Bias correction of KPSS test with structural break for reducing of size distortion*

Anton Skrobotov†
Gaidar Institute for Economic Policy,
The Russian Presidential Academy
of National Economy and Public Administration

November 15, 2012

Abstract

In this paper we extend the stationarity test proposed by Kurozumi and Tanaka (2010) for reducing size distortion with one structural break. We find the bias up to the order of $1/T$ for four types of models containing structural breaks. The simulations on finite samples show a reducing of size distortions in comparison with other tests, thus receiving higher power.

Key words: Stationarity tests, KPSS test, bias correction, size distortion, structural break.
JEL: C12, C22

1 Introduction

Unit root testing is a necessary element of data analysis. The standard and most widespread approach beginning from the work of Dicky and Fuller (Dickey and Fuller, 1979) is the hypothesis testing of unit root in a time series. But beginning from Kwiatkowski *et al.* (1992) (hereafter KPSS) the opposite direction of researches has had development with null hypothesis of stationarity about deterministic trend as a reversal complement of the unit root tests. The general and one of the basic problems in all of these tests is an assumption about the type of deterministic function. In works beginning by Perron (1989) it has been shown that usual unit root tests are inconsistent if the alternative hypothesis is that of stationary with structural break in the deterministic trend. A similar problem arises in stationarity tests, i.e. if a structural break occurs in the data there is serious size distortion in the usual KPSS test. Accordingly for solving this problem there were researches allowing breaks in stationarity tests.

Lee and Strazicich (2001) use an analogue of the KPSS test, considering two models: one model with a change in level for a non-trending series and another model with a change in both level and slope. In case of unknown break date the authors proposed the estimator of break date obtained at the value that minimizes the stationarity statistic. However it has been shown that the

*We thank Eiji Kurozumi for providing his GAUSS code and for his helpful comments, and Marina Turuntseva for her helping in preparing this article.

†E-mail: antonskrobotov@gmail.com

test has had low power because the minimizing of test statistics leads to the least favorable outcome against the alternative. Kurozumi (2002) considered the local asymptotic of the KPSS test allowing four types of structural changes. He obtained the limiting distributions of test statistics for all models, asymptotic local power function (using the approach stated in Tanaka (1996, chapter 9)) and investigated the power of tests depending on the location of the break. He also proposed another test in which limiting distribution didn't depend on the timing of the break. In case of unknown break date the author used an estimator of the break date obtained through minimizing the sum of squared residuals.

In Busetti and Harvey (2001) a similar problem of stationarity testing has been considered under a fixed alternative. The authors investigated KPSS tests for four types of models and received a corresponding limiting distribution and critical values. Note that Harvey and Mills (2003) found an error in final distribution for a model with a change in level though the proof was correct and critical values simulated using correct distribution. Harvey and Mills (2003) proposed modification of the standard KPSS statistic which limiting distribution didn't depend on the timing of the break and its generalization on a case of multiple structural changes.

Busetti and Harvey (2003) investigated different ways of estimating the break date if it is unknown. The test of Busetti and Harvey (2001) is based on the assumption of small breaks (the magnitude of the shift shrinks to zero as the sample size grows), and the break date is calculated using minimization of test statistics. Then there is a natural question how large the shift should be to satisfy this assumption because in the presence of large changes the rejection of the null hypothesis occurs too frequently. Another procedure proposed by Busetti and Harvey (2003) has good size properties but loses power under small changes. It consists of making a preliminary estimation of the break date by minimizing the sum of squared residuals, and then using the standard test statistics with the obtained estimate as the true date of break. Simulations have shown that minimization of test statistics too often rejects the null hypothesis of stationarity even for such small sizes of shift as one standard deviation of errors. Though for large shifts the empirical size is never above 0.17 the size of the test based on using a superconsistent estimate of break fraction as true is close to nominal at any size of break[1]. However the power of the test is too small for a small shift, but increases with its growth in magnitude, i.e. in a situation where it is easier to identify change. If the break actually doesn't occur, the infimum-test is more powerful than the test using a superconsistent estimate of break fraction obtained through minimization of the sum of squares residuals. Thus, if there is an uncertainty over the presence of a break, the infimum-test is preferable. On the other hand, if there is a confidence that a break exists, but its location isn't clear, it is preferable to use a two-step procedure in which the test statistic is calculated with the break date estimated through minimization of the sum of squares residuals.

Carrion-i-Silvestre and Sansó (2005) consider a possibility of two structural breaks in the KPSS test. They use an approach proposed by Sul *et al.* (2005) (hereafter SPC) for long run variance estimator. For each of seven models examined in the article the authors obtained the limiting distributions. A search of unknown break dates was made through minimization of the sum of squares residuals.

Another problem of KPSS type test has been considered, among others, in Carrion-i-Silvestre and Sansó (2006). They showed that on finite samples an SPC test with $AR(1)$ prewhitening was more preferable than others because it controls size. However the SPC test with $AR(1)$ prewhitening has serious size distortion when the data generating process (DGP) is $AR(2)$ (or

[1] This may not happen if the local asymptotic is considered.

higher order). Kurozumi and Tanaka (2010) (hereafter KT) extended the SPC approach (however with the absence of breaks) where for estimation of a long run variance they used autoregressive approximation using a boundary rule introduced by SPC for warranting of the consistency of long run variance. The authors also investigate the problem of downward bias in numerator of KPSS test statistics. For size correction they derived the finite sample bias and proposed the bias-corrected version of the KPSS test. Simulations showed that the empirical size of the modified test is well controlled in the case of $AR(2)$ errors, and the modified test also has higher power in comparison with SPC.

In this paper we propose the extension of the KT test on a case of a single structural break. Using modification of the boundary rule with autoregressive approximation as in KT we derive a finite sample bias of KPSS test statistics in a case of the presence of one break in DGP. If the break date is unknown it is possible to use its estimate obtained through minimization of the sum of squares residuals using the new approach of Harvey and Leybourne (2012). Simulations show the superiority of the modified test even in case of MA error.

The paper is organised as follows. In section 2 we discuss the model, test statistics, boundary rule of SPC with modification of long run variance proposed by KT. Also we derive parameter of bias for different types of models containing break. In section 3 we describe the possible estimators for unknown break date. In section 4 we detail the finite sample properties of the modified test. The results obtained are formulated in the Conclusion.

2 The Model

We consider the time series process $\{y\}$ generated according to the following model

$$y_t = d_t'\beta + u_t,\ t = 1, \dots, T, \tag{1}$$

where d_t is some deterministic component, and process u_t satisfy following standard assumption (see also Phillips and Solo (1992)).

Assumption 1 *Process u_t may be either I(0) or I(1):*

- *if $u_t \sim I(0)$, then it is linear process such that*

$$u_t = \gamma(L)e_t = \sum_{i=0}^{\infty} \gamma_i e_{t-i}$$

with $\gamma(z) \neq 0$ for all $|z| \leq 1$ u $\sum_{i=0}^{\infty} i|\gamma_i| < \infty$, where e_t − martingale difference sequence with conditional variance σ_e^2 and $\sup_t \mathbb{E}(e_t^4) < \infty$. Short run and long run variance are defined as $\sigma_u^2 = \mathbb{E}(u_t^2)$ and $\omega_u^2 = \lim_{T\to\infty} T^{-1}\mathbb{E}\left(\sum_{t=1}^{T} u_t\right)^2 = \sigma_e^2\gamma(1)^2$, respectively;

- *if $u_t \sim I(1)$, then it may be represented as $u_t = \sum_{j=1}^{t} e_t$, where $e_t \sim I(0)$.*

Also as in Perron (1989) (see also Perron (2006)) we consider four type of model: Model 0 (a change in level), respectively for a non-trending and for a trending series, Model I (a change in slope) and Model II (a change in both level and slope, mixed effect). Therefore deterministic component d_t can be written as:

$$d_t' = \begin{cases} (1, DU_t), & \text{for Model 0} \\ (1, t, DU_t), & \text{for Model 0t} \\ (1, t, DT_t), & \text{for Model I} \\ (1, t, DU_t, DT_t), & \text{for Model II} \end{cases},$$

where $DU_t = \mathbb{I}(t > T_1 + 1)$, $DT_t = (t - T_1)\mathbb{I}(t > T_1 + 1)$, $\mathbb{I}(\cdot)$ is the indicator function, T_1 is the break date. We define also break fraction as $\lambda_1 = T_1/T$.

We are testing null of stationarity ($u_t \sim I(0)$) against the alternative of a unit root ($u_t \sim I(1)$). It is usually used following KPSS test statistics for testing stationarity against unit root:

$$KPSS(\lambda_1) = \frac{T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{u}_s\right)^2}{\hat{\omega}_u^2} \quad (2)$$

where $\hat{u}_t = y_t - d_t'\hat{\beta}$ are OLS-residuals of y_t on d_t, where $d_t = [1, DU_t]'$, $\beta = (\mu_0, \mu_1)'$ for Model 0, $d_t = [1, t, DU_t]'$, $\beta = (\mu_0, \beta_0, \mu_1)'$ for Model 0t, $d_t = [1, t, DT_t]'$, $\beta = (\mu_0, \beta_0, \beta_1)'$ for Model I, $d_t = [1, t, DU_t, DT_t]'$, $\beta = (\mu_0, \beta_0, \mu_1, \beta_1)'$ for Model II, and long run variance estimator $\hat{\omega}_u^2$ is constructed according to the nonparametric approach using Bartlett or QS kernel.

