\boldsymbol{x}_p(\tfrac{j}{n})\boldsymbol{x}'_{p_0-p}(\tfrac{j}{n})\right)\boldsymbol{\beta}_{p_0-p,n} \\ &+h_t(\lambda)n^{-1}\left(\boldsymbol{x}'_{p_0-p}(\tfrac{t}{n})-\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p)\tfrac{1}{n}\sum_{j=[n\lambda]+1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})\boldsymbol{x}'_{p_0-p}(\tfrac{j}{n})\right)\boldsymbol{\theta}_{p_0-p,n} \end{aligned} \quad (2.33)$$

with $n^{-1/2}\cdot\hat{\eta}_{t,p}(\lambda)$ as in (A.14) in Appendix A with v = 1/2, $\boldsymbol{\beta}_{p_0-p,n}=n^{1/2}\cdot\boldsymbol{D}^{-1}_{p_0-p,n}\boldsymbol{\beta}_{p_0-p}$, and $\boldsymbol{\theta}_{p_0-p,n}=n^{1/2}\cdot\boldsymbol{D}^{-1}_{p_0-p,n}\boldsymbol{\theta}_{p_0-p}$, where we have used $\boldsymbol{D}_{p_0,n}=(\boldsymbol{D}_{p,n}:\boldsymbol{D}_{p_0-p,n})$. Each term in $\boldsymbol{\beta}_{p_0-p,n}$ and $\boldsymbol{\theta}_{p_0-p,n}$ is of order $O(n^{j+1/2})$ for $j=p+1, ..., p_0$, so that with finite values of the coefficients β_j and θ_j, $j=p+1, ..., p_0$, the scaled partial sum of the OLS residuals will diverge at a rate $O_p(n^{p_0+1/2})$ while the sample autocovariances based on these residuals will diverge at a rate $O_p(n^{2p_0})$. Thus, the KPSS test statistic will be $O_p(n/m_n)$, and will reject too often the null of stationarity. Unless β_j and θ_j where of order $O(n^{-(j+1/2+\alpha)})$ for each $j=p+1, ..., p_0$, with $0<\alpha<1/2$, so that all the coefficients $\beta_{j,n}$ and $\theta_{j,n}$ are asymptotically negligible, the omitted terms in the specification of the systematic component will cause divergence of the test statistics for testing the null of stationarity because $n^{-1/2}\hat{S}_{[nr],p}(\lambda)=O_p(n^{-\alpha})$. With α = 0, so that $\beta_{j,n}$ and $\theta_{j,n}$ are all finite constants and independent of the sample size and the order j, the asymptotic null distribution of these test statistics will change due to the additional limits of the above last two terms, while for $-1/2\le\alpha<0$, the scaled partial sum of OLS residuals will diverge to infinite at the given rate. This latter case will cause an increase in the empirical size of the stationarity tests. The above conditions on the size of the coefficients for the omitted terms could not be appropriate in this context. However, in other situations this formulation has been used in order to eliminate nuisance parameters or to preserve some asymptotic results even under a possible misspecification. This approach results in a kind of local sensitivity analysis, and it will be used extensively in

[1] The case $p_0<p$ will not cause, in general, significative effects on the properties of the test statistics as it corresponds to the inclusion of irrelevant regressors with zero coefficients (Hadri and Rao (2009)).

the rest of the paper.

Lee et. al. (1997) examine the effect of the omission of a structural break on KPSS test for stationarity, whereas Carrion-i-Silvestre (2003) shows that the KPSS test with $p = 0$ and 1 diverges under a misspecification of the break-point. We will extend these results to the variance-ratio unit root test, allowing the shift parameters could vary with the sample size which covers the case of no break actually occurring.

In the first case, when the structural break is ignored, the residuals are computed as in equation (2.13) with the error terms of the auxiliary regression given now by $\xi_t = \eta_t + \boldsymbol{x}'_{t,p} h_t(\lambda)\boldsymbol{\alpha}_p$. Thus, the sequence of OLS residuals are of the form

$$\hat{\xi}_{t,p} = \hat{\eta}_{t,p} + \boldsymbol{z}'_{p,n}(t,\lambda)\boldsymbol{\alpha}_{p,n} \tag{2.34}$$

with $\hat{\eta}_{t,p} = \eta_t - n^{-v}\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)n^{v-1}\sum_{j=1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})\eta_j$, and

$$\boldsymbol{z}'_{p,n}(t,\lambda) = \boldsymbol{x}'_p(\tfrac{t}{n})(h_t(\lambda) - \bar{\boldsymbol{Q}}_n^{-1}(p)\bar{\boldsymbol{U}}_{n-[n\lambda]}(p)), \tag{2.35}$$

and $\boldsymbol{\alpha}_{p,n} = \boldsymbol{D}_{p,n}^{-1}\boldsymbol{\alpha}_p = \boldsymbol{O}(n^p)$ if $\boldsymbol{\alpha}_p = \boldsymbol{O}(1)$. For stationarity tests, $v = 1/2$, $\eta_t = \varepsilon_{0,t}$, and

$$n^{-1/2}\hat{S}_{[nr],p} = O_p(1) + n^{-1}\sum_{t=1}^{[nr]}\boldsymbol{z}_{p,n}(t,\lambda)\,n^{1/2}\cdot\boldsymbol{\alpha}_{p,n} = O_p(1) + \boldsymbol{\alpha}_{p,n}(v) = O_p(n^{p+v}) \tag{2.36}$$

that will diverge unless $\alpha_i = O(n^{-(i+v+\alpha)})$ for $i = 0, 1, ..., p$ for $0 < \alpha \le 1/2$, with $\alpha_{i,n} = O(n^{-(v+\alpha)})$, and $\alpha_{i,n}(v) = O(n^{-\alpha})$. This is the assumption made in Busetti and Harvey (2001) when computing the KPSS test under a structural change with a unknown break point location. In this case the asymptotic distribution of the KPSS test is the same as without the break. With $\alpha = 0$, $\alpha_{i,n}(v) = O(1)$ and $n^{-1/2}\hat{S}_{[nr],p} = O_p(1)$ but with a different limit distribution that depends on the unknown values of $\alpha_{i,n}(v)$. For relatively large values of the shift parameters, with $-1/2 \le \alpha < 0$, then $n^{-1}\hat{S}^2_{[nr],p} = O_p(n^{-2\alpha})$ but $\hat{\omega}^2_n(m_n)$ is still a consistent estimator of the long-run variance. Also, from (2.34) with fixed shift parameters, $\alpha_i = O(1)$, and under stationarity, the jth order sample autocovariance is $\hat{\gamma}_n(j) = n^{-1}\sum_{t=j+1}^{n}\varepsilon_{0,t}\varepsilon_{0,t-j} + O_p(n^{2p})$, which implies that the kernel estimator of the long-run variance of $\varepsilon_{0,t}$ in (2.18) is of order $O_p(m_n \cdot n^{2p})$. Then, with finite shift parameters, the KPSS test statistic will diverge at the same rate as under the unit root alternative, $O_p(n/m_n)$, except in the case $p = 0$ where it is $O_p(n)$.

The same argument can be applied in the case of a misspecification in the break point location. Thus, with $\lambda_0 \in (0,1)$ the true location of the break, the error term in (2.10) is now given by $\xi_t = \eta_t + \boldsymbol{x}'_{t,p}[h_t(\lambda_0) - h_t(\lambda)]\boldsymbol{\alpha}_p \ = \eta_t + \boldsymbol{x}'_{t,p} d_t(\lambda_0)\boldsymbol{\alpha}_p$, where $d_t(\lambda_0)=1$ for $t = [n\lambda_0]+1, ..., [n\lambda]$ if $\lambda_0 < \lambda$, $d_t(\lambda_0)=-1$ for $t = [n\lambda]+1, ..., [n\lambda_0]$ if $\lambda_0 > \lambda$, and zero otherwise. The OLS residuals are given now as in (2.34) with $\hat{\eta}_{t,p}$ replaced by $\hat{\eta}_{t,p}(\lambda)$ in (2.14) and $\boldsymbol{z}'_{p,n}(t,\lambda)$ replaced by

$$\begin{aligned} \boldsymbol{z}'_{p,n}(t,\lambda_0) = \boldsymbol{x}'_p(\tfrac{t}{n})d_t(\lambda_0) - (1-h_t(\lambda)\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}^{-1}_{[n\lambda]}(p)\bar{\boldsymbol{Q}}_{[n\lambda]}(p,\lambda_0) \\ - h_t(\lambda)\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p)\bar{\boldsymbol{U}}_{n-[n\lambda]}(p,\lambda_0) \end{aligned} \tag{2.37}$$

where

$$\bar{\boldsymbol{Q}}_{[n\lambda]}(p,\lambda_0) = n^{-1}\sum_{j=[n\lambda_0]+1}^{[n\lambda]} \boldsymbol{x}_p(\tfrac{j}{n})\boldsymbol{x}'_p(\tfrac{j}{n}) \text{ if } \lambda_0 < \lambda,$$

and

$$\bar{\boldsymbol{U}}_{n-[n\lambda]}(p,\lambda_0) = -n^{-1}\sum_{j=[n\lambda]+1}^{[n\lambda_0]} \boldsymbol{x}_p(\tfrac{j}{n})\boldsymbol{x}'_p(\tfrac{j}{n}) \text{ when } \lambda_0 > \lambda.$$

With fixed and finite shift parameters and $p > 0$, it could be expected the same size distortion of the KPSS test as in Carrion-i-Silvestre, with a symmetric behavior around $\lambda_0 - \lambda = 0$. In order to complete this analysis and to compare with these results and the established in Proposition 1, the following Proposition 2 gives the basic sensitivity results for the variance-ratio unit root test.

