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

Anton Skrobotov ${ }^{\dagger}$<br>Gaidar Institute for Economic Policy, The Russian Presidential Academy<br>of National Economy and Public Administration

October 22, 2013


#### Abstract

In this paper we extend the stationarity test proposed by Kurozumi and Tanaka (2010) to reduce size distortion with one structural break in data generating process. We find the bias up to the order of $1 / T$ for four types of models containing structural breaks. Simulations on finite samples show a decrease of size distortions relative to 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 is hypothesis testing of unit root in a time series, first proposed in the work of Dickey and Fuller (Dickey and Fuller, 1979). Recently, Kwiatkowski et al. (1992) (hereafter KPSS) proposed an opposite direction of inquiry, employing the null hypothesis of stationarity about the deterministic trend as a reversal complement of the unit root tests. A serious challenge to these tests is an assumption about the type of deterministic function.

Results of a study by Perron (1989) indicate that conventional 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 a serious size distortion in the usual KPSS test (see, e.g., Lee et al. (1997)). To address this problem several studies were conducted, allowing for breaks in stationarity tests.

Lee and Strazicich (2001) used an analogue of the KPSS test, where two models were considered: one model with a change in level for a non-trending series and the other with a change in

[^0]both level and slope. In case of an unknown break date the authors used the test statistic obtained through minimizing the sequence of stationarity statistics for each possible break date (infimumtest). However, it was shown that the test has low power, as minimization 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. The author obtained limiting distributions of test statistics for all models, asymptotic local power function (approach stated in Tanaka (1996, chapter 9)) and investigated the power of tests relative to the location of the break. Furthermore, this study proposed another test, in which limiting distribution did not depend on the timing of the break. In case of an unknown break date, the author used an estimator of the break date - obtained by minimizing the sum of squared residuals.

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

Busetti and Harvey (2003) investigated different ways of estimating an unknown break date. 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. A natural question arises about how large the shift should be to satisfy this assumption, as it was shown that in the presence of large changes the rejection of the null hypothesis occurs too frequently. Another procedure proposed by Busetti and Harvey (2003) consists of making a preliminary estimation of the break date by minimizing the sum of squared residuals, and then using the standard test statistic with the obtained estimate as the true date of break. This approach was shown to contain good size properties, but loses power under small changes. Simulations have shown that minimization of test statistics rejects the null hypothesis of stationarity too often, even for shift as small as one standard deviation of errors.

While the empirical size is never above 0.17 for large shifts, the size of the test based on a superconsistent estimate of break fraction is close to nominal at any size of break ${ }^{1}$. However, the power of the test is too small for a small shift, increasing with growth in magnitude (i.e. in a situation where it is easier to identify change). If the break does not occur, the infimum-test is more powerful than a test relying on a superconsistent estimate of break fraction obtained through minimization of the sum of squares residuals. Thus, if there is uncertainty about the presence of a break, the infimum-test is preferable. However, if it can be shown with confidence that a break exists, but its location is not 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-i-Rosselló (2005) consider a possibility of two structural breaks in the KPSS test. This study uses an approach proposed by Sul et al. (2005) (hereafter SPC) for the long-run variance estimator. The authors obtained limiting distributions for each of the seven models examined in the article. A search of unknown break dates was made through minimization of the sum of squares residuals.

[^1]Another problem of KPSS type test has been considered in Carrion-i-Silvestre and Sansó-i-Rosselló (2006), among others. This study revealed that on finite samples the SPC test with AR(1) pre-whitening was more preferable than others, as it controls size. However, the SPC test with $A R(1)$ pre-whitening contains serious size distortion when the data generating process (DGP) is $\operatorname{AR}(2)$ (or higher order). Kurozumi and Tanaka (2010) (hereafter KT) extended the SPC approach (in absence of breaks) where the authors employed autoregressive approximation for estimation of a long-run variance, using a boundary rule introduced by SPC to assure the consistency of long-run variance. The authors also investigated the problem of downward bias in the numerator of the KPSS test statistic. For size correction they derived the finite sample bias and proposed a bias-corrected version of the KPSS test. Simulations showed that the empirical size of the modified test is well controlled in the case of $\operatorname{AR}(2)$ errors, and the modified test also has a higher power in comparison with the SPC. In this paper we propose the extension of the KT test to a case of a single structural break.