Test statistics (2) have following limiting distribution derived by Busetti and Harvey (2001) (with correction of Harvey and Mills (2003)):

Lemma 1 *Under the null*

$$KPSS(\lambda_1) \Rightarrow \int_0^1 \left(W^*(r, \lambda_1)\right)^2 dr,$$

where for Model 0:

$$W^*(r, \lambda_1) = \begin{cases} W(r) - \frac{r}{\lambda_1}W(\lambda_1), & \text{for } r \le \lambda_1 \\ (W(r) - W(\lambda_1)) - \frac{r-\lambda_1}{1-\lambda_1}\left(W(1) - W(\lambda_1)\right), & \text{for } r > \lambda_1 \end{cases};$$

for Model 0t:

$$W^*(r, \lambda_1) = \begin{cases} W(r) - \frac{r}{\lambda_1}W(\lambda_1) - \frac{6r(r-\lambda_1)}{1-3\lambda_1+3\lambda_1^2} & \\ \times\left[\int_0^1 r dW(r) - \frac{\lambda_1}{2}W(\lambda_1) - \frac{1+\lambda_1}{2}(W(1) - W(\lambda_1))\right], & \text{for } r \le \lambda_1 \\ (W(r) - W(\lambda_1)) - \frac{r-\lambda_1}{1-\lambda_1}\left(W(1) - W(\lambda_1)\right) - \frac{6(r-1)(r-\lambda_1)}{1-3\lambda_1+3\lambda_1^2} & \\ \times\left[\int_0^1 r dW(r) - \frac{\lambda_1}{2}W(\lambda_1) - \frac{1+\lambda_1}{2}(W(1) - W(\lambda_1))\right], & \text{for } r > \lambda_1 \end{cases};$$

for Model I:

$$W^*(r,\lambda_1)=\begin{cases} W(r)-rW(1)-\frac{3}{\lambda_1^3(1-\lambda_1)^3} \\ \times\left[\left(a\frac{r^2}{2}-a\lambda_1 r+\frac{r}{2}(a\lambda_1^2-b(1-\lambda_1)^2)\right)J_1\right. \\ \left.+\left(b\frac{r^2}{2}-b\lambda_1 r+\frac{r}{2}(b\lambda_1^2-c(1-\lambda_1)^2)\right)J_2\right], & \textit{for } r\le\lambda_1 \\ W(r)-rW(1)-\frac{3}{\lambda_1^3(1-\lambda_1)^3} \\ \times\left[\left(-a\frac{\lambda_1^2}{2}+b\frac{r^2-\lambda_1^2}{2}-b\lambda_1(r-\lambda_1)+\frac{r}{2}(a\lambda_1^2-b(1-\lambda_1)^2)\right)J_1\right. \\ \left.+\left(-b\frac{\lambda_1^2}{2}+c\frac{r^2-\lambda_1^2}{2}-c\lambda_1(r-\lambda_1)+\frac{r}{2}(b\lambda_1^2-c(1-\lambda_1)^2)\right)J_2\right], & \textit{for } r>\lambda_1 \end{cases};$$

for Model II:

$$W^*(r,\lambda_1)=\begin{cases} W(r)-\frac{r}{\lambda_1}W(\lambda_1)-\frac{6r(r-\lambda_1)}{\lambda_1^3} \\ \times\left[\int_0^{\lambda_1} rdW(r)-\frac{\lambda_1}{2}W(\lambda_1)\right], & \textit{for } r\le\lambda_1 \\ (W(r)-W(\lambda_1))-\frac{r-\lambda_1}{1-\lambda_1}(W(1)-W(\lambda_1))-\frac{6(r-1)(r-\lambda_1)}{(1-\lambda_1)^3} \\ \times\left[\int_{\lambda_1}^1 rdW(r)-\frac{1+\lambda_1}{2}(W(1)-W(\lambda_1))\right], & \textit{for } r>\lambda_1 \end{cases};$$

Here $a=(1-\lambda_1)^3(1+\lambda_1)$, $b=-3\lambda_1^2(1-\lambda_1)^2$, $c=\lambda_1^3(4-3\lambda_1)$, $J_1=\int_0^{\lambda_1} rdW(r)-\lambda_1 W(\lambda_1)+\frac{\lambda_1^2}{2}W(1)$ *and* $J_1=\int_{\lambda_1}^1 rdW(r)-\lambda_1(W(1)-W(\lambda_1))-\frac{(1-\lambda_1)^2}{2}W(1)$.

The KPSS test can control the size of the test asymptotically but in finite samples it has serious size distortion. For their reducing SPC proposed $AR(1)$ prewhitening method with a boundary rule. I.e. we first estimate $AR(p)$ model for residuals of regression (1), $\hat{u}_t$:

$$\hat{u}_t=\varphi_1\hat{u}_{t-1}+\cdots+\varphi_p\hat{u}_{t-p}+e_t.$$

Then the long run variance estimator is constructed as

$$\hat{\omega}_u^2=\frac{\hat{\omega}_e^2}{(1-\tilde{\varphi})^2}, \tag{3}$$

where $\tilde{\phi}=\min\left\{\sum_{j=1}^p\hat{\phi}_j, 1-1/\sqrt{T}\right\}$ and $\hat{\omega}_e$ is the long run variance estimator of residuals $\hat{e}_t$[2]. KT proposed the modification of SPC for the case of autoregerssive long run variance estimator ω_u^2, i.e.:

$$\hat{\omega}_{u,AR}^2=\frac{\hat{\sigma}_e^2}{(1-\tilde{\phi})^2}, \tag{4}$$

where $\hat{\sigma}_e=T^{-1}\sum_{t=1}^T\hat{e}_t^2$ and $\tilde{\phi}=\min\left\{\sum_{j=1}^p\hat{\phi}_j, 1-c/\sqrt{T}\right\}$ with c is some finite constant. However while autoregressive estimator of long run variance is applied well enough there is a problem of downward bias of test statistics (2) in finite samples. KT showed for the cases of constant and

[2]For $AR(1)$ error Carrion-i-Silvestre and Sansó (2006) showed that the size of SPC test with $AR(1)$ prewhitening is close to the nominal and more preferably other tests when true DGP is $AR(1)$ process while it has liberal size distortion if true DGP is $AR(2)$ process.

trend that their modification of long run variance estimator still leads too rare null rejection due to this downwards bias. This leads to essential power losses under the alternative hypothesis. For prevention bias in the numerator of KPSS test statistics in finite samples KT proposed its bias-corrected version:

$$KPSS = \frac{T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{u}_s\right)^2 - \hat{b}_T}{\hat{\omega}^2_{u,AR}}. \tag{5}$$

Here the term b_T is responsible for bias. To calculate its value authors suggest to use Beveridge-Nelson decomposition. Let $u_t = \psi(L)e_t$, then process u_t can be written as

$$u_t = \psi(1)e_t + \upsilon_{t-1} - \upsilon_t,$$

where $\upsilon_t = \sum_{j=0}^{\infty}\tilde{\psi}_j e_{t-j}$, $\tilde{\psi}_j = \sum_{i=j+1}^{\infty}\psi_i$. The residuals $\hat{u}_t$ is defined as

$$\begin{aligned}
\hat{u}_t &= u_t - d_t'\left(\sum_{t=1}^{T} d_t d_t'\right)^{-1}\sum_{t=1}^{T} d_t u_t \\
&= \psi(1)e_t + \upsilon_{t-1} - \upsilon_t - d_t'\left(\sum_{t=1}^{T} d_t d_t'\right)^{-1}\sum_{t=1}^{T} d_t(\psi(1)e_t + \upsilon_{t-1} - \upsilon_t) \\
&= \psi(1)\hat{e}_t - \widehat{\Delta\upsilon_t},
\end{aligned}$$

where $\hat{e}_t$ and $\widehat{\Delta\upsilon_t}$ are the residuals of regression of e_t and $\Delta\upsilon_t$ on d_t, respectively. Then KT decomposed the numerator of (2) into three terms:

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{u}_s\right)^2 &= \frac{\psi^2(1)}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{e}_s\right)^2 \\
&\quad + \frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\widehat{\Delta\upsilon_s}\right)^2 - \frac{2\psi(1)}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{e}_s\right)\left(\sum_{s=1}^{t}\widehat{\Delta\upsilon_s}\right) \\
&= \frac{\psi^2(1)}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{e}_s\right)^2 + R_1 - R_2.
\end{aligned}$$

The second and third terms are $o_p(1)$ while the first term have nondegenerate limiting distribution. Thus bias of numerator depends on R_1 and R_2. KT determines this bias as expectation of $R_1 - R_2$ up to $O(T^{-1})$. This bias is defined as b_T:

$$\mathbb{E}[R_1 - R_2] = b_T + o(T^{-1}). \tag{6}$$

Proposition 1 *Let $\gamma_0 = \mathbb{E}[\upsilon_t^2]$, lag polynomial $\varphi(L) = c(L)(1-\rho L)$ and $\varphi'(1) = d\varphi(z)/dz|_{z=1}$. Then bias b_T in the numerator of KPSS test statistic* (5) *is expressed as*

$$b_T = \frac{b_0}{T}\left(\gamma_0 + \sigma_e^2\frac{\varphi'(1)}{\varphi^3(1)}\right) \tag{7}$$

where $b_0 = 5/3$ *for Model 0,* $b_0 = \frac{285\lambda_1^4 - 576\lambda_1^3 + 498\lambda_1^2 - 213\lambda_1 + 38}{30(1-3\lambda_1+3\lambda_1^2)^2}$ *for Model 0t,* $b_0 = 7/6$ *for Model I u* $b_0 = 19/15$ *for Model II*[3].