Proposition 2. *Under misspecification of the model in (2.10), both for omission of the structural break or misspecification of the break-point location, and under the unit root hypothesis, with* $\alpha_{j,n} = \alpha_j \cdot n^j$ *and* $\alpha_{j,n}(v) = n^{j+v}\cdot\alpha_j$, $j = 0, 1, ..., p$, $v = -1/2$, *then we have:*

(*i*) *With fixed shift parameters the variance ratio test statistic will be O(1) with a non-random limit, except for p=0 where the misspecification have no effect, and*

(*ii*) *When the shift parameters are of orders* $\alpha_j = c_j \cdot O(n^{-(j+v+\alpha)})$, *for* $0 < \alpha \leq 1/2$, *there is no effect of the misspecification, while for* $\alpha = 0$, *the test statistic is* $O_p(1)$ *but with a different limit distribution. For* $-1/2 \leq \alpha < 0$, *the test statistic will be O(1) for all p.*

Proof. For part (*i*), with $\alpha_{0,n}(v) = n^{-1/2}\cdot\alpha_0$ and $\alpha_{j,n}(v) = n^{j-1/2}\cdot\alpha_j$, $j = 1, ..., p$, finite values of the α_j, and $n^{-1/2}\hat{\xi}_{t,p}(\lambda) = n^{-1/2}\hat{\eta}_{t,p}(\lambda) + \boldsymbol{z}'_{p,n}(t,\lambda)O(n^{p-1/2})$, the result follows easily for p=0 (demeaned observations). For $p \geq 1$, both the numerator and the denominator diverges at the same rate $O_p(n^{2p+1})$, so that in the limit behaves as it were a finite constant value. For part (*ii*), under the given assumption on the size of the shift parameters we have that $\alpha_{j,n}=O(n^{1/2-\alpha})$ and $\alpha_{j,n}(v)=O(n^{-\alpha})$ for all j=0,1,...,p. For values $0<\alpha\leq 1/2$, $\frac{1}{\sqrt{n}}\hat{\xi}_{t,p}(\lambda) = \frac{1}{\sqrt{n}}\hat{\eta}_{t,p}(\lambda) + o(1)$, while for α=0 $\frac{1}{\sqrt{n}}\hat{\xi}_{t,p}(\lambda) = \frac{1}{\sqrt{n}}\hat{\eta}_{t,p}(\lambda) + \boldsymbol{z}'_{p,n}(t,\lambda)\boldsymbol{c}_p$, so that $\frac{1}{\sqrt{n}}\hat{\xi}_{[nr],p}(\lambda) \Rightarrow \omega_1 V_p(r,\lambda) + \boldsymbol{z}'_p(r,\lambda)\boldsymbol{c}_p$, with the corresponding limit of (2.35) or (2.37). Otherwise, for $\alpha < 0$, it is of application the same argument as in part (*i*).

■

The rest of this section is devoted to the case of existence of outliers in the sample and their effects on the stationarity and unit root tests. The effects of additive outliers on parametric tests for unit roots, as the Dickey-Fuller test, has been well established by Franses and Haldrup (1994) and Shin et.al. (1996). Thus, the limiting distribution of the OLS estimator of the unit root parameter in the AR(1) model is affected by additive outliers (AO) and may produce spurious stationarity, but is unaffected by innovational outliers. The performance of stationarity tests is studied in Darné (2004) and Otero and Smith (2005) for the KPSS test, and Afonso-Rodríguez (2009) for the KPSS, V/S and KS tests. The numerical evidence found in these papers indicate that the stationarity tests are quite robust under stationarity in presence of outliers, isolated or in patches, except in the case of very persistent outliers. Even for moderately large sample sizes, the mean reversion effect induced by the outliers may result in low power of these tests. However, in spite of the large numerical evidence on this effects, there is no a complete analytical work that allows to clarify these findings. In Afonso-Rodríguez (2010) there is an attempt to formulate this problem through different possible representations of the outlier contaminating process. Thus, the error term in (2.10) is given now by $\xi_t = \eta_t + Z_t(\boldsymbol{\theta})$, with $Z_t(\boldsymbol{\theta})$ a proper deterministic or stochastic function that contains information about the magnitudes, locations, number and persistence properties of the outliers present in the sample. Following Rodrigues and Rubia (2010) and Otero and Smith (2005), we consider a Bernoulli-type stochastic jump process defined by

$$Z_t(\boldsymbol{\theta}) = B_t(\boldsymbol{\theta})(\lambda_0 + \lambda_1 v_t), \quad B_t(\boldsymbol{\theta}) = (1-\phi L)^{-1}\tilde{B}_t(\boldsymbol{\pi}) \tag{2.38}$$

where λ_j are real finite parameters, $j = 0, 1$, $v_t \sim iid(0,\sigma_v^2)$, $|\phi| < 1$ and $\tilde{B}_t(\boldsymbol{\pi})$ is an iid Bernoulli-type sequence with support (1, −1, 0) and probabilities $\boldsymbol{\pi} = (\pi_1,\pi_2,1-(\pi_1+\pi_2))$. Furthermore, the processes v_t and $\tilde{B}_t(\boldsymbol{\pi})$ are mutually independent. Then, the OLS residuals from (2.13) (that is, without structural break), are given by

$$\hat{\xi}_{t,p} = \hat{\eta}_{t,p} + \mu_1(\boldsymbol{\gamma})m_{t,n} + \tilde{Z}_t(\boldsymbol{\gamma}) - \boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)\tfrac{1}{n}\sum_{j=1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})\tilde{Z}_j(\boldsymbol{\gamma}) \tag{2.39}$$

with $\hat{\eta}_{t,p}$ as in (2.34), $m_{t,n} = 1 - \boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)\tfrac{1}{n}\sum_{j=1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})$, and $\tilde{Z}_t(\boldsymbol{\gamma}) = Z_t(\boldsymbol{\gamma}) - \mu_1(\boldsymbol{\gamma})$, where $\mu_1(\gamma) = E[Z_t(\gamma)] = \lambda_0 \cdot E[B_t(\pi)]$, $E[B_t(\pi)] = (1-\phi)^{-1} \cdot E[\tilde{B}_t(\pi)] = (1-\phi)^{-1} \cdot (\pi_1 - \pi_2)$. Under this conditions, $\tilde{Z}_t(\gamma)$ is a zero-mean stationary sequence that satisfy the FCLT for martingale differences, that is $n^{-1/2}\sum_{j=1}^{[nr]}\tilde{Z}_t(\gamma) \Rightarrow \omega(\gamma) \cdot W_\gamma(r)$, with $\omega^2(\gamma)$ the long-run variance defined below. Thus, we can formulate the following result concerning the asymptotic distribution of the scaled residuals and partial sum of residuals needed to build the variance-ratio and stationarity tests under this general type of outlier contamination.

Proposition 3. *With the assumption* $\pi_1 = \pi_2 = \pi$, *so that* $\mu_1(\gamma) = 0$, *and the usual regularity conditions on the error terms in (2.1)-(2.2), we have that:*

(*i*) *Under the unit root hypothesis, then:*

$$n^{-1/2}\hat{\xi}_{[nr],p} = n^{-1/2}\hat{\eta}_{[nr],p} + n^{-1/2}\tilde{Z}_{[nr]}(\gamma) + O_p(n^{-1}) \Rightarrow \omega_1 V_p(r) \tag{2.40}$$

(*ii*) *Under the stationarity hypothesis, then:*

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{\xi}_{t,p} \Rightarrow \omega_0[B_p(r) + \tfrac{\omega(\gamma)}{\omega_0}B_{p,\gamma}(r)] \tag{2.41}$$

with $B_{p,\gamma}(r)$ *as* $B_p(r)$ *with* $W_0(r)$ *replaced by* $W_\gamma(r)$ *and* $\omega^2(\gamma) = \frac{2\pi}{1-\phi^2}(\lambda_1^2\sigma_v^2 + \lambda_0^2(\frac{1+\phi}{1-\phi}))$.

Proof. The proof of (*ii*) follows from Afonso-Rodríguez (2010). For (*i*), given the stationarity of the sequence $\tilde{Z}_t(\gamma)$, then $\tilde{Z}_t(\gamma)$, and $n^{-1/2}\sum_{j=1}^{n}\boldsymbol{x}_p(\frac{j}{n})\tilde{Z}_j(\gamma)$ are both $O_p(1)$, and thus $\hat{\xi}_{t,p} = \hat{\eta}_{t,p} + \tilde{Z}_t(\gamma) + O_p(n^{-1/2})$ in (2.39). The result then follows under scaling.

■

From (2.41), and under the assumption that the outlier generating process has no effect on the level of the series, it is seen that the limiting null distribution of the stationarity tests can be seriously distorted when $\omega^2(\gamma)$ is large compared with ω_0^2. This long-run variance grows with the degree of persistence of the outlier, for large positive values of ϕ, and also for large values of λ_0 and λ_1. This particular choice of the outlier generating process does not cause any distortion in the variance-ratio unit root test, unless we consider a local-to-unity framework of analysis through the assumption that $\omega^2(\gamma) = c^2(\gamma) \cdot n$, with $c(\gamma)$ a finite constant, in which case in (*i*) we have $n^{-1/2}\hat{\xi}_{t,p} \Rightarrow \omega_1[V_p(r) + \frac{c(\gamma)}{\omega_1}W_\gamma(r)]$.

3. Augmented basic DGP by an outlier intervention model

Let us now assume that the true DGP is given by the following augmented version of the basic DGP by an additional deterministic term, $\varphi_t(\boldsymbol{\theta}_0)$, that takes into account the effect of an additive-type outlier (AO), that is,

$$Y_t = d_t(p,\lambda) + \eta_t + \delta\varphi_t(\boldsymbol{\theta}_0)\,,\ \boldsymbol{\theta}_0 = (\tau_0, \phi_0)' \tag{3.1}$$

where $0 < \tau_0 \leq 1$ is the relative location of the AO in the sample, and $|\phi_0| < 1$ is a parameter that allows to capture the possible persistent effect of the perturbation. In the case $\phi_0 = 0$, with $\boldsymbol{\theta}_0 = (\tau_0, 0)'$, we define $\varphi_t(\boldsymbol{\theta}_0)$ simply as a step function

$$\varphi_t(\boldsymbol{\theta}_0) = I_t(\tau_0) = I(t = k_0) \tag{3.2}$$

while in the case $|\phi_0| < 1$, with $\phi_0 \neq 0$, we have

$$\varphi_t(\boldsymbol{\theta}_0) = I_t(\tau_0) + \phi_0^{t-k_0} h_t(\tau_0) = \phi_0^{t-k_0} H_t(\tau_0) \tag{3.3}$$

with, $H_t(\tau_0) = I(t \geq k_0) = I_t(\tau_0) + h_t(\tau_0)$ and $k_0 = [n\tau_0]$. In (3.2) we consider the possibility of a single isolated pure AO with true location $k_0 = [n\tau_0]$, which can differ from the specified in the auxiliary regression of the tests as can be seen below. In (3.3) we consider the case of a single isolated but persistent AO, with initial impact δ at $k_0 = [n\tau_0]$ and subsequent effect that decrease at a geometric rate of $\delta\cdot\phi_0^{t-k_0}$ for all $t = k_0+1, ..., n$. The value of ϕ_0 determines the extent of the duration of the effect, with mean and median life of the impact given by $|\phi_0|/(1-|\phi_0|)$ and $log(0.5)/log(|\phi_0|)-1$, respectively. To take into account the outlier effect in the computation of the stationarity and unit root tests we consider the following auxiliary regression

$$Y_t = d_t(p,\lambda) + \delta\cdot I_t(\tau) + \xi_t \tag{3.4}$$

where, the error term ξ_t is given by $\xi_t = \eta_t + \delta\cdot[\varphi_t(\boldsymbol{\theta}_0) - I_t(\tau)] = \eta_{0,t} - \delta I_t(\tau)$, with

$$\eta_{0,t} = \eta_t + \delta\cdot\varphi_t(\boldsymbol{\theta}_0) \tag{3.5}$$

In the case of no structural breaks ($\lambda = 1$), the OLS residuals from (3.4) are given by

$$\begin{aligned}\hat{\xi}_{t,p}(\tau) &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n}(\tau) - I_t(\tau)\hat{\delta}_n(\tau)\\ &= Y_t - \boldsymbol{x}'_{t,p}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{Q}_n^{-1}(p)\boldsymbol{x}'_{k,p}\hat{\delta}_n(\tau)) - I_t(\tau)\hat{\delta}_n(\tau)\\ &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n} - \hat{\delta}_n(\tau)m_{t,k}(\tau)\end{aligned}$$

where $k = [n\tau]$, $m_{t,k}(\tau) = I_t(\tau) - \frac{1}{n}\boldsymbol{x}'_p(\frac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)\boldsymbol{x}_p(\frac{k}{n})$, $\hat{\boldsymbol{\beta}}_{p,n} = \boldsymbol{\beta}_p + \boldsymbol{Q}_n^{-1}(p)\sum_{j=1}^{n}\boldsymbol{x}_{j,p}\eta_{0,j}$ the OLS estimation of vector parameter without the AO dummy as in section 2, and