In this work we derive a finite sample bias of KPSS test statistic in a case where one break in DGP is present, using a modification of the boundary rule with autoregressive approximation, as in KT. If the break date is unknown, we use its estimate obtained through a minimization of the sum of squares of residuals using the approach of Harvey and Leybourne (2013). Results (finite sample simulations) derived from this study suggest superiority of the modified test, even in case of MA error.

The paper is organized in the following order. In section 2 we discuss the model, test statistic and the boundary rule of SPC with modification of long-run variance proposed by KT. Also we derive the parameter of bias for different types of models containing a structural 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. Obtained results are formulated in the Conclusion.

## 2 The Model

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

$$
\begin{equation*}
y_{t}=d_{t}^{\prime} \beta+u_{t}, t=1, \ldots, T \tag{1}
\end{equation*}
$$

where $d_{t}$ is some deterministic component, and process $u_{t}$ satisfy following standard assumptions (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 a linear process such that

$$
u_{t}=c(L) e_{t}=\sum_{i=0}^{\infty} c_{i} e_{t-i}
$$

with $c(z) \neq 0$ for all $|z| \leq 1$ and $\sum_{i=0}^{\infty} i\left|c_{i}\right|<\infty$, where $e_{t}$ - martingale difference sequence with conditional variance $\sigma_{e}^{2}$ and $\sup _{t} \mathbb{E}\left(e_{t}^{4}\right)<\infty$. Short run and long-run variance are defined as $\sigma_{u}^{2}=\mathbb{E}\left(u_{t}^{2}\right)$ and $\omega_{u}^{2}=\lim _{T \rightarrow \infty} T^{-1} \mathbb{E}\left(\sum_{t=1}^{T} u_{t}\right)^{2}=\sigma_{e}^{2} c(1)^{2}$, respectively;

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

Also, as in Perron (1989) (see also Perron (2006)) we consider four types of models: Model 0 (a change in level), respectively for a non-trending and 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:

$$
d_{t}^{\prime}= \begin{cases}\left(1, D U_{t}\right), & \text { for Model } 0 \\ \left(1, t, D U_{t}\right), & \text { for Model 0t } \\ \left(1, t, D T_{t}\right), & \text { for Model I } \\ \left(1, t, D U_{t}, D T_{t}\right), & \text { for Model II }\end{cases}
$$

where $D U_{t}=\mathbb{I}\left(t \geq T_{1}+1\right), D T_{t}=\left(t-T_{1}\right) \mathbb{I}\left(t \geq T_{1}+1\right), \mathbb{I}(\cdot)$ is the indicator function, $T_{1}$ is the break date. We also define the break fraction as $\lambda_{1}=T_{1} / T$.

We are testing null hypothesis of stationarity $\left(u_{t} \sim I(0)\right)$ against the alternative of a unit root $\left(u_{t} \sim I(1)\right)$. It is usually used following the KPSS test statistic for testing stationarity against the unit root:

$$
\begin{equation*}
\operatorname{KPSS}\left(\lambda_{1}\right)=\frac{T^{-2} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{u}_{s}\right)^{2}}{\hat{\omega}_{u}^{2}} \tag{2}
\end{equation*}
$$