Remark Note that for Model 0, I and II the bias doesn't depend on break fraction λ. But this is not for Model 0t.

Parameter γ_0 can be constructed recursively solving Yule-Walker equations (see for details in Kurozumi and Tanaka (2010, section 3.2)).

3 The case with unknown break date

The test considered in previous sections is based on assumptions that the timing of the structural breaks is known a priori. However in many cases it cannot be known to the researcher. Then it is possible to replace a known break date with its consistent estimate. Then the limiting distribution of the test statistics remains the same. The consistent estimator of the break fraction $\lambda_1 = T_1/T$ can be obtained through minimizing the sum of squared residuals in the model over all possible break dates. It is possible to show that this estimator is superconsistent under both the $I(0)$ and $I(1)$ (see, e.g., Perron and Zhu (2005)).

The alternative procedure of searching an unknown break date proposed by Carrion-i Silvestre *et al.* (2009), using preliminary (quasi) GLS-detrending of y_t. I.e., let us estimate the following regression:

$$y_t = X_t(\lambda_1)\phi + u_t, \tag{8}$$

where $X_t(\lambda_1)$ includes all regressors, and ϕ is a corresponding parameters. Then GLS-estimate $\hat{\phi}$ of vector ϕ is the OLS-estimate of coefficient vector in equation

$$y_t^{\bar{\rho}} = X_t^{\bar{\rho}}(\lambda_1)\phi + u_t^{\bar{\rho}}, \tag{9}$$

where

$$y_t^{\bar{\rho}} = [y_1, (1-\bar{\rho}L)y_2, \ldots, (1-\bar{\rho}L)y_T]',$$
$$X_t^{\bar{\rho}}(\lambda_1) = [x_1, (1-\bar{\rho}L)x_2, \ldots, (1-\bar{\rho}L)x_T]'.$$

Carrion-i Silvestre *et al.* (2009) suggest to choose $\bar{\rho} = 1 + \bar{c}/T$ depending on timing of break, because we do not know a priory whether the series u_t is $I(0)$ or $I(1)$. Then the estimator of break date is:

$$\hat{\lambda}_1^{\bar{\rho}} = \arg\min_{\lambda_1 \in \Lambda(e)} S(\bar{\rho}, \lambda_1), \tag{10}$$

where $S(\bar{\rho}, \lambda_1)$ is the sum of squared residuals in regression (9). The obtained estimate of the break fraction will be superconsistent under both the $I(0)$ and $I(1)$ cases.

Harvey and Leybourne (2012) proposed the modification of this break date estimator using additional information if the series is $I(1)$. In this case estimate of break fraction obtaining for GLS detrended series will be also consistent but the estimate obtaining for the first differenced series will be efficient. Authors propose to use a hybrid estimator:

[3] KT showed that in case of constant only $b_0 = 5/3$ and in case of constant and trend $b_0 = 19/15$.

$$\hat{\lambda}_1^{D_m} = \arg \min_{\lambda_1 \in \Lambda(e), \bar{\rho} \in D_m} S(\bar{\rho}, \lambda_1), \tag{11}$$

where $D_m = \{\rho'_1, \rho'_2, \ldots, \rho'_{m-1}, 1\}$ is the m element set, where $|\rho'_i| < 1$ for all i and, without loss of generality, $-1 < \rho'_1 < \rho'_2 < \cdots < \rho'_{m-1} < 1$.

Asymptotic results show that it is necessary to set ρ'_{m-1} close enough to unit that the estimate (11) had demanded asymptotic properties (if the true value of $\rho > \rho'_{m-1}$, then break fraction estimator will be inefficient). As set D_m authors suggest to use $D_m = \{0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.975, 1\}$. The reason is that a negative serial correlation usually is not observed in practice, and also the serial correlation often can be very strongly positive, therefore inclusion of value 0.975 allows a small enough interval $0.975 < \rho < 1$ for adequate asymptotic choice.

We will use this estimator further in the next section. Note that as an estimate of break fraction is consistent it is possible to use the same critical values (obtained by Busetti and Harvey (2001) or Kurozumi (2002)) as in known break date.

4 Finite sample properties

In this section we investigate the finite sample behavior of the tests. Consider following DGP:

$$y_t = d'_t\beta + u_t,\ u_t = \alpha u_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}, \tag{12}$$

i.e. we allow MA component in DGP. Also $\varepsilon_t \sim i.i.d.N(0, 1)$. We set the values of parameter α from 0.5 to 1, values of parameter θ in MA component $\{-0.8, -0.4, 0.0, 0.4, 0.8\}$, break fraction $\lambda_1 = 0.5$. The parameters of deterministic component $\mu_0 = 0$ without loss of generality, $\mu_1 = -4$, $\beta_0 = 0.3$, $\beta_1 = -0.1$. The significance level is 0.05, the number of replications is 5,000. The initial value u_0 is set to 0 for simplicity.

We consider the behavior of four tests. The first is a bias-corrected KPSS test with correction factor, obtained under structural change in Section 2 (it is marked in graphs as BC). The second is SPC test with $AR(1)$ prewhitening (it is marked as SPC). Also for comparison we consider a KPSS test without correcting factor $\hat{b}_T$ (it is marked as NC), and a KPSS test with break with proposed by Kurozumi (2002) bandwidth:

$$l_A k = \min\left\{1.447\left(\frac{\hat{\alpha}^2 T}{(1+\hat{\alpha})^2(1-\hat{\alpha})^2}\right)^{1/3}, 1.447\left(\frac{4k^2 T}{(1+k)^2(1-k)^2}\right)^{1/3}\right\}.$$

with $k = 0.8$ (it is marked as K). In all cases the lag length is selected using BIC.

The Figure 1 shows the size and power of the considered tests for Model II (the results for other models are similar and omitted for brevity) for known break date ($\lambda_1 = 0.5$) and boundary value equal to $1 - c/\sqrt{T} = 0.8$. This value suppose for α is greater than 0.8 the null hypothesis will tend to be rejected more often if $\theta = 0$. If $\theta \neq 0$ then the measure of the strength of persistence should be based on the autoregressive representation of the series. In our $ARMA(1, 1)$ rewrite DGP for u_t as

$$(1 - \alpha L)(1 - \theta L)^{-1} u_t = \varepsilon_t$$

Then the measure of the strength of persistence $\varphi_1 + \cdots + \varphi_p$ should be $(\alpha - \theta)/(1 - \theta)$. We expect that the null hypothesis will be rejected more often than the nominal size when this measure is greater than 0.8. That is in Figure 1 the bias-corrected KPSS test should control the size for

$\alpha < 0.64$ and over-reject the null otherwise for $\theta = -0.8$. For $\theta = -0.4$, $\theta = 0.4$ and $\theta = 0.8$ the test should control the size for $\alpha < 0.72$, $\alpha < 0.88$ and $\alpha < 0.96$, respectively. For $\theta = -0.8$, $\theta = -0.4$ and $\theta = 0$ this is confirmed by results in Figure 1 while for $\theta = 0.4$ and $\theta = 0.8$ the test begins to reject the null more often for smaller α, because the boundary value is large.

In cases $\theta = -0.8$ and $\theta = -0.4$ the obtained bias-corrected test has a size closer to nominal than in other tests. For $\theta = -0.8$ any of tests has no power above than of bias-corrected test. On the other hand for $\theta = -0.4$ case there are some size distortion for $\alpha > 0.7$. This leads to higher power in comparison with a test with no correction. In simple $AR(1)$ case the results are similar to, i.e. Carrion-i-Silvestre and Sansó (2005) and KT. Here the bias-corrected test obviously surpasses the others in size and power. For negative MA component the test proposed by Kurozumi (2002) has higher power simultaneously with higher size (for all values of θ this test is characterized by oversizing). Though the bias-corrected test has lower power it controls the size well.