$\hat{\delta}_n(\tau) = m_{k,k}^{-1}(\tau)(Y_k - \boldsymbol{x}'_{k,p}\hat{\boldsymbol{\beta}}_{p,n})$. Under correct location of the pure AO, $\boldsymbol{\theta}_0 = \boldsymbol{\theta} = (\tau, 0)'$, and using (3.5) with (3.2) then

$$\hat{\xi}_{t,p}(\tau) = \eta_{0,t} - \boldsymbol{x}'_{t,p}\boldsymbol{Q}_n^{-1}(p)\tfrac{1}{n}\sum_{j=1}^{n}\boldsymbol{x}_{j,p}\eta_{0,j} - \hat{\delta}_n(\tau)m_{t,k}(\tau) = \hat{\eta}_{t,p} - [\hat{\delta}_n(\tau) - \delta]\cdot m_{t,k}(\tau)$$

where $\hat{\eta}_{t,p} = \eta_t - n^{-v}\cdot\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)n^{v-1}\sum_{j=1}^{n}\boldsymbol{x}_p(\tfrac{j}{n})\eta_j$ and v determines the scaling factor needed to obtain a finite limit distribution of the test statistics, with v = 1/2 under stationarity and $v = -1/2$ for the unit root test. Then, for $v = -1/2$,

$$n^{-1/2}\cdot\hat{\xi}_{t,p}(\tau) = n^{-1/2}\hat{\eta}_{t,p} - M_{t,k}(\tau)n^{-1/2}\hat{\eta}_{k,p} \tag{3.6}$$

with $M_{t,k}(\tau) = \frac{m_{t,k}(\tau)}{m_{k,k}(\tau)} = I(\tau) + O(n^{-1})$, while for v = 1/2

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\tau) = n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p} - (\tfrac{1}{\sqrt{n}}\hat{\eta}_{k,p})\sum_{t=1}^{[nr]}M_{t,k}(\tau) = \tfrac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p} + O_p(n^{-1/2}) \tag{3.7}$$

where $n^{-1/2}(\hat{\delta}_n(\tau) - \delta) = n^{-1/2}\hat{\eta}_{k,p} = O_p(n^{-1/2})$. From (3.6) and (3.7), it is expected that the inclusion of the dummy variable will not have any significant effect on the asymptotic distribution of the corresponding test statistics. However, if the true location of the outlier is at $k_0 = [n\tau_0]$, with $\tau \neq \tau_0$, then

$$\hat{\xi}_{t,p}(\tau) = \hat{\eta}_{t,p} - M_{t,k}(\tau)\hat{\eta}_{k,p} + \delta(\varphi_{t,p}(\tau_0) - M_{t,k}(\tau)\varphi_{k,p}(\tau_0)), \tag{3.8}$$

with

$$\varphi_{t,p}(\tau_0) = I_t(\tau_0) - \tfrac{1}{n}\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)\boldsymbol{x}'_p(\tfrac{k_0}{n}) \tag{3.9}$$

and

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\tau) = \tfrac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p} + \tfrac{\delta}{\sqrt{n}}\left(\sum_{t=1}^{[nr]}\varphi_{t,p}(\tau_0) - \varphi_{k,p}(\tau_0)\sum_{t=1}^{[nr]}M_{t,k}(\tau)\right) + O_p(\tfrac{1}{\sqrt{n}}) \tag{3.10}$$

In (3.10), as in (3.8) under scaling by $n^{-1/2}$, it is seen that the behavior of the resulting test statistics can be affected mainly due to the size of the scaled outlier magnitude and, depending on this, the possible differences in location and the persistence nature of the outlier. To make the analysis as general as possible we consider the main results of this section in the case of the OLS estimation with a break in the polynomial trend in (3.1),

$$Y_t = \boldsymbol{x}'_{t,p}\boldsymbol{\beta}_{p,t}(\lambda) + \delta\cdot I_t(\tau) + \xi_t \tag{3.11}$$

with residuals denoted by $\hat{\xi}_{t,p}(\lambda, \tau)$, and a possible misspecification in the location and type of the outlier, as in (3.2)-(3.3). When $\phi_0 \neq 0$, with (3.3), we will encounter terms of the form

$$\sum_{t=1}^{[nr]} \boldsymbol{x}_p(\tfrac{t}{n})\varphi_t(\boldsymbol{\theta}_0) = \sum_{t=k_0}^{[nr]} \phi_0^{t-k_0} \boldsymbol{x}_p(\tfrac{t}{n}) = \sum_{t=k_0}^{[nr]} \boldsymbol{x}_p(\tfrac{t}{n}) + \sum_{t=k_0+1}^{[nr]} (\phi_0^{t-k_0} - 1)\boldsymbol{x}_p(\tfrac{t}{n})$$

which must have a finite limit depending on the value of the persistence parameter, ϕ. Thus, to obtain a closed expression for the limit results that involve the persistence parameter ϕ_0 in the case where the true contaminating process is not a pure additive outlier, we introduce the following assumption.

Assumption 1. *In the case where the true outlier process is not a pure additive outlier (AO) but instead we allow for a certain degree of persistence, then the persistence parameter* ϕ_0 *is given by* $\phi_0 = exp(\gamma_0/n)$ *with* $\gamma_0 < 0$ *so that* $0 < \phi_0 < 1$.

Remark 4. Under this assumption, and the limit result in (2.8), we have that

$$\begin{aligned} \boldsymbol{G}_{p,n}(r,\boldsymbol{\theta}_0) &= \frac{1}{n}\sum_{t=1}^{[nr]} \boldsymbol{x}_p(\tfrac{t}{n})\varphi_t(\boldsymbol{\theta}_0) = \frac{1}{n}\sum_{t=1}^{[nr]} \phi_0^{t-k_0}\boldsymbol{x}_p(\tfrac{t}{n})H_t(\tau_0) = \frac{1}{n}\sum_{t=k_0}^{[nr]} \phi_0^{n(t/n-\tau_0)}\boldsymbol{x}_p(\tfrac{t}{n}) \\ &= \sum_{t=k_0}^{[nr]} e^{(t/n-\tau_0)\gamma_0}\int_{(t-1)/n}^{t/n} d\boldsymbol{z}_{p,n}(s) = \sum_{t=[n\tau_0]}^{[nr]} \int_{(t-1)/n}^{t/n} e^{(s-\tau_0)\gamma_0} d\boldsymbol{z}_{p,n}(s) \\ &= \int_{\tau_0}^{r} e^{(s-\tau_0)\gamma_0} d\boldsymbol{z}_{p,n}(s) \to \int_{\tau_0}^{r} e^{(s-\tau_0)\gamma_0} d\boldsymbol{z}_p(s) = \boldsymbol{G}_{p,\gamma_0}(r,\boldsymbol{\theta}_0) \end{aligned} \tag{3.12}$$

Similarly, for the normalized partial sum process of the error term $\varepsilon_{i,t}$ (i = 0, 1) and using (2.3), we have

$$\begin{aligned} J_{n,i}(r,\boldsymbol{\theta}_0) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} \varepsilon_{i,t}\varphi_t(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} \phi_0^{t-k_0}\varepsilon_{i,t}H_t(\tau_0) = \frac{1}{\sqrt{n}}\sum_{t=k_0}^{[nr]} \phi_0^{n(t/n-\tau_0)}\varepsilon_{i,t} \\ &= \sum_{t=k_0}^{[nr]} e^{(t/n-\tau_0)\gamma_0}\int_{(t-1)/n}^{t/n} dW_{n,i}(s) = \sum_{t=[n\tau_0]}^{[nr]} \int_{(t-1)/n}^{t/n} e^{(s-\tau_0)\gamma_0} dW_{n,i}(s) \\ &= \int_{\tau_0}^{r} e^{(s-\tau_0)\gamma_0} dW_{n,i}(s) \to \omega_i \cdot \int_{\tau_0}^{r} e^{(s-\tau_0)\gamma_0} dW_i(s) = \omega_i J_{\gamma_0,i}(r,\boldsymbol{\theta}_0) \end{aligned} \tag{3.13}$$

as $n\to\infty$ for all $r \geq \tau_0$. In the case $p = 0$, with $\boldsymbol{x}_p(s) = 1$ and $\boldsymbol{z}_p(s) = s$, then (3.12) simplifies to $G_{0,\gamma_0}(r,\boldsymbol{\theta}_0) = \int_{\tau_0}^{r} e^{(s-\tau_0)\gamma_0} ds = \gamma_0^{-1}[e^{(r-\tau_0)\gamma_0} - 1]$. The above finite sample function and integrals only exist for $r \geq \tau_0$. The limit process $J_{\gamma_0,i}(r,\boldsymbol{\theta}_0)$ in (3.13) is similar to the Ornstein-Uhlenbeck process as the stochastic limit of a first-order autoregression with autoregressive parameter depending on the sample size.

Also, from expressions (3.8) and (3.10) that are the leading terms needed to build the variance-ratio and stationarity tests, the outlier magnitude appears scaled by $n^{-1/2}$, $\delta_n = n^{-1/2}\cdot\delta$. This can partially explain the results found in previous studies where in neither case δ_n exceeds 0.7 or, at most 1, which correspond with very small magnitudes of δ. Then, in order to assign a more realistic sense to what can be considered an outlier in this context we introduce the assumption that for a value of $\alpha \in [-1/2, 1/2]$, the magnitude of the outlier is of order $\delta = c\cdot O(n^{1/2-\alpha})$, where the constant c must be small. The case $\alpha = 1/2$ correspond to a fixed and finite value of the outlier magnitude, $\delta = c$, while for $0 < \alpha < 1/2$, even for small values of c, the outlier magnitude could be very high. The case $-1/2 \leq \alpha < 0$ corresponds to a extremely large outlier, that rarely are encountered in practice but that we consider from the theoretical view of the analysis. With this, we denote by $\delta_n(\alpha) = n^{-1/2}\cdot\delta = c\cdot O(n^{-\alpha})$ the scaled outlier magnitude, as a function of the constant c, the sample size n and α. The corresponding limit value of $\delta_n(\alpha)$ will be c for $\alpha = 0$, $o(1)$ for $0<\alpha<1/2$, and c times a divergent term for $-1/2<\alpha<0$, as $n\to\infty$. The main conclusions founded are summarized in the following Proposition 4.