where $\hat{u}_{t}=y_{t}-d_{t}^{\prime} \hat{\beta}$ are OLS-residuals of $y_{t}$ on $d_{t}$, where $d_{t}=\left[1, D U_{t}\right]^{\prime}, \beta=\left(\mu_{0}, \mu_{1}\right)^{\prime}$ for Model $0, d_{t}=\left[1, t, D U_{t}\right]^{\prime}, \beta=\left(\mu_{0}, \beta_{0}, \mu_{1}\right)^{\prime}$ for Model $0 t, d_{t}=\left[1, t, D T_{t}\right]^{\prime}, \beta=\left(\mu_{0}, \beta_{0}, \beta_{1}\right)^{\prime}$ for Model I, $d_{t}=\left[1, t, D U_{t}, D T_{t}\right]^{\prime}, \beta=\left(\mu_{0}, \beta_{0}, \mu_{1}, \beta_{1}\right)^{\prime}$ for Model II, and long-run variance estimator $\hat{\omega}_{u}^{2}$ is constructed according to the nonparametric approach using the Bartlett or QS kernel.

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

## Lemma 1 Under the null hypothesis

$$
K P S S\left(\lambda_{1}\right) \Rightarrow \int_{0}^{1}\left(W^{*}\left(r, \lambda_{1}\right)\right)^{2} d r
$$

where for Model 0:

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

for Model 0t:

$$
W^{*}\left(r, \lambda_{1}\right)= \begin{cases}W(r)-\frac{r}{\lambda_{1}} W\left(\lambda_{1}\right)-\frac{6 r\left(r-\lambda_{1}\right)}{1-3 \lambda_{1}+3 \lambda_{1}^{2}} & \\ \times\left[\int_{0}^{1} r d W(r)-\frac{\lambda_{1}}{2} W\left(\lambda_{1}\right)-\frac{1+\lambda_{1}}{2}\left(W(1)-W\left(\lambda_{1}\right)\right)\right], & \text { for } r \leq \lambda_{1} \\ \left(W(r)-W\left(\lambda_{1}\right)\right)-\frac{r-\lambda_{1}}{1-\lambda_{1}}\left(W(1)-W\left(\lambda_{1}\right)\right)-\frac{6(r-1)\left(r-\lambda_{1}\right)}{1-3 \lambda_{1}+3 \lambda_{1}^{2}} & \\ \times\left[\int_{0}^{1} r d W(r)-\frac{\lambda_{1}}{2} W\left(\lambda_{1}\right)-\frac{1+\lambda_{1}}{2}\left(W(1)-W\left(\lambda_{1}\right)\right)\right], & \text { for } r>\lambda_{1}\end{cases}
$$

for Model I:

$$
W^{*}\left(r, \lambda_{1}\right)= \begin{cases}W(r)-r W(1)-\frac{3}{\lambda_{1}^{3}\left(1-\lambda_{1}\right)^{3}} & \\ \times\left[\left(a \frac{r^{2}}{2}-a \lambda_{1} r+\frac{r}{2}\left(a \lambda_{1}^{2}-b\left(1-\lambda_{1}\right)^{2}\right)\right) J_{1}\right. & \\ \left.+\left(b \frac{r^{2}}{2}-b \lambda_{1} r+\frac{r}{2}\left(b \lambda_{1}^{2}-c\left(1-\lambda_{1}\right)^{2}\right)\right) J_{2}\right], & \text { for } r \leq \lambda_{1} \\ W(r)-r W(1)-\frac{3}{\lambda_{1}^{3}\left(1-\lambda_{1}\right)^{3} ;} & \\ \times\left[\left(-a \frac{\lambda_{1}^{2}}{2}+b \frac{r^{2}-\lambda_{1}^{2}}{2}-b \lambda_{1}\left(r-\lambda_{1}\right)+\frac{r}{2}\left(a \lambda_{1}^{2}-b\left(1-\lambda_{1}\right)^{2}\right)\right) J_{1}\right. \\ \left.+\left(-b \frac{\lambda_{1}^{2}}{2}+c \frac{r^{2}-\lambda_{1}^{2}}{2}-c \lambda_{1}\left(r-\lambda_{1}\right)+\frac{r}{2}\left(b \lambda_{1}^{2}-c\left(1-\lambda_{1}\right)^{2}\right)\right) J_{2}\right], & \text { for } r>\lambda_{1}\end{cases}
$$