The results for other values of boundary value are similar and omitted. It can be noticed that for higher values the superiority of the bias-corrected test is more appreciable though it has lower power (for other tests the size distortion increases), and for lower values higher power is observed which is compensated by size distortions after boundary value.

The figure 2 shows size and power for unknown break date estimated by Harvey and Leybourne (2012) procedure. Apparently results practically have not changed and are qualitatively similar.

5 Conclusion

In this paper we considered the generalization of the test proposed by Kurozumi and Tanaka (2010) on the case of a single structural break. Using boundary rule in SPC with $AR(1)$ prewhitening we found the finite sample bias of numerator in case of structural change occurring in DGP. The results are similar to the case of absence of break. In simulation analysis the case of unknown break date has been considered using the approach proposed by Harvey and Leybourne (2012). The finite sample results show the superiority of the obtained modification in the presence of break because the test controls the size better. Therefore use of the bias-corrected KPSS test should be useful in empirical applications.

A Appendix

Proof of Theorem 1[4]. Consider the most general case with $d_t' = (1, t, DU_t, DT_t)$. As in KT, we decompose R_1 into three term:

$$R_1 = R_{11} - R_{12} + R_{13},$$

where

[4] Some matrix calculations (multiplication, inversion) are carried out using Wolfram Mathematica 8.

$$\begin{aligned}
R_{11} &= \frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\Delta\upsilon_s\right)^2 \\
R_{12} &= = -\frac{2}{T^2}\left(\sum_{t=1}^{T}\sum_{s=1}^{t}\Delta\upsilon_s\sum_{s=1}^{t}d_s'\right)\left(\sum_{t=1}^{T}d_t d_t'\right)^{-1}\sum_{t=1}^{T}d_t\Delta\upsilon_t \\
R_{13} &= \frac{1}{T^2}\sum_{t=1}^{T}\Delta\upsilon_t d_t'\left(\sum_{t=1}^{T}d_t d_t'\right)^{-1}\left(\sum_{t=1}^{T}\sum_{s=1}^{t}d_s\sum_{s=1}^{t}d_s'\right)\left(\sum_{t=1}^{T}d_t d_t'\right)^{-1}\sum_{t=1}^{T}d_t\Delta\upsilon_t.
\end{aligned}$$

Notice that $\sum_{s=1}^{t}\Delta\upsilon_s = \upsilon_t - \upsilon_0$, then (see KT):

$$\mathbb{E}[R_{11}] = \frac{2}{T}\gamma_0 + O(T^{-2}). \tag{13}$$

The second term is expressed as:

$$\mathbb{E}[R_{12}] = \frac{2}{T^2}tr\left\{\left(\sum_{t=1}^{T}d_t d_t'\right)^{-1}\mathbb{E}\left[\left(\sum_{t=1}^{T}d_t\Delta\upsilon_t\right)\left(\sum_{t=1}^{T}\sum_{s=1}^{t}\Delta\upsilon_s\sum_{s=1}^{t}d_s'\right)\right]\right\} \tag{14}$$

For proof we use the following results:

$$\sum_{t=1}^{T}d_t\Delta\upsilon_t = \begin{bmatrix}\sum_{t=1}^{T}\Delta\upsilon_t \\ \sum_{t=1}^{T}t\Delta\upsilon_t \\ \sum_{t=1}^{T}DU_t\Delta\upsilon_t \\ \sum_{t=1}^{T}DT_t\Delta\upsilon_t\end{bmatrix} = \begin{pmatrix}\upsilon_T - \upsilon_0 \\ (T+1)\upsilon_T - \upsilon_0 - \sum_{t=1}^{T}\upsilon_t \\ \upsilon_T - \upsilon_{T_1} \\ (T+1-T_1)\upsilon_T - \upsilon_{T_1} - \sum_{t=T_1+1}^{T}\upsilon_t\end{pmatrix}. \tag{15}$$

$$\left(\sum_{t=1}^{T}\sum_{s=1}^{t}\Delta\upsilon_s\sum_{s=1}^{t}d_s'\right) = \begin{pmatrix}\sum_{t=1}^{T}t(\upsilon_t-\upsilon_0) \\ \sum_{t=1}^{T}(\upsilon_t-\upsilon_0)\sum_{s=1}^{t}s \\ \sum_{t=1}^{T}(\upsilon_t-\upsilon_0)\sum_{s=1}^{t}DU_s \\ \sum_{t=1}^{T}(\upsilon_t-\upsilon_0)\sum_{s=1}^{t}DT_s\end{pmatrix}' \tag{16}$$

$$\sum_{t=1}^{T}\sum_{s=1}^{t}DU_s = \frac{T^2(1-\lambda_1)^2}{2} + o(T^2) \tag{17}$$

$$\sum_{t=1}^{T}\sum_{s=1}^{t}DT_s = \frac{T^3(1-\lambda_1)^3}{3} + o(T^3) \tag{18}$$

$$\sum_{t=1}^{T}tDU_t = \frac{(T+T_1+1)(T-T_1)}{2} \tag{19}$$

$$\sum_{t=1}^{T}tDT_t = \frac{(T-T_1)(T-T_1+1)(2T+T_1+1)}{6} \tag{20}$$

$$\sum_{t=1}^{T}DT_tDT_t = \frac{(T-T_1)(T-T_1+1)(2T-2T_1+1)}{6} \tag{21}$$

Using (19)-(21) it can be shown that

$$\sum_{t=1}^{T} d_t d_t' =$$

$$\begin{pmatrix} T & \frac{T(T+1)}{2} & T-T_1 & \frac{(T-T_1)(T-T_1+1)}{2} \\ \frac{T(T+1)}{2} & \frac{T(T+1)(2T+1)}{6} & \frac{(T-T_1)(T+T_1+1)}{2} & \frac{(T-T_1)(T-T_1+1)(2T+T_1+1)}{6} \\ T-T_1 & \frac{(T-T_1)(T+T_1+1)}{2} & T-T_1 & \frac{(T-T_1)(T-T_1+1)}{2} \\ \frac{(T-T_1)(T-T_1+1)}{2} & \frac{(T-T_1)(T-T_1+1)(2T+T_1+1)}{6} & \frac{(T-T_1)(T-T_1+1)}{2} & \frac{(2T-2T_1+1)(T-T_1)(T-T_1+1)}{6} \end{pmatrix}. \quad (22)$$

Then

$$\left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} =$$

$$\begin{pmatrix} \frac{2(2T_1+1)}{(T_1-1)T_1} & -\frac{6}{(T_1-1)T_1} & \frac{2}{T_1} & \frac{6}{(T_1-1)T_1} \\ -\frac{6}{(T_1-1)T_1} & \frac{12}{(T_1-1)T_1(T_1+1)} & -\frac{6}{T_1(T_1+1)} & -\frac{12}{(T_1-1)T_1(T_1+1)} \\ \frac{2}{T_1} & -\frac{6}{T_1(T_1+1)} & \frac{2T\left(2TT_1-T-2T_1^2+2T_1+1\right)}{T_1(T_1+1)(T-T_1-1)(T-T_1)} & \frac{6T(T-2T_1-1)}{T_1(T_1+1)(T-T_1-1)(T-T_1)} \\ \frac{6}{(T_1-1)T_1} & -\frac{12}{(T_1-1)T_1(T_1+1)} & \frac{6T(T-2T_1-1)}{T_1(T_1+1)(T-T_1-1)(T-T_1)} & \frac{12T\left(T^2-3TT_1+3T_1^2-1\right)}{(T_1-1)T_1(T_1+1)(T-T_1-1)(T-T_1)(T-T_1+1)} \end{pmatrix}. \quad (23)$$

Consider the expectation of right hand side of (14) using (15) and (16). The expectation of (1,1) element equal to (see KT):

$$\mathbb{E}[(1,1) \text{ element}] = \frac{T^2}{2}\gamma_0 + o(T^2), \quad (24)$$

Similarly for the (1,2) element (see KT):

$$\mathbb{E}[(1,2) \text{ element}] = \frac{T^3}{6}\gamma_0 + o(T^3), \quad (25)$$

Consider (1,3) element.

$$\begin{aligned} \mathbb{E}[(1,3) \text{ element}] &= \mathbb{E}\left[(\upsilon_T - \upsilon_0)\sum_{t=1}^{T}(\upsilon_t - \upsilon_0)\sum_{s=1}^{t} DU_s\right] \quad (26) \\ &= \sum_{t=1}^{T}\gamma_{T-t}\sum_{s=1}^{t} DU_s - \gamma_T\sum_{t=1}^{T}\sum_{s=1}^{t} DU_t - \sum_{t=1}^{T}\gamma_t\sum_{s=1}^{t} DU_t + \gamma_0\sum_{t=1}^{T}\sum_{s=1}^{t} DU_t \\ &= \frac{T^2(1-\lambda_1)^2)}{2}\gamma_0 + o(T^2), \end{aligned}$$

as γ_t absolutely summable and $\gamma_T = o(1)$, and also using the equation (17). Similarly expectation of (1,4) element (we use (18)):

$$\mathbb{E}[(1,4) \text{ element}] = \frac{T^3(1-\lambda_1)^3)}{6}\gamma_0 + o(T^3). \quad (27)$$

As in KT (2,1) and (2,2) elements are $o(T^3)$ and $o(T^4)$, respectively. Clearly that (2,3) and (2,4) elements have the same orders $o(T^3)$ and $o(T^4)$, respectively, because DU_t and DT_t have the same orders as constant and trend, respectively.