Proposition 4. *Under the outlier generating process in (3.2) and (3.3), and the usual regularity conditions on the error terms, we have that:*

(*i*) *Under the unit root hypothesis, then* $n^{-1/2}\hat{\xi}_{[nr],p}(\lambda,\tau) \Rightarrow \omega_1 V_p(r,\lambda)$, *irrespective of the persistent nature and the true location of the outlier, for* $0\leq\alpha\leq 1/2$. *For extremely large outliers,* $-1/2\leq\alpha<0$, *then* $n^{-1/2}\hat{\xi}_{[nr],p}(\lambda,\tau) \to^p \delta(\alpha)/\omega_1$.

(*ii*) *Under the stationarity hypothesis, then* $n^{-1/2}\hat{S}_{[nr],p}(\lambda,\tau) \Rightarrow \omega_0 B_p(r,\lambda)$, *when* $\phi = 0$, *and* $\tau=\tau_0$, *while* $n^{-1/2}\hat{S}_{[nr],p}(\lambda,\tau) \Rightarrow \omega_0[B_p(r,\lambda) + \frac{\delta(\alpha)}{\omega_0} m_{r,p}(\lambda,\tau_0)]$ *when* $\tau \neq \tau_0$, *for* $0\leq\alpha\leq 1/2$. *For a extremely large outlier,* $-1/2\leq\alpha<0$, $n^{-1/2}\hat{S}_{[nr],p}(\lambda,\tau) \to^p \frac{\delta(\alpha)}{\omega_1} m_{r,p}(\lambda,\tau_0)$, *with* $m_{r,p}(\lambda,\tau_0)$ *defined in (B.25).*

For a persistent outlier and $\alpha\in[-½,0]$, $n^{-1/2}\hat{S}_{[nr],p}(\lambda,\tau) \Rightarrow \omega_0[B_p(r,\lambda) + \frac{\delta^*(\alpha)}{\omega_0} G_p(r,\boldsymbol{\theta}_0)]$, *with* $G_p(r,\boldsymbol{\theta}_0)$ *defined in (B.28) and* $\delta^*(\alpha) = lim_{n\to\infty} c\cdot O(n^{1-\alpha})$. *Also, for the Bartlett kernel, we have the following limits for the long-run variance estimator:*

$$\hat{\omega}_n^2(m_n) \to^p \begin{cases} \omega_0^2(1+\frac{c^2}{\omega_0^2} m_n \cdot O_p(n^{-2\alpha})) & \phi_0 = 0, \tau \neq \tau_0 \\ \omega_0^2(1+\frac{c^2}{\omega_0^2} m_n \cdot O_p(n^{1-2\alpha})) & \phi_0 \neq 0 \end{cases} \qquad (3.14)$$

Proof. See Appendix B.

The result in (*i*) support the previous finding of robustness of the variance-ratio test statistic in Proposition 3(*i*), even for relatively large outliers. Also, for non-persistent outliers with magnitudes as large as $c\cdot O(n^{1/2})$, the stationarity tests will have a good performance, except in the case of wrong location (which can also be interpreted as an ignored outlier) where it can be expected a size reduction due to the reduction in the

$\hat{\omega}_n^2(m_n)$ as can see in (3.14), with $n^{-1}\cdot\hat{\omega}_n^{-2}(m_n)\cdot n^{-1}\hat{S}^2_{[nr],p}(\lambda,\tau) = O_p(m_n^{-1}) = o_p(1)$ in the extreme case of $\alpha = -1/2$. For a persistent outlier, the effect is contrary, with an increase in the empirical size of the tests, irrespective of the location with an impact in the scaled partial sum of OLS residuals of higher order than in the long-run variance estimator.

We have performed an extensive simulation study using this framework of analysis that confirm this theoretical findings, even for relatively small sample sizes

The same procedure and all the above results can also be applied in the case where there are multiple outliers, $m_0 \geq 1$, that is

$$Y_t = d_t(p,\lambda) + \eta_{m_0,t} = d_t(p,\lambda) + \eta_t + \sum_{i=1}^{m_0} \delta_i \varphi_t(\boldsymbol{\theta}_{0i}), \tag{3.15}$$

with $\boldsymbol{\theta}_{0i} = (\tau_{0i}, \phi_i)'$, $i = 1,..., m_0$, and $\varphi_t(\boldsymbol{\theta}_{0i})$ as in (3.2)-(3.3), but the estimated model is

$$Y_t = d_t(p,\lambda) + \sum_{i=1}^{m} \delta_i I_t(\tau_i) + \xi_t \tag{3.16}$$

with m the number of specified interventions, not necessarily equal to m_0. When $m = 1$ and there is no structural break in the trend function, the OLS residuals are given by

$$\hat{\xi}_{t,p}(\tau) = \hat{\eta}_{t,p} - M_{t,k}(\tau)\hat{\eta}_{k,p} + \sum_{i=1}^{m_0} \delta_i(\varphi_{t,p}(\boldsymbol{\theta}_{0i}) - M_{t,k}(\tau)\varphi_{k,p}(\boldsymbol{\theta}_{0i})) \tag{3.17}$$

with

$$\varphi_{t,p}(\boldsymbol{\theta}_{0i}) = \varphi_t(\boldsymbol{\theta}_{0i}) - \boldsymbol{x}_p'(\tfrac{t}{n})\bar{\boldsymbol{Q}}_n^{-1}(p)\tfrac{1}{n}\sum_{j=1}^{n} \boldsymbol{x}_p'(\tfrac{j}{n})\varphi_j(\boldsymbol{\theta}_{0i}) \tag{3.18}$$

Thus, for the stationarity tests we have that from

$$n^{-1/2}\hat{S}_{[nr],p}(\lambda,\tau) = n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p} + \sum_{i=1}^{m_0}\frac{\delta_i}{\sqrt{n}}\left\{\sum_{t=1}^{[nr]}\varphi_{t,p}(\boldsymbol{\theta}_{0i}) - \varphi_{k,p}(\boldsymbol{\theta}_{0i})\sum_{t=1}^{[nr]}M_{t,k}(\tau)\right\} + O_p(n^{-1/2})$$

the omission of the outliers will have no effect, at least in this component of the numerator of the test statististics, for small magnitudes δ_i as compared with the sample size, $\delta_i = c_i \cdot o(n^{1/2})$, with c_i not very large.

References

Afonso-Rodríguez, J.A. (2009), "Covariance stationarity tests with additive outliers and random measurement errors", in: Computers and Simulation in Modern Science, 154-159. Eds.: Mastorakis, N., M. Demiralp, I. Rudas, and C.A. Bulucea. WSEAS Press.

Afonso-Rodríguez, J.A. (2010), "Contrastes de estacionariedad en series temporales con outliers aditivos persistentes", Working Paper DT-E-2010-14. Instituto Universitario de Desarrollo Regional, University of La Laguna.

Andrews, D.W.K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation", Econometrica, 59, 817-858.

Breitung, J. (2002). "Nonparametric tests for unit roots and cointegration", Journal of Econometrics, 108, 343-363.

Busetti, F., and A. Harvey (2001), "Testing for the presence of a random walk in series with structural breaks", Journal of Time Series Analysis, 22(2), 127-150.

Busetti, F., and A. Harvey (2003), "Further comments on stationarity tests in series with structural breaks at unknown points", Journal of Time Series Analysis, 24(2), 137-140.

Carrion-i-Silvestre, J.L. (2003), "Breaking date misspecification error for the level shift KPSS test", Economics Letters, 81, 365-571.

Carrion-i-Silvestre, J.L., and A. Sansó (2006), "A guide to the computation of stationarity tests", Empirical Economics, 31, 433-448.

Darné, O. (2004), "The effects of additive outliers on stationarity tests: a monte carlo study", Economics Bulletin, 3(16), 1-8.

de Jong, R.M., and J. Davidson (2000), "Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices", Econometrica, 68(2), 407-423.

Elliot, G. (1999), "Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution", International Economic Review, 40(3), 767-783.

Franses, P.H., and N. Haldrup (1994), "The effects of additive outliers on tests for unit roots and cointegration", Journal of Business and Economic Statistics, 12(4), 471-478.

Giraitis, L.P., R.L. Kokoszka, and G. Teyssiere (2003), "Rescaled variance and related tests for long memory in volatility and levels", Journal of Econometrics, 112(2), 265-294.

Hadri, K., and Y. Rao (2009), "KPSS test and model misspecifications", Applied Economics Letters, 16(2), 1187-1190.

Hobijn, B., P.H. Franses, and M. Ooms (2004), "Generalizations of the KPSS-test for stationarity", Statistica Neerlandica, 58(4), 483-502.

Jönsson, K. (2006), "Finite-sample stability of the KPSS test". Working Paper 2006:23, Department of Economics, Lund University, Sweden.

Kurozumi, E. (2002), "Testing for stationarity with a break", Journal of Econometrics, 108, 63-99.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root", Journal of Econometrics, 54(1-3), 159-178.

Lee, J., C.J. Huang, and Y. Shin (1997), "On stationarity tests in the presence of structural breaks", Economic Letters, 55, 165-172.

Müller, U.K. (2005), "Size and power of tests of stationarity in highly autocorrelated time series", Journal of Econometrics, 128, 195-213.

Müller, U.K., and G. Elliot (2003), "Tests for unit roots and the initial condition", Econometrica, 71(4), 1269-1286.

Otero, J., and J. Smith (2005), "The KPSS test with outliers", Computational Economics, 26, 241-249.

Rodrigues, P.M.M. and A. Rubia (2010), "The effects of additive outliers and measurement errors when testing for structural breaks in variance", Working Paper 11-2010, Banco de Portugal.

Shin, D.W., S. Sarkar, and J.H. Lee (1996), "Unit root tests for time series with outliers", Statistics and Probability Letters, 30, 189-197.

Xiao, Z. (2001), "Testing the null hypothesis of stationarity against an autoregressive unit root alternative", Journal of Time Series Analysis, 22(1), 87-103.

Xiao, Z., and L.R. Lima (2007), "Testing covariance stationarity", Econometric Reviews, 6(6), 643-667.