for Model II:

$$
W^{*}\left(r, \lambda_{1}\right)= \begin{cases}W(r)-\frac{r}{\lambda_{1}} W\left(\lambda_{1}\right)-\frac{6 r\left(r-\lambda_{1}\right)}{\lambda_{1}^{3}} & \\ \times\left[\int_{0}^{\lambda_{1}} r d W(r)-\frac{\lambda_{1}}{2} W\left(\lambda_{1}\right)\right], & \text { for } r \leq \lambda_{1} \\ \left(W(r)-W\left(\lambda_{1}\right)\right)-\frac{r-\lambda_{1}}{1-\lambda_{1}}\left(W(1)-W\left(\lambda_{1}\right)\right)-\frac{6(r-1)\left(r-\lambda_{1}\right)}{\left(1-\lambda_{1}\right)^{3}} & \\ \times\left[\int_{\lambda_{1}}^{1} r d W(r)-\frac{1+\lambda_{1}}{2}\left(W(1)-W\left(\lambda_{1}\right)\right)\right], & \text { for } r>\lambda_{1}\end{cases}
$$

Here $a=\left(1-\lambda_{1}\right)^{3}\left(1+\lambda_{1}\right), b=-3 \lambda_{1}^{2}\left(1-\lambda_{1}\right)^{2}, c=\lambda_{1}^{3}\left(4-3 \lambda_{1}\right), J_{1}=\int_{0}^{\lambda_{1}} r d W(r)-\lambda_{1} W\left(\lambda_{1}\right)+$ $\frac{\lambda_{1}^{2}}{2} W(1)$ and $J_{1}=\int_{\lambda_{1}}^{1} r d W(r)-\lambda_{1}\left(W(1)-W\left(\lambda_{1}\right)\right)-\frac{\left(1-\lambda_{1}\right)^{2}}{2} W(1)$.

The KPSS test can control the size of the test asymptotically, but in finite samples it has a serious size distortion. For their reduction SPC proposed the AR(1) pre-whitening method with a boundary rule. Thus, we first estimate the $\operatorname{AR}(\mathrm{p})$ model for residuals of regression (1), $\hat{u}_{t}$ :

$$
\hat{u}_{t}=\phi_{1} \hat{u}_{t-1}+\cdots+\phi_{p} \hat{u}_{t-p}+e_{t} .
$$

Then, the long-run variance estimator is constructed as

$$
\begin{equation*}
\hat{\omega}_{u}^{2}=\frac{\hat{\omega}_{e}^{2}}{(1-\tilde{\phi})^{2}}, \tag{3}
\end{equation*}
$$

where $\tilde{\phi}=\min \left\{\sum_{j=1}^{p} \hat{\phi}_{j}, 1-1 / \sqrt{T}\right\}$ and $\hat{\omega}_{e}^{2}$ is the long-run variance estimator of residuals $\hat{e}_{t} .{ }^{2}$ KT proposed the modification of SPC for the case of autoregressive long-run variance estimator $\omega_{u}^{2}$, i.e.:

$$
\begin{equation*}
\hat{\omega}_{u, A R}^{2}=\frac{\hat{\sigma}_{e}^{2}}{(1-\tilde{\phi})^{2}}, \tag{4}
\end{equation*}
$$

where $\hat{\sigma}_{e}^{2}=T^{-1} \sum_{t=1}^{T} \hat{e}_{t}^{2}$ 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 with significant accuracy in this case, there is a problem of a downward bias of test statistic (2) in finite samples. KT showed