Elements of third row of matrix differ that in corresponding expressions γ_0 replaces on γ_{T_1}, but as $T_1 > p$ then $\gamma_{T_1} = o(1)$. Then elements of third row are $o(T^2)$, $o(T^3)$, $o(T^2)$, $o(T^3)$. Elements of fourth row equal to corresponding elements of second row as they have the same order. Hence we can write the expectation as

$$\mathbb{E}\left[\left(\sum_{t=1}^{T} d_t \Delta\upsilon_t\right)\left(\sum_{t=1}^{T}\sum_{s=1}^{t}\Delta\upsilon_s \sum_{s=1}^{t} d'_s\right)\right] = \begin{pmatrix} \frac{T^2}{2}\gamma_0 + o(T^2) & \frac{T^3}{6}\gamma_0 + o(T^3) & \frac{T^2(1-\lambda_1)^2}{2}\gamma_0 + o(T^2) & \frac{T^3(1-\lambda_1)^3}{6}\gamma_0 + o(T^3) \\ o(T^3) & o(T^4) & o(T^3) & o(T^4) \\ o(T^2) & o(T^3) & o(T^2) & o(T^3) \\ o(T^3) & o(T^4) & o(T^3) & o(T^4) \end{pmatrix} \quad (28)$$

Then using (23) и (28) we obtain:

$$\mathbb{E}[R_{12}] = 2\frac{-\lambda_1^2\gamma_0 + 3\lambda_1^3\gamma_0 - 3\lambda_1^4\gamma_0 + \lambda_1^5\gamma_0}{T(\lambda_1 - 1)^3\lambda_1^2} + o(T^{-1}) = \frac{2}{T}\gamma_0 + o(T^{-1}). \quad (29)$$

Consider the third term of R_{13}:

$$\mathbb{E}[R_{13}] = \frac{1}{T^2} tr\left\{\left(\sum_{t=1}^{T} d_t d'_t\right)^{-1}\left(\sum_{t=1}^{T}\sum_{s=1}^{t} d_s \sum_{s=1}^{t} d'_s\right)\left(\sum_{t=1}^{T} d_t d'_t\right)^{-1} \mathbb{E}\left[\sum_{t=1}^{T} d_t \Delta\upsilon_t \sum_{t=1}^{T}\Delta\upsilon_t d'_t\right]\right\}. \quad (30)$$

It can be shown using (15) and the fact that γ_{T_1} and γ_{T-T_1} are $o(1)$ that

$$\mathbb{E}\left[\sum_{t=1}^{T} d_t \Delta\upsilon_t \sum_{t=1}^{T}\Delta\upsilon_t d'_t\right] = \begin{pmatrix} 2\gamma_0 + o(1) & T\gamma_0 + o(T) & \gamma_0 + o(1) & T(1-\lambda_1)\gamma_0 + o(T) \\ T\gamma_0 + o(T) & T^2\gamma_0 + o(T^2) & T\gamma_0 + o(T) & T^2(1-\lambda_1)\gamma_0 + o(T^2) \\ \gamma_0 + o(1) & T\gamma_0 + o(T) & 2\gamma_0 + o(1) & T(1-\lambda_1)\gamma_0 + o(T) \\ T(1-\lambda_1)\gamma_0 + o(T) & T^2(1-\lambda_1)\gamma_0 + o(T^2) & T(1-\lambda_1)\gamma_0 + o(T) & T^2(1-\lambda_1)^2\gamma_0 + o(T^2) \end{pmatrix} \quad (31)$$

Consider the component $DD := \left(\sum_{t=1}^{T} d_t d'_t\right)^{-1}\left(\sum_{t=1}^{T}\sum_{s=1}^{t} d_s \sum_{s=1}^{t} d'_s\right)\left(\sum_{t=1}^{T} d_t d'_t\right)^{-1}$. We can show that

$$\left(\sum_{t=1}^{T}\sum_{s=1}^{t} d_s \sum_{s=1}^{t} d'_s\right) = (\xi_1, \xi_2, \xi_3, \xi_4),$$

where

$$\xi_1 = \begin{pmatrix} \frac{T(T+1)(2T+1)}{6} \\ \frac{T(T+1)(T+2)(3T+1)}{24} \\ \frac{(T-T_1)(T-T_1+1)(2T+T_1+1)}{6} \\ \frac{(T-T_1)(T-T_1+1)(T-T_1+2)(3T+T_1+1)}{24} \end{pmatrix},$$

$$\xi_2 = \begin{pmatrix} \frac{T(T+1)(T+2)(3T+1)}{24} \\ \frac{T(T+1)(T+2)\left(3T^2+6T+1\right)}{60} \\ \frac{(T-T_1)(T-T_1+1)\left(3T^2+2TT_1+7T+T_1^2+3T_1+2\right)}{24} \\ \frac{(T-T_1)(T-T_1+1)(T-T_1+2)\left(6T^2+3TT_1+12T+T_1^2+3T_1+2\right)}{120} \end{pmatrix},$$

$$\xi_3 = \begin{pmatrix} \frac{(T-T_1)(T-T_1+1)(2T+T_1+1)}{6} \\ \frac{(T-T_1)(T-T_1+1)\left(3T^2+2TT_1+7T+T_1^2+3T_1+2\right)}{24} \\ \frac{(2T-2T_1+1)(T-T_1)(T-T_1+1)}{6} \\ \frac{(3T-3T_1+1)(T-T_1)(T-T_1+1)(T-T_1+2)}{24} \end{pmatrix},$$

$$\xi_4 = \begin{pmatrix} \frac{(T-T_1)(T-T_1+1)(T-T_1+2)(3T+T_1+1)}{24} \\ \frac{(T-T_1)(T-T_1+1)(T-T_1+2)\left(6T^2+3TT_1+12T+T_1^2+3T_1+2\right)}{120} \\ \frac{(3T-3T_1+1)(T-T_1)(T-T_1+1)(T-T_1+2)}{24} \\ \frac{(T-T_1)(T-T_1+1)(T-T_1+2)\left(3T^2-6TT_1+6T+3T_1^2-6T_1+1\right)}{60} \end{pmatrix},$$

Thus using (23),

$$DD = (\kappa_1, \kappa_2, \kappa_3, \kappa_4), \tag{32}$$

where

$$\kappa_1 = \begin{pmatrix} \frac{15TT_1^2-15TT_1+2T_1^3+22T_1^2-8T_1+2}{15(T_1-1)T_1} \\ -\frac{11T_1^2-5T_1+6}{10(T_1-1)T_1} \\ -\frac{(T_1-2)(T_1+2)}{30T_1} \\ \frac{(T_1+2)(T_1+3)}{10(T_1-1)T_1} \end{pmatrix},$$

$$\kappa_2 = \begin{pmatrix} -\frac{11T_1^2-5T_1+6}{10(T_1-1)T_1} \\ \frac{6\left(T_1^2+1\right)}{5(T_1-1)T_1(T_1+1)} \\ -\frac{(T_1-3)(T_1-2)}{10T_1(T_1+1)} \\ -\frac{6\left(T_1^2+1\right)}{5(T_1-1)T_1(T_1+1)} \end{pmatrix},$$

$$\kappa_3 = \begin{pmatrix} -\frac{(T_1-2)(T_1+2)}{30T_1} \\ -\frac{(T_1-3)(T_1-2)}{10T_1(T_1+1)} \\ \frac{2T\left(T^2T_1^2+T^2T_1-2TT_1^3-3TT_1^2+7TT_1-T+T_1^4+2T_1^3-7T_1^2+2T_1+1\right)}{15T_1(T_1+1)(T-T_1-1)(T-T_1)} \\ -\frac{3T\left(TT_1-T-T_1^2+2T_1+1\right)}{5T_1(T_1+1)(T-T_1-1)(T-T_1)} \end{pmatrix},$$

$$\kappa_4 = \begin{pmatrix} \frac{(T_1+2)(T_1+3)}{10(T_1-1)T_1} \\ -\frac{6\left(T_1^2+1\right)}{5(T_1-1)T_1(T_1+1)} \\ -\frac{3T\left(TT_1-T-T_1^2+2T_1+1\right)}{5T_1(T_1+1)(T-T_1-1)(T-T_1)} \\ \frac{6T\left(T^2T_1^2+T^2-2TT_1^3-4TT_1+T_1^4+4T_1^2-1\right)}{5(T_1-1)T_1(T_1+1)(T-T_1-1)(T-T_1)(T-T_1+1)} \end{pmatrix},$$