Appendix A. Proof of Proposition 1

From auxiliary regression (2.10) and standard results from OLS estimation, conditional on a given value of $\lambda \in (0, 1)$, we have that

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{p,n}(\lambda) &= \left(\sum_{t=1}^{[n\lambda]} \boldsymbol{x}_{t,p}\boldsymbol{x}'_{t,p}\right)^{-1} \sum_{t=1}^{[n\lambda]} \boldsymbol{x}_{t,p}Y_t = \boldsymbol{\beta}_p + \boldsymbol{D}_{p,n}\left(\frac{1}{n}\sum_{t=1}^{[n\lambda]} \boldsymbol{x}_p(\tfrac{t}{n})\boldsymbol{x}'_p(\tfrac{t}{n})\right)^{-1} \frac{1}{n}\sum_{t=1}^{[n\lambda]} \boldsymbol{x}_p(\tfrac{t}{n})\eta_t \\ &= \boldsymbol{\beta}_p + \boldsymbol{D}_{p,n}\overline{\boldsymbol{Q}}_{[n\lambda]}^{-1}(p)n^{-1}\cdot\boldsymbol{H}_{[n\lambda]}(p)\end{aligned} \tag{A.1}$$

and

$$\hat{\boldsymbol{\alpha}}_{p,n}(\lambda) = -\hat{\boldsymbol{\beta}}_{p,n}(\lambda) + \hat{\boldsymbol{\theta}}_{p,n}(\lambda) \tag{A.2}$$

with

$$\hat{\boldsymbol{\theta}}_{p,n}(\lambda) = \left(\sum_{t=[n\lambda]+1}^{n} \boldsymbol{x}_{t,p}\boldsymbol{x}'_{t,p} \right)^{-1} \sum_{t=[n\lambda]+1}^{n} \boldsymbol{x}_{t,p} Y_t = \boldsymbol{\theta}_p + \boldsymbol{D}_{p,n} \left(\tfrac{1}{n} \sum_{t=[n\lambda]+1}^{n} \boldsymbol{x}_p(\tfrac{t}{n})\boldsymbol{x}'_p(\tfrac{t}{n}) \right)^{-1} n^{-1} \cdot \sum_{t=[n\lambda]+1}^{n} \boldsymbol{x}_p(\tfrac{t}{n})\eta_t$$

$$= \boldsymbol{\theta}_p + \boldsymbol{D}_{p,n} \bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p) n^{-1} \cdot \boldsymbol{J}_{n-[n\lambda]}(p) \tag{A.3}$$

where $\bar{\boldsymbol{U}}_{n-[n\lambda]}(p) = \boldsymbol{Q}_n(p) - \boldsymbol{Q}_{[n\lambda]}(p)$, $\boldsymbol{J}_{n-[n\lambda]}(p) = \boldsymbol{H}_n(p) - \boldsymbol{H}_{[n\lambda]}(p)$ and it is assumed that the inverses in (A.1) and (A.3) exist and are finite, with

$$\bar{\boldsymbol{Q}}_{[n\lambda]}(p) \to \int_0^\lambda \boldsymbol{x}_p(s)\boldsymbol{x}'_p(s)ds = \bar{\boldsymbol{Q}}_\lambda(p) \tag{A.4}$$

and

$$\bar{\boldsymbol{U}}_{n-[n\lambda]}(p) \to \int_\lambda^1 \boldsymbol{x}_p(s)\boldsymbol{x}'_p(s)ds = \bar{\boldsymbol{Q}}_{1-\lambda}(p) = \bar{\boldsymbol{Q}}_1(p) - \bar{\boldsymbol{Q}}_\lambda(p) \tag{A.5}$$

From (A.1) and (A.3) we can write

$$n^v \cdot \boldsymbol{D}^{-1}_{p,n}(\hat{\boldsymbol{\beta}}_{p,n}(\lambda) - \boldsymbol{\beta}_p) = \bar{\boldsymbol{Q}}^{-1}_{[n\lambda]}(p) \cdot n^{v-1} \boldsymbol{H}_{[n\lambda]}(p) \tag{A.6}$$

and

$$n^v \cdot \boldsymbol{D}^{-1}_{p,n}(\hat{\boldsymbol{\theta}}_{p,n}(\lambda) - \boldsymbol{\theta}_p) = \bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p) \cdot n^{v-1} \boldsymbol{J}_{n-[n\lambda]}(p) \tag{A.7}$$

with different values for v depending on the scaling needed to get finite asymptotic limits for (A.6) and (A.7). Thus, v takes the value $-1/2$ under the null of a unit root, and $v = 1/2$ under the null of stationarity together with $\sigma_1^2 = 0$. Using (2.11), we have that

$$n^{-1/2}\eta_{[nr]} = O_p(n^{-1/2}) + n^{-1/2}\sum_{j=1}^{[nr]} \varepsilon_{1,j} \Rightarrow \omega_1 W_1(r) \tag{A.8}$$

under a unit root process (where $\sigma_1^2 > 0$), and

$$n^{-1/2}\sum_{t=1}^{[nr]} \eta_t = n^{-1/2}\sum_{t=1}^{[nr]} \varepsilon_{0,t} + O_p(n^{-1/2}) + n\left\{ n^{-1}\sum_{t=1}^{[nr]} (n^{-1/2}\sum_{j=1}^{t} \varepsilon_{1,j}) \right\} \tag{A.9}$$

with weak limit $n^{-1/2}\sum_{t=1}^{[nr]}\eta_t = n^{-1/2}\sum_{t=1}^{[nr]}\varepsilon_{0,t} + O_p(n^{-1/2}) \Rightarrow \omega_0 W_0(r)$ under stationarity (that is, when $\sigma_1^2 = 0$ so that the last term in (A.9) vanishes). If $\sigma_1^2 > 0$, from (A.9), we have instead the following limit result

$$n^{-3/2}\sum_{t=1}^{[nr]} \eta_t = O_p(n^{-1}) + O_p(n^{-3/2}) + n^{-1}\sum_{t=1}^{[nr]} (n^{-1/2}\sum_{j=1}^{t} \varepsilon_{1,j}) \Rightarrow \omega_1 \int_0^r W_1(s)ds \tag{A.10}$$

that will be used to establish the asymptotic null distribution of Breitung's variance-ratio unit root test statistic and also the the consistency and asymptotic distribution of stationarity tests under the unit root alternative. Thus, with (A.8) and (A.9) we have that

$$n^{v-1}\boldsymbol{H}_{[n\lambda]}(p) \Rightarrow \begin{cases} \omega_1 \int_0^\lambda \boldsymbol{x}_p(s) W_1(s)ds & v = -1/2 \\ \omega_0 \int_0^\lambda \boldsymbol{x}_p(s) dW_0(s) & v = 1/2 \end{cases}$$

and

$$n^{v-1}\boldsymbol{J}_{n-[n\lambda]}(p) \Rightarrow \begin{cases} \omega_1 \int_\lambda^1 \boldsymbol{x}_p(s) W_1(s)ds & v = -1/2 \\ \omega_0 \int_\lambda^1 \boldsymbol{x}_p(s) dW_0(s) & v = 1/2 \end{cases}$$

as $n \to \infty$ in (A.6) and (A.7). Given now that the OLS residuals from (2.10) are given by

$$\hat{\eta}_{t,p}(\lambda) = \eta_t - \boldsymbol{x}'_{t,p}(\hat{\boldsymbol{\beta}}_{p,t}(\lambda) - \boldsymbol{\beta}_{p,t}(\lambda)) \tag{A.11}$$

where

$$\hat{\beta}_{p,t}(\lambda) = \hat{\beta}_{p,n}(\lambda) + h_t(\lambda)\cdot\hat{\alpha}_{p,n}(\lambda) = (1-h_t(\lambda))\cdot\hat{\beta}_{p,n}(\lambda) + h_t(\lambda)\cdot\hat{\theta}_{p,n}(\lambda) \tag{A.12}$$

and thus

$$\hat{\beta}_{p,t}(\lambda) - \beta_{p,t}(\lambda) = (1-h_t(\lambda))\cdot(\hat{\beta}_{p,n}(\lambda) - \beta_p) + h_t(\lambda)\cdot(\hat{\theta}_{p,n}(\lambda) - \theta_p) \tag{A.13}$$

then we can write

$$\begin{aligned}\hat{\eta}_{t,p}(\lambda) = \eta_t - n^{-v}\cdot\boldsymbol{x}'_p(\tfrac{t}{n})((1-h_t(\lambda))\cdot[n^v\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p)] \\ + h_t(\lambda)\cdot[n^v\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\theta}_{p,n}(\lambda)-\theta_p)])\end{aligned} \tag{A.14}$$

Scaling (A.14) by $n^{-1/2}$, using (A.8) and setting $v = -1/2$, then we have that

$$\begin{aligned}n^{-1/2}\hat{\eta}_{[nr],p}(\lambda) = n^{-1/2}\eta_{[nr]} - \boldsymbol{x}'_p(\tfrac{[nr]}{n})((1-h_{[nr]}(\lambda))\cdot[n^{-1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p)] \\ + h_{[nr]}(\lambda)\cdot[n^{-1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\theta}_{p,n}(\lambda)-\theta_p)])\end{aligned} \tag{A.15}$$

with weak limit $V_p(r,\lambda)$ as defined in the proposition, where $I(r \le \lambda) = 1 - h_{[nr]}(\lambda)$ and $I(r > \lambda) = h_{[nr]}(\lambda)$ as $n \to \infty$. The final result follows using the CMP. With $v = 1/2$, we get

$$\begin{aligned}n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p}(\lambda) = n^{-1/2}\sum_{t=1}^{[nr]}\eta_t - n^{-1}\sum_{t=1}^{[nr]}\boldsymbol{x}'_p(\tfrac{t}{n})((1-h_t(\lambda)\cdot[n^{1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p)] \\ + h_t(\lambda)\cdot[n^{1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\theta}_{p,n}(\lambda)-\theta_p)])\end{aligned}$$

that is

$$\begin{aligned}n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p}(\lambda) &= n^{-1/2}\sum_{t=1}^{[nr]}\eta_t - n^{-1}\sum_{t=1}^{[nr]}\boldsymbol{x}'_p(\tfrac{t}{n})[n^{1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p)]\cdot I(r \le \lambda) \\ &= n^{-1/2}\sum_{t=1}^{[nr]}\eta_t - n^{-1}\sum_{t=1}^{[n\lambda]}\boldsymbol{x}'_p(\tfrac{t}{n})[n^{1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p)] \\ &\quad - n^{-1}\sum_{t=[n\lambda]+1}^{[nr]}\boldsymbol{x}'_p(\tfrac{t}{n})\cdot[n^{1/2}\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\theta}_{p,n}(\lambda)-\theta_p)]\cdot I(r > \lambda)\end{aligned} \tag{A.16}$$

Thus, under stationarity, $\sigma_1^2 = 0$, we get the weak limit $B_p(r,\lambda)$ as defined in equations (2.22)-(2.24), and the final result follows from the CMP.

■

Appendix B. Proof of Proposition 4

If we consider OLS fitting of the auxiliary regression in (3.4) for observations in the first subsample, $t = 1, ..., [n\cdot\lambda]$, we have

$$\hat{\beta}_{p,n}(\lambda,\tau) = \hat{\beta}_{p,n}(\lambda) - \boldsymbol{Q}^{-1}_{[n\lambda]}(p)\boldsymbol{x}_{k,p}\cdot\hat{\delta}_n(\lambda,\tau)\cdot I(\tau \le \lambda) \tag{B.1}$$

$$n^v\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\beta}_{p,n}(\lambda)-\beta_p) = \bar{\boldsymbol{Q}}^{-1}_{[n\lambda]}(p)n^{v-1}\boldsymbol{H}_{0,[n\lambda]}(p) \tag{B.2}$$

with

$$\boldsymbol{H}_{0,[n\lambda]}(p) = \sum_{t=1}^{[n\lambda]}\boldsymbol{x}_p(\tfrac{t}{n})\eta_{0,t} = \boldsymbol{H}_{[n\lambda]}(p) + \delta\sum_{t=1}^{[n\lambda]}\boldsymbol{x}_p(\tfrac{t}{n})\varphi_t(\theta_0) \tag{B.3}$$

while for the second subsample, that is, for $t = [n\cdot\lambda]+1,...,n$, we get

$$\hat{\theta}_{p,n}(\lambda,\tau) = \hat{\theta}_{p,n}(\lambda) - \boldsymbol{U}^{-1}_{n-[n\lambda]}(p)\boldsymbol{x}_{k,p}\cdot\hat{\delta}_n(\lambda,\tau)\cdot I(\tau > \lambda) \tag{B.4}$$

$$\hat{\alpha}_{p,n}(\lambda,\tau) = -\hat{\beta}_{p,n}(\lambda,\tau) + \hat{\theta}_{p,n}(\lambda,\tau) \tag{B.5}$$

$$n^v\cdot\boldsymbol{D}^{-1}_{p,n}(\hat{\theta}_{p,n}(\lambda)-\theta_p) = \bar{\boldsymbol{U}}^{-1}_{n-[n\lambda]}(p)n^{v-1}\boldsymbol{J}_{0,n-[n\lambda]}(p) \tag{B.6}$$

with

$$\boldsymbol{J}_{0,n-[n\lambda]}(p) = \sum_{t=[n\lambda]+1}^{n}\boldsymbol{x}_p(\tfrac{t}{n})\eta_{0,t} = \boldsymbol{J}_{n-[n\lambda]}(p) + \delta\sum_{t=[n\lambda]+1}^{n}\boldsymbol{x}_p(\tfrac{t}{n})\varphi_t(\theta_0) \tag{B.7}$$

by using the structure of the error terms $\eta_{0,t}$ in (3.5) and where we have called $k = [n\cdot\tau]$ the specified location of the outlier in the sample. In (B.1) and (B.4), $\hat{\delta}_n(\lambda,\tau)$ is the OLS estimator of the outlier magnitude given by

$$\hat{\delta}_n(\lambda,\tau) = m_{k,k}^{-1}(\lambda,\tau)\cdot\hat{\eta}_{0k,p}(\lambda,\tau) \tag{B.8}$$

with $\hat{\eta}_{0t,p}(\lambda,\tau)$ be the sequence of OLS residuals obtained using the parameter vector estimator without the dummy variable for the observations in the subsample which contains the specified location of the outlier, that is

$$\hat{\eta}_{0t,p}(\lambda,\tau) = Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n}(\lambda)\cdot I(\tau\le\lambda) - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\theta}}_{p,n}(\lambda)\cdot I(\tau>\lambda) \tag{B.9}$$

and $m_{k,k}(\lambda,\tau) = m_{t,k}(\lambda,\tau)$ with $t = k$, where

$$m_{t,k}(\lambda,\tau) = 1-(1-h_t(\lambda))\tfrac{1}{n}\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_{[n\lambda]}^{-1}(p)\boldsymbol{x}_p(\tfrac{k}{n}) - h_t(\lambda)\tfrac{1}{n}\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{U}}_{n-[n\lambda]}^{-1}(p)\boldsymbol{x}_p(\tfrac{k}{n}) \tag{B.10}$$

Using now (B.1) and (B.4), we have, as in (A.12), that

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{p,t}(\lambda,\tau) &= (1-h_t(\lambda))\cdot\hat{\boldsymbol{\beta}}_{p,n}(\lambda,\tau) + h_t(\lambda)\cdot\hat{\boldsymbol{\theta}}_{p,n}(\lambda,\tau)\\ &= (1-h_t(\lambda))\cdot\hat{\boldsymbol{\beta}}_{p,n}(\lambda) + h_t(\lambda)\cdot\hat{\boldsymbol{\theta}}_{p,n}(\lambda)\\ &\quad -\hat{\delta}_n(\lambda,\tau)\cdot[\boldsymbol{Q}_{[n\lambda]}^{-1}(p)\boldsymbol{x}_{k,p}\cdot I(\tau\le\lambda) + \boldsymbol{U}_{n-[n\lambda]}^{-1}(p)\boldsymbol{x}_{k,p}\cdot I(\tau>\lambda)]\end{aligned} \tag{B.11}$$

so that the sequence of residuals from the OLS fitting of the auxiliary regression in (3.4) is given by

$$\begin{aligned}\hat{\xi}_{t,p}(\lambda,\tau) &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,t}(\lambda,\tau) - I_t(\tau)\cdot\hat{\delta}_n(\lambda,\tau)\\ &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n}(\lambda)(1-h_t(\lambda)) - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\theta}}_{p,n}(\lambda)h_t(\lambda)\\ &\quad -\hat{\delta}_n(\lambda,\tau)[I_t(\tau) - (1-h_t(\lambda))\boldsymbol{x}'_{t,p}\boldsymbol{Q}_{[n\lambda]}^{-1}(p)\boldsymbol{x}_{k,p}I(\tau\le\lambda)\\ &\qquad - h_t(\lambda)\boldsymbol{x}'_{t,p}\boldsymbol{U}_{n-[n\lambda]}^{-1}(p)\boldsymbol{x}_{k,p}\cdot I(\tau>\lambda)]\end{aligned} \tag{B.12}$$

that is

$$\hat{\xi}_{t,p}(\lambda,\tau) = \hat{\eta}_{0t,p}(\lambda) - \hat{\delta}_n(\lambda,\tau)\cdot m_{t,k}^{*}(\lambda,\tau) \tag{B.13}$$

where

$$\begin{aligned}\hat{\eta}_{0t,p}(\lambda) &= Y_t - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n}(\lambda)(1-h_t(\lambda)) - \boldsymbol{x}'_{t,p}\hat{\boldsymbol{\theta}}_{p,n}(\lambda)h_t(\lambda)\\ &= \eta_{0,t} - (1-h_t(\lambda))n^{-v}\boldsymbol{x}'_p(\tfrac{t}{n})[n^{v}\cdot\boldsymbol{D}_{p,n}^{-1}(\hat{\boldsymbol{\beta}}_{p,n}(\lambda)-\boldsymbol{\beta}_p)]\\ &\qquad - h_t(\lambda)n^{-v}\boldsymbol{x}'_p(\tfrac{t}{n})[n^{v}\cdot\boldsymbol{D}_{p,n}^{-1}(\hat{\boldsymbol{\theta}}_{p,n}(\lambda)-\boldsymbol{\theta}_p)]\\ &= \hat{\eta}_{t,p}(\lambda) + \delta\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)\end{aligned} \tag{B.14}$$

and $m_{t,k}^{*}(\lambda,\tau) = m_{t,k}(\lambda,\tau)$ for $t = k$, with

$$\begin{aligned}m_{t,k}^{*}(\lambda,\tau) &= I_t(\tau) - (1-h_t(\lambda))\cdot[\tfrac{1}{n}\boldsymbol{x}'_p(\tfrac{t}{n})]\bar{\boldsymbol{Q}}_{[n\lambda]}^{-1}(p)\boldsymbol{x}_p(\tfrac{k}{n})\\ &\qquad - h_t(\lambda)\cdot[\tfrac{1}{n}\boldsymbol{x}'_p(\tfrac{t}{n})]\bar{\boldsymbol{U}}_{n-[n\lambda]}^{-1}(p)\boldsymbol{x}_p(\tfrac{k}{n})\end{aligned} \tag{B.15}$$

From the OLS residuals defined in (B.9) and (B.14), we have that $\hat{\eta}_{0k,p}(\lambda) = \hat{\eta}_{0k,p}(\lambda,\tau)$ in (B.8) because $1-h_k(\lambda) = I(k\le[n\lambda]) = I(\tau\le\lambda)$ and $h_k(\lambda) = I(k>[n\lambda]) = I(\tau>\lambda)$. In (B.14), $\hat{\eta}_{t,p}(\lambda)$ is as in equation (A.14) in Appendix A, $\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) = m_{t,k_0}^{*}(\lambda,\tau_0)$ when $\phi_0 = 0$, and

$$\begin{aligned}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) &= \phi_0^{t-k_0}H_t(\tau_0) - (1-h_t(\lambda))\cdot\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{Q}}_{[n\lambda]}^{-1}(p)\boldsymbol{G}_{p,n}(\lambda,\boldsymbol{\theta}_0)\\ &\qquad - h_t(\lambda)\cdot\boldsymbol{x}'_p(\tfrac{t}{n})\bar{\boldsymbol{U}}_{n-[n\lambda]}^{-1}(p)\boldsymbol{G}_{p,n}(1,\boldsymbol{\theta}_0)\end{aligned} \tag{B.16}$$

where $\boldsymbol{G}_{p,n}(a,\boldsymbol{\theta}_0)$ as in (3.10), with $a = \lambda$ for $\tau_0 \le \lambda$ and $a = 1$ when $\tau_0 > \lambda$ in the case $\phi_0 \ne 0$. Now, from equation (B.13), using (B.8) and (B.14), the sequence of OLS residuals is given by

$$\hat{\xi}_{t,p}(\lambda,\tau) = \hat{\eta}_{t,p}(\lambda) + \delta\cdot\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) - M_{t,k}(\lambda,\tau)\cdot[\hat{\eta}_{k,p}(\lambda) + \delta\cdot\varphi_{k,p}(\lambda,\boldsymbol{\theta}_0)] \tag{B.17}$$

with $M_{t,k}(\lambda,\tau)=m_{k,k}^{-1}(\lambda,\tau)\cdot m_{t,k}^{*}(\lambda,\tau)=I_t(\tau)+O(n^{-1})$. The leading term for the variance-ratio test statistic is given by

$$n^{-1/2}\cdot\hat{\xi}_{t,p}(\lambda,\tau)=n^{-1/2}\cdot\hat{\eta}_{t,p}(\lambda)+\delta_n(\alpha)\cdot\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) \\ -M_{t,k}(\lambda,\tau)\cdot[n^{-1/2}\cdot\hat{\eta}_{k,p}(\lambda)+\delta_n(\alpha)\cdot\varphi_{k,p}(\lambda,\boldsymbol{\theta}_0)] \tag{B.18}$$