We then obtain that

$$\mathbb{E}[R_{13}] = \frac{19}{15T}\gamma_0 + o(T^{-1}); \tag{33}$$

Combining (13), (29) and (33) we obtain

$$\mathbb{E}[R_1] = \frac{19}{15T}\gamma_0 + o(T^{-1}). \tag{34}$$

We now consider the expectation of R_2. We decompose R_2 into four terms (see KT) except for the scalar $2\psi(1)/T^2$:

$$R_2 = R_{21} - R_{22} - R_{23} + R_{24}, \tag{35}$$

where

$$\begin{aligned} R_{21} &= \sum_{t=1}^{t}\sum_{s=1}^{T} e_s \sum_{s=1}^{t} \Delta v_s \\ R_{22} &= \sum_{t=1}^{T}\sum_{s=1}^{t} \Delta v_s \sum_{s=1}^{t} d_s' \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \sum_{t=1}^{T} d_t e_t \\ R_{23} &= \sum_{t=1}^{T}\sum_{s=1}^{t} e_s \sum_{s=1}^{t} d_s' \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \sum_{t=1}^{T} d_t \Delta v_t \\ R_{24} &= \sum_{t=1}^{T} e_t d_t' \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \left(\sum_{t=1}^{T}\sum_{s=1}^{t} d_s \sum_{s=1}^{t} d_s'\right) \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \sum_{t=1}^{T} d_t \Delta v_t \end{aligned}$$

The expectation of the first term is (see KT):

$$\mathbb{E}[R_{21}] = \sigma_e^2 T \sum_{t=0}^{T-1} \left(1 - \frac{t}{T}\right) \tilde{\psi}_t \tag{36}$$

Consider the second term:

$$\mathbb{E}[R_{22}] = tr\left\{ \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \mathbb{E}\left[\sum_{t=1}^{T} d_t e_t \sum_{t=1}^{T}\sum_{s=1}^{t} \Delta v_s \sum_{s=1}^{t} d_s'\right] \right\}. \tag{37}$$

Next we use Lemma provided in KT with some generalizations for dummy variables:

Lemma 2 *Let f_t and g_t be deterministic sequences for $t = 1, \dots, T$. Then*

$$E\left[\left(\sum_{t=1}^{T} f_t e_t\right)\left(\sum_{t=1}^{T} g_t \upsilon_t\right)\right] = \sigma_e^2 \sum_{t=0}^{T-1}\left(\sum_{s=1}^{T-t} f_s g_{s+t}\right)\tilde{\psi}_t, \tag{38}$$

$$E\left[\left(\sum_{t=1}^{T} f_t e_t\right)\left(\sum_{t=1}^{T} g_t \Delta\upsilon_t\right)\right] = \sigma_e^2 \sum_{t=0}^{T-1}\left(\sum_{s=1}^{T-t} f_s g_{s+t} - \sum_{s=1}^{T-t-1} f_s g_{s+t+1}\right)\tilde{\psi}_t, \tag{39}$$

$$\sum_{t=1}^{T} f_t \sum_{s=1}^{t} e_s = \sum_{t=1}^{T}\left(\sum_{s=t}^{T} f_s\right) e_t \tag{40}$$

Also, because f_t can be dummy variable being zero up to the moment T_1, therefore for convenience it is necessary to transform (40) *as follows:*

$$\sum_{t=1}^{T} f_t \sum_{s=1}^{t} e_s = \sum_{t=1}^{T}\left(\sum_{s=T_1+1}^{T} f_s\right) e_t - \sum_{t=1}^{T}\left(\sum_{s=1}^{t} f_{s-1}^{b}\right) e_t, \tag{41}$$

where the expression in a bracket of the second term of right hand side is dummy variable (it is marked by the top index b, summation on all t is equivalent to summation from the moment of $T_1 + 1$) while expression $\left(\sum_{s=T_1+1}^{T} f_s\right)$ in the first term is a constant and does not depend on t.

Using (38) (see KT), we obtain the following elements of expectation of right hand side of (37):

$$
\begin{aligned}
\mathbb{E}[(1,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{T^2}{2} - \frac{t^2}{2} \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(1,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{T^3}{6} - \frac{t^3}{6} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^3}{6} - \frac{tT^2}{2} + \frac{T^3}{3} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^4}{24} - \frac{tT^3}{6} + \frac{T^4}{8} \right) \tilde{\psi}_t + O(T^3) \\
\mathbb{E}[(3,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{T^2}{2} - \frac{1}{2}(t+\lambda_1 T)^2 \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(3,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{T^3}{6} - \frac{1}{6}(t+\lambda_1 T)^3 \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{1}{2}T^2(t+\lambda_1 T) + \frac{1}{6}(t+\lambda_1 T)^3 + \frac{T^3}{3} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{1}{6}T^3(t+\lambda_1 T) + \frac{1}{24}(t+\lambda_1 T)^4 + \frac{T^4}{8} \right) \tilde{\psi}_t + O(T^3) \\
\mathbb{E}[(1,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{2}(1-\lambda_1)^2 T^2 \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(1,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{6}(1-\lambda_1)^3 T^3 \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{6}(1-\lambda_1)^2(\lambda_1+2)T^3 - \frac{1}{2}(1-\lambda_1)^2 tT^2 \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{24}(1-\lambda_1)^3(\lambda_1+3)T^4 - \frac{1}{6}(1-\lambda_1)^3 tT^3 \right) \tilde{\psi}_t + O(T^3) \\
\mathbb{E}[(3,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{2}(1-\lambda_1)^2 T^2 - \frac{t^2}{2} \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(3,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{6}(1-\lambda_1)^3 T^3 - \frac{t^3}{6} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^3}{6} - \frac{1}{2}(1-\lambda_1)^2 tT^2 + \frac{1}{3}(1-\lambda_1)^3 T^3 \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^4}{24} - \frac{1}{6}(1-\lambda_1)^3 tT^3 + \frac{1}{8}(1-\lambda_1)^4 T^4 \right) \tilde{\psi}_t + O(T^3)
\end{aligned}
$$

Using (23) we obtain the final expression of expectation of the second term:

$$\mathbb{E}[R_{22}] = \sigma_e^2 T \sum_{t=0}^{T-1} \frac{1}{2} \left(1 + f_1 \frac{t}{T} + f_2 \frac{t^2}{T^2} + f_3 \frac{t^4}{T^4}\right) \tilde{\psi}_t + O(1), \tag{42}$$

where f_1, f_2 and f_3 are some functions which may depend only on λ_1.

Consider the R_{23} term:

$$\mathbb{E}[R_{23}] = tr \left\{ \left(\sum_{t=1}^{T} d_t d_t'\right)^{-1} \mathbb{E}\left[\left(\sum_{t=1}^{T} d_t \Delta \upsilon_t\right) \left(\sum_{t=1}^{T} \sum_{s=1}^{t} e_s \sum_{s=1}^{t} d_s'\right)\right]\right\}. \tag{43}$$

Using (40) and (41) we can show that:

$$\begin{aligned} \sum_{t=1}^{T} \sum_{s=1}^{t} e_s \sum_{s=1}^{t} d_s' = \Bigg[&\sum_{t=1}^{T} \left(\frac{T^2 - t^2}{2} + O(T)\right) e_t, \\ \sum_{t=1}^{T} \left(\frac{T^3 - t^3}{6} + O(T^2)\right) e_t, &\sum_{t=1}^{T} \left(\frac{T^2(1-\lambda_1)^2}{2} + O(T)\right) e_t - \sum_{t=1}^{T} \left(\frac{(t-\lambda_1 T)^2}{2} + O(T)\right) e_t, \\ &\sum_{t=1}^{T} \left(\frac{T^3(1-\lambda_1)^3}{6} + O(T^2)\right) e_t - \sum_{t=1}^{T} \left(\frac{(t-\lambda_1 T)^3}{6} + O(T^2)\right) e_t, \Bigg]. \end{aligned} \tag{44}$$

Notice that in third and fourth elements the second terms are equal to 0 up to time T_1, i.e. they are dummy variables. Therefore the calculation of interior sum in (38) and (39) the calculation is performed given this fact.

Then using (39) we obtain the following elements of expectation of right hand side of (42):

$$
\begin{aligned}
\mathbb{E}[(1,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(tT - \frac{t^2}{2} \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(1,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^3}{6} - \frac{t^2 T}{2} + \frac{tT^2}{2} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{t^3}{6} + tT^2 - \frac{T^3}{3} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^4}{24} - \frac{t^2 T^2}{4} + \frac{tT^3}{2} - \frac{T^4}{8} \right) \tilde{\psi}_t + O(T^3) \\
\mathbb{E}[(3,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\lambda_1 tT + tT + \frac{\lambda_1^2 T^2}{2} - \frac{T^2}{2} \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(3,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{2}\lambda_1 t^2 T - \frac{t^2 T}{2} - \frac{1}{2}\lambda_1^2 tT^2 + \frac{tT^2}{2} + \frac{\lambda_1^3 T^3}{6} - \frac{T^3}{6} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,1)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{1}{2}\lambda_1^2 tT^2 - \lambda_1 tT^2 + \frac{tT^2}{2} - \frac{1}{6}\lambda_1^3 T^3 + \frac{\lambda_1 T^3}{2} - \frac{T^3}{3} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,2)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{1}{4}\lambda_1^2 t^2 T^2 + \frac{1}{2}\lambda_1 t^2 T^2 - \frac{t^2 T^2}{4} \right. \\
&\quad \left. + \frac{1}{6}\lambda_1^3 tT^3 - \frac{1}{2}\lambda_1 tT^3 + \frac{tT^3}{3} - \frac{1}{24}\lambda_1^4 T^4 + \frac{\lambda_1 T^4}{6} - \frac{T^4}{8} \right) \tilde{\psi}_t + O(T^3) \\
\mathbb{E}[(1,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{t^2}{2} + Tt - T\lambda_1 t \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(1,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^3}{6} - \frac{Tt^2}{2} + \frac{1}{2}T\lambda_1 t^2 + \frac{T^2 t}{2} + \frac{1}{2}T^2\lambda_1^2 t - T^2\lambda_1 t \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,3)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{t^3}{6} - \frac{1}{2}T\lambda_1 t^2 + T^2 t - T^2\lambda_1 t - \frac{T^3}{3} - \frac{T^3\lambda_1^3}{6} + \frac{T^3\lambda_1}{2} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(2,4)\text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^4}{24} + \frac{1}{6}T\lambda_1 t^3 - \frac{T^2 t^2}{4} + \frac{1}{4}T^2\lambda_1^2 t^2 + \frac{T^3 t}{2} + \frac{1}{2}T^3\lambda_1^2 t - T^3\lambda_1 t \right. \\
&\quad \left. - \frac{T^4}{8} + \frac{T^4\lambda_1^4}{24} - \frac{T^4\lambda_1^2}{4} + \frac{T^4\lambda_1}{3} \right) \tilde{\psi}_t + O(T^3)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[(3,3) \text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{t^2}{2} + Tt - T\lambda_1 t \right) \tilde{\psi}_t + O(T) \\
\mathbb{E}[(3,4) \text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^3}{6} - \frac{Tt^2}{2} + \frac{1}{2} T\lambda_1 t^2 + \frac{T^2 t}{2} + \frac{1}{2} T^2 \lambda_1^2 t - T^2 \lambda_1 t \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,3) \text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(-\frac{t^3}{6} - \frac{1}{2} T\lambda_1 t^2 + T^2 t - T^2 \lambda_1 t - \frac{T^3}{3} - \frac{T^3 \lambda_1^3}{6} + \frac{T^3 \lambda_1}{2} \right) \tilde{\psi}_t + O(T^2) \\
\mathbb{E}[(4,4) \text{ element}] &= \sigma_e^2 T \sum_{t=0}^{T-1} \left(\frac{t^4}{24} + \frac{1}{6} T\lambda_1 t^3 - \frac{T^2 t^2}{4} + \frac{1}{4} T^2 \lambda_1^2 t^2 + \frac{T^3 t}{2} + \frac{1}{2} T^3 \lambda_1^2 t - T^3 \lambda_1 t \right. \\
&\quad \left. - \frac{T^4}{8} + \frac{T^4 \lambda_1^4}{24} - \frac{T^4 \lambda_1^2}{4} + \frac{T^4 \lambda_1}{3} \right) \tilde{\psi}_t + O(T^3)
\end{aligned}
$$

Thus using (23) we obtain the final expression for R_{23}:

$$
\mathbb{E}[R_{23}] = \sigma_e^2 T \sum_{t=0}^{T-1} \frac{1}{2} \left(1 + f_4 \frac{t}{T} + f_5 \frac{t^2}{T^2} + f_6 \frac{t^4}{T^4} \right) \tilde{\psi}_t + O(1), \tag{45}
$$

where f_4, f_5 and f_6 are some functions determining as for (42).

Similarly, we obtain an expression for R_{24}:

$$
\mathbb{E}[R_{24}] = tr \left\{ DD \times \mathbb{E} \left[\left(\sum_{t=1}^{T} d_t \Delta v_t \right) \left(\sum_{t=1}^{T} e_t d_t' \right) \right] \right\}. \tag{46}
$$

where the expression for DD has been obtained in (32). Considering expectation in right hand side and using (39) we obtain:

$$
\mathbb{E} \left[\left(\sum_{t=1}^{T} d_t \Delta v_t \right) \left(\sum_{t=1}^{T} e_t d_t' \right) \right] = \sigma_e^2 (\eta_1, \eta_2, \eta_3, \eta_4), \tag{47}
$$

where

$$
\eta_1 = \begin{pmatrix} \sum_{t=0}^{T-1} \tilde{\psi}_t \\ \sum_{t=0}^{T-1} t\tilde{\psi}_t + O(1) \\ 0 \\ 0 \end{pmatrix},
$$

$$
\eta_2 = \begin{pmatrix} \sum_{t=0}^{T-1} (T-t)\tilde{\psi}_t \\ \sum_{t=0}^{T-1} \frac{T^2 - t^2}{2} \tilde{\psi}_t + O(T) \\ \sum_{t=0}^{T-1} (T - \lambda_1 T)\tilde{\psi}_t + O(T) \\ \sum_{t=0}^{T-1} \frac{(\lambda_1 - 1)^2 T^2}{2} \tilde{\psi}_t + O(T) \end{pmatrix},
$$

$$
\eta_3 = \begin{pmatrix} \sum_{t=0}^{T-1} \tilde{\psi}_t \\ \sum_{t=0}^{T-1} (t + \lambda_1 T)\tilde{\psi}_t + O(1) \\ \sum_{t=0}^{T-1} \tilde{\psi}_t \\ \sum_{t=0}^{T-1} t\tilde{\psi}_t + O(1) \end{pmatrix},
$$

$$\eta_4 = \begin{pmatrix} \sum_{t=0}^{T-1}(T-\lambda_1 T - t)\tilde{\psi}_t + O(1) \\ \sum_{t=0}^{T-1}\frac{1}{T^2(1+\lambda_1)^2 - t^2}\tilde{\psi}_t \\ \sum_{t=0}^{T-1}(T-\lambda_1 T - t)\tilde{\psi}_t + O(1) \\ \sum_{t=0}^{T-1}\frac{1}{T^2(1-\lambda_1)^2 - t^2}\tilde{\psi}_t + O(T) \end{pmatrix},$$

Using obtained expression of DD matrix it can be shown that:

$$\mathbb{E}[R_{24}] = \sigma_e^2 T \sum_{t=0}^{T-1} \frac{1}{2}\left(\frac{19}{15} + f_7\frac{t}{T} + f_8\frac{t^2}{T^2}\right)\tilde{\psi}_t + O(1), \tag{48}$$

where f_7 and f_8 are defined as earlier.

Further, from (36), (42), (45) and (48) we obtain:

$$\mathbb{E}[R_2] = \frac{2\sigma_e^2\psi(1)}{T}\sum_{t=0}^{T-1}\left(\frac{19}{30} + f_6\frac{t}{T} + f_7\frac{t^2}{T^2} + f_8\frac{t^4}{T^4}\right)\tilde{\psi}_t + o(T^{-1}). \tag{49}$$

As $\sum_{j=0}^{\infty}|\tilde{\psi}_j| < \infty$ the sum in (49) converges to $\sum_{j=0}^{\infty}(19/30)\tilde{\psi}_j$. Also, notice that $\psi(1) = 1/\phi(1)$ and $\sum_{j=0}^{\infty}\tilde{\psi}_t = \psi'(1) = \left(\frac{1}{\phi(1)}\right)' = -\frac{\phi'(1)}{\phi^2(1)}$ we obtain:

$$\mathbb{E}[R_2] = -\frac{1}{T}\frac{19}{15}\frac{\sigma_e^2\phi'(1)}{\phi^3(1)}. \tag{50}$$

■

References

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

Busetti, F. and Harvey, A.C. (2003). Further Comments On Stationarity Tests In Series With Structural Breaks At Unknown Points. *Journal of Time Series Analysis*, **24**, 137–140.

Carrion-i Silvestre, J.L., Kim, D., and Perron, P. (2009). GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. *Econometric Theory*, **25**, 1754–1792.

Carrion-i-Silvestre, J.L. and Sansó, A. J. (2005). The KPSS test with two structural breaks. *Spanish Economic Review*, **9**, 105–127.

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

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427–431.

Harvey, D.I. and Leybourne, S.J. (2012). Break date estimation for models with deterministic structural change. Unpublished manuscript.