while for the stationarity tests we need to evaluate the behaviour of

$$n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\lambda,\tau)=n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\eta}_{t,p}(\lambda)+\delta_n(\alpha)\cdot\sum_{t=1}^{[nr]}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) \\ -[n^{-1/2}\cdot\hat{\eta}_{k,p}(\lambda)+\delta_n(\alpha)\cdot\varphi_{k,p}(\lambda,\boldsymbol{\theta}_0)]\sum_{t=1}^{[nr]}M_{t,k}(\lambda,\tau) \tag{B.19}$$

where $\delta_n(\alpha)=n^{-1/2}\cdot\delta$. Under the assumption that the outlier magnitude parameter δ is of order $\delta=c\cdot O(n^{1/2-\alpha})$ with $0 < c < 1$ and $-1/2 \le \alpha \le 1/2$, then $\delta_n(\alpha)=c\cdot O(n^{-\alpha})$. With $\alpha = 0$, $\delta_n(0)=c$, while for $0 < \alpha \le 1/2$, $\delta = c\cdot o(n^{1/2})$, and $\delta_n(\alpha)=o(1)$ so that the effect of the outlier will become asymptotically negligible. When $\alpha = 1/2$, we have the case of fixed magnitude, $\delta = c$. Otherwise, for $-1/2 \le \alpha < 0$ $\delta_n(\alpha)\to\infty$ as $n \to \infty$ but this corresponds to a extremely large outlier magnitude of order greater than $n^{1/2}$, which gives very little realistic values. In what follows let $\delta(\alpha)$ denotes the corresponding limit of $\delta_n(\alpha)$ associated with the value of α. In the case of a pure AO ($\phi_0 = 0$), $\varphi_{[nr],p}(\lambda,\boldsymbol{\theta}_0)\to\varphi_{r,p}(\lambda,\boldsymbol{\theta}_0)=I(r=\tau_0)$ as $n\to\infty$, while for a persistent outlier effect ($\phi_0 \ne 0$) and under Assumption 1,

$$\varphi_{[nr],p}(\lambda,\boldsymbol{\theta}_0)\to\varphi_{r,p}(\lambda,\boldsymbol{\theta}_0)=e^{(r-\tau_0)\gamma_0}I(r\ge\tau_0)-\boldsymbol{x}'_p(r)\boldsymbol{Q}_\lambda^{-1}(p)\boldsymbol{G}_{p,\gamma_0}(\lambda,\boldsymbol{\theta}_0)\cdot I(r\le\lambda) \\ -\boldsymbol{x}'_p(r)\boldsymbol{Q}_{1-\lambda}^{-1}(p)\boldsymbol{G}_{p,\gamma_0}(1,\boldsymbol{\theta}_0)\cdot I(r>\lambda) \tag{B.20}$$

so that

$$n^{-1/2}\cdot\hat{\xi}_{[nr],p}(\lambda,\tau)\Rightarrow\omega_1[V_p(r,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{r,p}(\lambda,\boldsymbol{\theta}_0)] \\ -\omega_1 I(r=\tau)[V_p(\tau,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\cdot\varphi_{\tau,p}(\lambda,\boldsymbol{\theta}_0)] \tag{B.21}$$

with $V_p(r, \lambda)$ as in the standard case under the unit root hypothesis (see Proposition 1), and

$$n^{-3/2}\cdot\hat{S}_{[nr],p}(\lambda,\tau)=n^{-1}\sum_{t=1}^{[nr]}n^{-1/2}\cdot\hat{\xi}_{t,p}(\lambda,\tau)\Rightarrow\omega_1\int_0^r[V_p(s,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{s,p}(\lambda,\boldsymbol{\theta}_0)]ds \\ -\omega_1[V_p(\tau,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\cdot\varphi_{\tau,p}(\lambda,\boldsymbol{\theta}_0)]\int_0^r I(s=\tau)ds \\ =\omega_1\int_0^r[V_p(s,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{s,p}(\lambda,\boldsymbol{\theta}_0)]ds \tag{B.22}$$

where the last equality results from the integral of the point indicator function and the assumption that $\delta(\alpha)$ does not diverge. The same results follows directly from (B.21) in the case of a non-persistent outlier ($\phi_0 = 0$) where $\varphi_{r,p}(\lambda,\boldsymbol{\theta}_0)-I(r=\tau)\varphi_{\tau,p}(\lambda,\boldsymbol{\theta}_0)$ is zero for $\tau = \tau_0$ and $I(r=\tau_0)$ otherwise. Also, for the denominator of the variance-ratio UR test statistic we have from (B.21) that

$$\frac{1}{n}\sum_{t=1}^{n}(\tfrac{1}{\sqrt{n}}\hat{\xi}_{t,p}(\lambda,\tau))^2\Rightarrow\omega_1^2\cdot\int_0^1[V_p(s,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{s,p}(\lambda,\boldsymbol{\theta}_0)]^2ds \\ +\omega_1^2[V_p(\tau,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{\tau,p}(\lambda,\boldsymbol{\theta}_0)]^2\int_0^1 I(s=\tau)ds \\ -2\omega_1^2[V_p(\tau,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{\tau,p}(\lambda,\boldsymbol{\theta}_0)]\int_0^1 I(s=\tau)\cdot[V_p(s,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{s,p}(\lambda,\boldsymbol{\theta}_0)]ds$$

where the last two terms involving the integral of the point indicator function vanish, so that

$$\frac{1}{n}\sum_{t=1}^{n}(\tfrac{1}{\sqrt{n}}\hat{\xi}_{t,p}(\lambda,\tau))^2\Rightarrow\omega_1^2\cdot\int_0^1[V_p(s,\lambda)+\tfrac{\delta(\alpha)}{\omega_1}\varphi_{s,p}(\lambda,\boldsymbol{\theta}_0)]^2ds \tag{B.23}$$

In the simplest case $p = 0$, the limit equation in (B.20) is given by

$$\varphi_{r,0}(\lambda,\boldsymbol{\theta}_0) = e^{(r-\tau_0)\gamma_0} I(r \geq \tau_0) - \gamma_0^{-1}\left\{\tfrac{1}{\lambda}[e^{(\lambda-\tau_0)\gamma_0} - 1] I(r,\tau_0 \leq \lambda) + \tfrac{1}{1-\lambda}[e^{(1-\tau_0)\gamma_0} - 1] I(r,\tau_0 > \lambda)\right\}$$

When $\phi_0 = 0$, $\varphi_{r,p}(\lambda,\boldsymbol{\theta}_0) = I(r = \tau_0)$ so it is of application the same argument as before concerning the effect of the possible misspecification of the AO location in (B.22) and (B.23). Thus, irrespective of the true location of the outlier for any $0 \leq \alpha \leq 1/2$ the effect is asymptotically negligible. Except in the case of correct location of the outlier, $\tau = \tau_0$, for a extremely large outlier, $-1/2 \leq \alpha < 0$, both (B.22) and (B.23) will diverge to infinity at the same rate $O_p(n^{-\alpha})$. For a persistent outlier ($\phi_0 \neq 0$) and $0 \leq \alpha \leq 1/2$, $\delta(\alpha)$ will dominate in (B.22) and (B.23) so that the effect is asymptotically negligible irrespective of the location of the outlier. Given (B.20), and only under wrong location and an extremely large outlier, $-1/2 \leq \alpha < 0$, we will encounter with the above divergences.

Considering now the term in (B.19) under stationarity, where $n^{-1/2}\cdot\hat{\eta}_{k,p}(\lambda) = O_p(n^{-1/2})$, then

$$\begin{aligned} n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\lambda,\tau) &= n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\eta}_{t,p}(\lambda) \\ &+ \delta_n(\alpha)\cdot\left\{\sum_{t=1}^{[nr]}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) - \varphi_{k,p}(\lambda,\boldsymbol{\theta}_0)\cdot\sum_{t=1}^{[nr]}M_{t,k}(\lambda,\tau)\right\} + O_p(n^{-1/2}) \end{aligned} \tag{B.24}$$

From (B.15), we have that

$$\sum_{t=1}^{[nr]} m_{t,k}^*(\lambda,\tau) = I(r \geq \tau) - \tfrac{1}{n}\left\{\sum_{t=1}^{[nr]}(1-h_t(\lambda))\boldsymbol{x}_p'(\tfrac{t}{n})\bar{\boldsymbol{Q}}_{[n\lambda]}^{-1}(p) + \sum_{t=1}^{[nr]}h_t(\lambda)\boldsymbol{x}_p'(\tfrac{t}{n})\bar{\boldsymbol{U}}_{n-[n\lambda]}^{-1}(p)\right\}\boldsymbol{x}_p(\tfrac{k}{n})$$

where $\sum_{t=1}^{[nr]} m_{t,k}^*(\lambda,\tau) \to m_{r,p}(\lambda,\tau)$ as $n\to\infty$, with limit $m_{r,p}(\lambda,\tau)$ given by

$$\begin{aligned} m_{r,p}(\lambda,\tau) &= I(r \geq \tau) - \int_0^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_\lambda^{-1}(p)\,\boldsymbol{x}_p(\tfrac{\tau}{n}) & r \leq \lambda \\ &= I(r \geq \tau) - \left\{\int_0^\lambda \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_\lambda^{-1}(p) + \int_\lambda^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_{1-\lambda}^{-1}(p)\right\}\boldsymbol{x}_p(\tfrac{\tau}{n}) & r > \lambda \end{aligned} \tag{B.25}$$

and

$$\sum_{t=1}^{[nr]} M_{t,k}(\lambda,\tau) = \tfrac{1}{m_{k,k}(\lambda,\tau)}\sum_{t=1}^{[nr]} m_{t,k}^*(\lambda,\tau) = \sum_{t=1}^{[nr]} m_{t,k}^*(\lambda,\tau) + O(n^{-1}) \to m_{r,p}(\lambda,\tau)$$

then, using that $\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) = m_{t,k_0}^*(\lambda,\tau_0)$ when $\phi_0 = 0$,

$$n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\lambda,\tau) \Rightarrow \omega_0[B_p(r,\lambda) + \tfrac{\delta(\alpha)}{\omega_0}(m_{r,p}(\lambda,\tau_0) - I(\tau=\tau_0)\cdot m_{r,p}(\lambda,\tau)] \tag{B.26}$$