Harvey, D.I. and Mills, T.C. (2003). A note on Busetti-Harvey tests for stationarity in series with structural breaks. *Journal of Time Series Analysis*, **24**, 159–164.

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

Kurozumi, E. and Tanaka, S. (2010). Reducing the size distortion of the KPSS test. *Journal of Time Series Analysis*, **31**, 415–426.

Kwiatkowski, D., P.C.B., Phillips, Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics*, **54**, 159–178.

Lee, J. and Strazicich, M.C. (2001). Testing the null of stationarity in the presence of a structural break. *Applied Economics Letters*, **8**, 377–382.

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

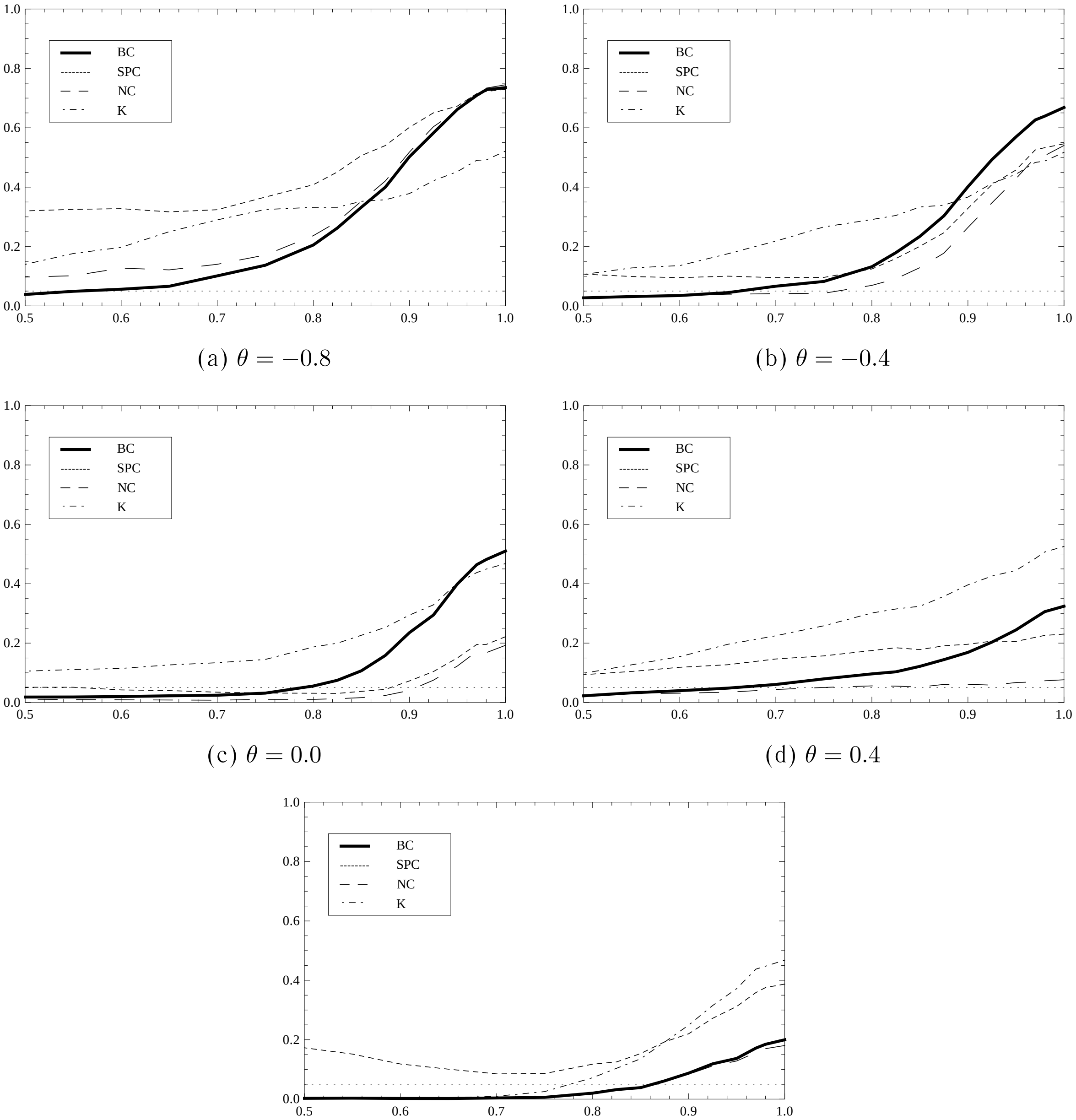
Perron, P. (2006). *Dealing with Structural Breaks*, Chapter 1: Econometric theory. Palgrave Macmillan.

Perron, P. and Zhu, X. (2005). Structural breaks with deterministic and stochastic trends. *Journal of Econometrics*, **129**, 65–119.

Phillips, P.C.B. and Solo, V. (1992). Asymptotics for linear processes. *Annals of Statistics*, **20**, 971–1001.

Sul, D., P.C.B., Phillips, and Choi, C.-Y. (2005). Prewhitening Bias in HAC Estimation. *Oxford Bulletin of Economics and Statistics*, **67**, 517–546.

Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. Wiley, New York.



(a) $\theta = -0.8$

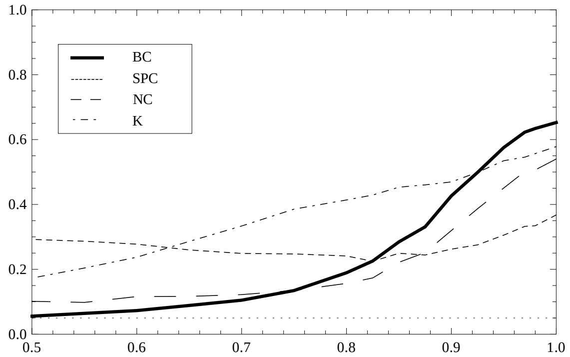
(b) $\theta = -0.4$

(c) $\theta = 0.0$

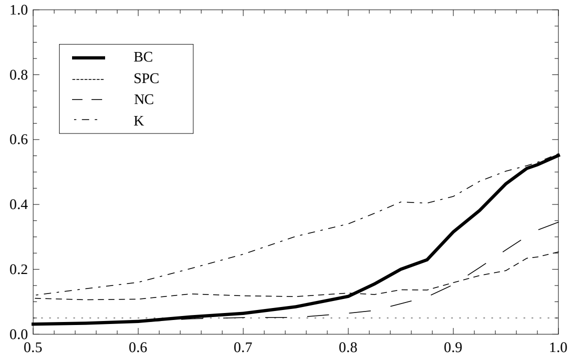
(d) $\theta = 0.4$

(e) $\theta = 0.8$

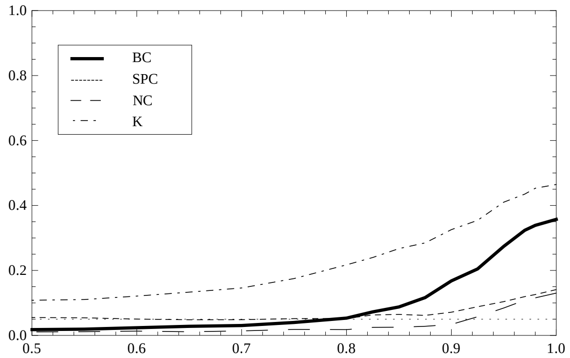
Figure 1. Size and power, known break date



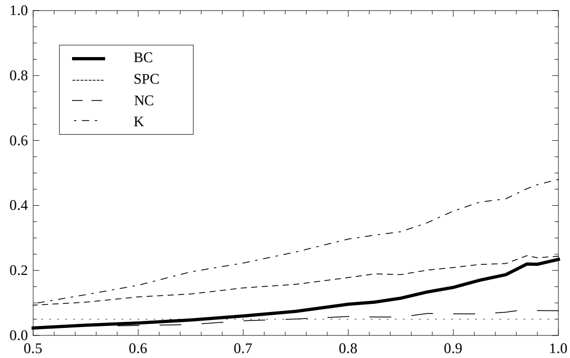
(a) $\theta = -0.8$



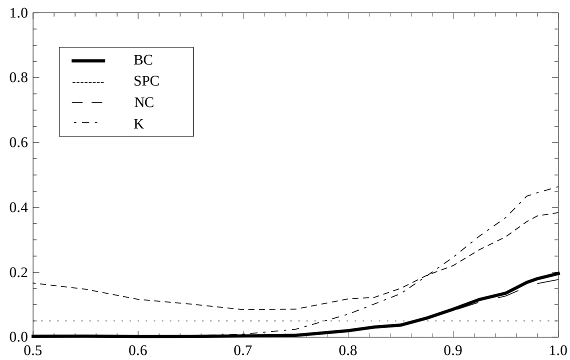
(b) $\theta = -0.4$



(c) $\theta = 0.0$



(d) $\theta = 0.4$



(e) $\theta = 0.8$

Figure 2. Size and power, unknown break date