From (B.26), the influence of the AO in the asymptotic behavior of the scaled partial sum of OLS residuals depends crucially on the magnitude of $\delta(\alpha)$ and the possible difference between the true and specified location of the outlier. The term between parenthesis is zero when $\tau = \tau_0$ and $m_{r,p}(\lambda,\tau_0)$ otherwise. When $\tau \neq \tau_0$, the final effect depends on the magnitude of $\delta(\alpha)$ and the true location of the non-persistent AO. For values of $\delta(\alpha) = c$ or $o_p(1)$ the only effect must come from the long-run variance estimator, while for a extremely large outlier (B.26) will diverge at the rate $O_p(n^{-\alpha})$, $-1/2 \leq \alpha < 0$. In the case $\phi_0 \neq 0$, (B.24) must be written as

$$n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\lambda,\tau) = n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\eta}_{t,p}(\lambda) + \delta_n^*(\alpha)\cdot\left\{\tfrac{1}{n}\sum_{t=1}^{[nr]}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)\right\} + O_p(n^{-1/2}) \tag{B.27}$$

where now $\delta_n^*(\alpha) = \delta_n(\alpha-1) = c\cdot O(n^{1-\alpha})$ diverges to infinity for all given values of α at the rate $O(n^{1-\alpha})$. In this case $n^{-1/2}\cdot\hat{S}_{[nr],p}(\lambda,\tau) = c\cdot O_p(n^{1-\alpha})$, and the term between brackets has a finite limit as $n\to\infty$ given by

$$n^{-1}\sum_{t=1}^{[nr]}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)\to G_p(r,\boldsymbol{\theta}_0)=G_{0,\gamma_0}(r,\boldsymbol{\theta}_0)-\int_0^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_\lambda^{-1}(p)\boldsymbol{G}_{p,\gamma_0}(\lambda,\boldsymbol{\theta}_0) \quad \text{if } r\le\lambda$$
$$=G_{0,\gamma_0}(r,\boldsymbol{\theta}_0)-\Big\{\int_0^\lambda \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_\lambda^{-1}(p)\boldsymbol{G}_{p,\gamma_0}(\lambda,\boldsymbol{\theta}_0)$$
$$+\int_\lambda^r \boldsymbol{x}_p'(s)ds\,\boldsymbol{Q}_{1-\lambda}^{-1}(p)\boldsymbol{G}_{p,\gamma_0}(1,\boldsymbol{\theta}_0)\Big\} \quad \text{if } r>\lambda \tag{B.28}$$

under Assumption 1, with $G_{0,\gamma_0}(r,\boldsymbol{\theta}_0)\neq 0$ only for $r \ge \tau_0$. The effect of this term differs depending on the values of r, τ_0 and p. When the only systematic component is the dummy variable, then (B.28) is given by $G(r,\boldsymbol{\theta}_0)=G_{0,\gamma_0}(r,\boldsymbol{\theta}_0)I(r\ge\tau_0)$, while for p = 0 (demeaned data) without structural break, then $G_0(r,\boldsymbol{\theta}_0)=G_{0,\gamma_0}(r,\boldsymbol{\theta}_0)I(r\ge\tau_0)-r\cdot G_{0,\gamma_0}(1,\boldsymbol{\theta}_0)$. With this, the limit of (B.27) as $n\to\infty$ is given by

$$n^{-1/2}\cdot\sum_{t=1}^{[nr]}\hat{\xi}_{t,p}(\lambda,\tau)\Rightarrow\omega_0\left(B_p(r,\lambda)+\tfrac{\delta^*(\alpha)}{\omega_0}G_p(r,\boldsymbol{\theta}_0)\right) \tag{B.29}$$

From (B.28) we can see that the wrong location of the outlier in the sample has no effect on the behavior of (B.29), and that for small values of c the divergence at the rate $O_p(n^{1-\alpha})$ could be controlled for $0 < \alpha \le 1/2$. On the other hand, for the computation of the nonparametric long-run variance estimator, the sample autocovariance of order j based on the OLS residuals in (B.17) is given by

$$\hat{\gamma}_n(j,\tau)=\frac{1}{n}\sum_{t=j+1}^{n}\hat{\xi}_{t,p}(\lambda,\tau)\hat{\xi}_{t-j,p}(\lambda,\tau)=\frac{1}{n}\sum_{t=j+1}^{n}\hat{\eta}_{t,p}(\lambda)\hat{\eta}_{t-j,p}(\lambda)$$
$$+\delta_n(\alpha)\frac{1}{\sqrt{n}}\sum_{t=j+1}^{n}[\hat{\eta}_{t,p}(\lambda)\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)+\hat{\eta}_{t-j,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)]$$
$$+\delta_n^2(\alpha)\sum_{t=j+1}^{n}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0) \tag{B.30}$$
$$-\hat{\eta}_{0k,p}(\lambda)\frac{1}{n}\sum_{t=j+1}^{n}[\hat{\eta}_{0t-j,p}(\lambda)M_{t,k}(\lambda,\tau)+\hat{\eta}_{0t,p}(\lambda)M_{t-j,k}(\lambda,\tau)]$$
$$+\hat{\eta}_{0k,p}^2(\lambda)\frac{1}{n}\sum_{t=j+1}^{n}M_{t,k}(\lambda,\tau)M_{t-j,k}(\lambda,\tau)$$

where the first component converges to $\gamma_0(j)$ under stationarity. All the terms involving the function $M_{t,k}(\lambda,\tau)$ are asymptotically negligible because the selected element is zero or $n^{-1}\cdot\hat{\eta}_{0k,p}^2(\lambda)$, with $\hat{\eta}_{0k,p}^2(\lambda)$ of order $O_p(1)$ or, at most, of order $c^2\cdot O_p(n^{1-2\alpha})$. If we write the terms involving products of residuals and $\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)$ as

$$\sum_{t=j+1}^{n}\hat{\eta}_{t,p}(\lambda)\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)=\sum_{t=1}^{n}\hat{\eta}_{t,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)-\sum_{t=1}^{j}\hat{\eta}_{t,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)$$
$$-\sum_{t=j+1}^{n}\hat{\eta}_{t,p}(\lambda)[\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)-\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)],$$
$$\sum_{t=j+1}^{n}\hat{\eta}_{t-j,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)=\sum_{t=1}^{n-j}\hat{\eta}_{t,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)$$
$$+\sum_{t=j+1}^{n}\hat{\eta}_{t-j,p}(\lambda)[\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)-\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)]$$

and

$$\sum_{t=j+1}^{n}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)=\sum_{t=j+1}^{n}\varphi_{t,p}^2(\lambda,\boldsymbol{\theta}_0)-\sum_{t=j+1}^{n}\varphi_{t,p}^2(\lambda,\boldsymbol{\theta}_0)$$
$$-\sum_{t=j+1}^{n}\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)[\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)-\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)]$$

they are dominated by the first term because the differences $\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)-\varphi_{t-j,p}(\lambda,\boldsymbol{\theta}_0)$ are zero or asymptotically negligible due to the fact that $\boldsymbol{x}'_p(\frac{t}{n}) - \boldsymbol{x}'_p(\frac{t-j}{n}) = \sum_{i=1}^{j}[\boldsymbol{x}'_p(\frac{t-i+1}{n}) - \boldsymbol{x}'_p(\frac{t-i}{n})]$ are of order $O(n^{-1})$. For $\phi_0 = 0$ all the terms are $O_p(1)$ due to the orthogonality between OLS residuals and the regressors, except $\sum_{t=1}^{n}\varphi_{t,p}^2(\lambda,\boldsymbol{\theta}_0) = 1 + O(n^{-1})$ so that

$$\gamma_n(j,\tau) = n^{-1}\sum_{t=j+1}^{n}\varepsilon_{0,t}\varepsilon_{0,t-j} + \delta_n^2(\alpha) + O_p(n^{-1/2}) \tag{B.31}$$

Thus, $\hat{\omega}_n^2(m_n) \to^p \omega_0^2(1 + \frac{c^2}{\omega_0^2} m_n \cdot O_p(n^{-2\alpha}) \cdot K)$, so that $\hat{\omega}_n^2(m_n)$ diverges at the rate $O_p(m_n \cdot n^{-2\alpha})$, but with $0 < c < 1$ the effect can be negligible except for very large outlier magnitudes, $-1/2 \leq \alpha < 0$, where $n^{-1} \cdot \hat{\omega}_n^{-2}(m_n) \cdot n^{-1}\hat{S}_{[nr],p}^2(\lambda,\tau) = O_p(m_n^{-1}) = o_p(1)$ in the case of wrong location.

For a persistent outlier ($\phi_0 \neq 0$), the quantity $\sum_{t=1}^{n}\varphi_{t,p}^2(\lambda,\boldsymbol{\theta}_0)$ is $O(n)$ by the same argument that for $\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0)$, and by the orthogonality between OLS residuals and the regressors

$$\sum_{t=1}^{n}\hat{\eta}_{t,p}(\lambda)\varphi_{t,p}(\lambda,\boldsymbol{\theta}_0) = \sqrt{n}\left(n^{-1/2}\sum_{t=1}^{n}\hat{\eta}_{t,p}(\lambda)\varphi_t(\boldsymbol{\theta}_0)\right) \tag{B.32}$$

where the term between parenthesis has a finite limit under Assumption 1, given by

$$n^{-1/2}\sum_{t=1}^{n}\hat{\eta}_{t,p}(\lambda)\varphi_t(\boldsymbol{\theta}_0) \Rightarrow \omega_0 \cdot J_{p,\gamma_0}(\boldsymbol{\theta}_0), \tag{B.33}$$

$$J_{p,\gamma_0}(\boldsymbol{\theta}_0) = J_{\gamma_0,0}(1,\boldsymbol{\theta}_0) - \left\{\boldsymbol{G}'_{p,\gamma_0}(\lambda,\boldsymbol{\theta}_0)\boldsymbol{Q}_\lambda^{-1}(p)\int_0^\lambda \boldsymbol{x}_p(s)dW_0(s) + \boldsymbol{G}'_{p,\gamma_0}(1,\boldsymbol{\theta}_0)\boldsymbol{Q}_{1-\lambda}^{-1}(p)\int_\lambda^1 \boldsymbol{x}_p(s)dW_0(s)\right\}$$

where $J_{\gamma_0,0}(1,\boldsymbol{\theta}_0)$ is defined in (3.11). Then, we have $\hat{\gamma}_n(j,\tau) = \gamma_0(j) + \delta_n(\alpha)\omega_0 O_p(1) + \delta_n^2(\alpha)O(n) + o_p(1)$, that is $\hat{\gamma}_n(j,\tau) = \gamma_0(j) + c^2 O_p(n^{1-2\alpha}) + o_p(1)$, and thus

$$\hat{\omega}_n^2(m_n) = \sum_{j=-(n-1)}^{n-1} w(j,m_n)\hat{\gamma}_n(j,\tau) \to^p \omega_0^2(1 + \frac{c^2}{\omega_0^2} m_n \cdot O_p(n^{1-2\alpha}) \cdot K)$$

For $\alpha = 0$, the sample aucovariance is dominated by the term $c^2 \cdot O_p(n)$ and we will have the same divergence rate as under the unit root alternative, $O_p(m_n \cdot n)$, so we will expect a slight increase in the empirical size of the tests, greater for higher values of ϕ_0. For $\alpha \to 1/2$, the leading term in the autocovariance will be $o_p(n)$ with the limit c^2 for $\alpha = 1/2$, thus making no significant effect.

■