[^2]that for the cases of a constant and a trend their modification of the long-run variance estimator still leads to rare rejection of the null hypothesis due to this downward bias. This leads to essential power losses under the alternative hypothesis. For prevention of bias in the numerator of the KPSS test statistic in finite samples KT proposed this bias-corrected version:
\[

$$
\begin{equation*}
K P S S=\frac{T^{-2} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{u}_{s}\right)^{2}-\hat{b}_{T}}{\hat{\omega}_{u, A R}^{2}} \tag{5}
\end{equation*}
$$

\]

Here, the term $b_{T}$ is responsible for the bias. To calculate its value, KT suggested the use of the Beveridge-Nelson decomposition. Let $u_{t}=\psi(L) e_{t}$, then the process $u_{t}$ can be written as

$$
u_{t}=\psi(1) e_{t}+v_{t-1}-v_{t}
$$

where $v_{t}=\sum_{j=0}^{\infty} \tilde{\psi}_{j} e_{t-j}, \tilde{\psi}_{j}=\sum_{i=j+1}^{\infty} \psi_{i}$. The residuals $\hat{u}_{t}$ are defined as

$$
\begin{align*}
\hat{u}_{t} & =u_{t}-d_{t}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} u_{t}  \tag{6}\\
& =\psi(1) e_{t}+v_{t-1}-v_{t}-d_{t}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t}\left(\psi(1) e_{t}+v_{t-1}-v_{t}\right) \\
& =\psi(1) \hat{e}_{t}-\widehat{\Delta v_{t}}
\end{align*}
$$

where $\hat{e}_{t}$ and $\widehat{\Delta v_{t}}$ are the residuals of regression of $e_{t}$ and $\Delta v_{t}$ on $d_{t}$, respectively. Then, KT decomposed the numerator of (2) into three terms:

$$
\begin{align*}
\frac{1}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{u}_{s}\right)^{2}= & \frac{\psi^{2}(1)}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{e}_{s}\right)^{2} \\
& +\frac{1}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \widehat{\Delta v_{s}}\right)^{2}-\frac{2 \psi(1)}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{e}_{s}\right)\left(\sum_{s=1}^{t} \widehat{\Delta v_{s}}\right) \\
= & \frac{\psi^{2}(1)}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \hat{e}_{s}\right)^{2}+R_{1}-R_{2} . \tag{7}
\end{align*}
$$

The second and third terms are $o_{p}(1)$, while the first term have non-degenerate 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 $O\left(T^{-1}\right)$. This bias is defined as $b_{T}$ :

$$
\begin{equation*}
\mathbb{E}\left[R_{1}-R_{2}\right]=b_{T}+o\left(T^{-1}\right) \tag{8}
\end{equation*}
$$

For further analysis, we need the following additional assumption.
Assumption 2 Let $\gamma_{T_{1}}$ and $\gamma_{T-T_{1}}$ are $o(1)$, where $\gamma_{j}=\mathbb{E}\left[v_{t} v_{t-j}\right]$.
Assumption 2 is sufficient for our purposes, as it is usually assumed that the shift is in the range of $(0.15 T, 0.85 T)$, so we can ignore the covariance between the $T_{1}$ observations. Thus, the following theorem holds.

Proposition 1 Let $\gamma_{0}=\mathbb{E}\left[v_{t}^{2}\right]$, lag polynomial $\varphi(L)=c(L)(1-\rho L)$ and $\varphi^{\prime}(1)=d \varphi(z) /\left.d z\right|_{z=1}$. Then, bias $b_{T}$ in the numerator of KPSS test statistic (5) is expressed as

$$
\begin{equation*}
b_{T}=\frac{b_{0}}{T}\left(\gamma_{0}+\sigma_{e}^{2} \frac{\varphi^{\prime}(1)}{\varphi^{3}(1)}\right), \tag{9}
\end{equation*}
$$

where $b_{0}=5 / 3$ for Model $0, b_{0}=\frac{285 \lambda_{1}^{4}-570 \lambda_{1}^{3}+498 \lambda_{1}^{2}-213 \lambda_{1}+38}{30\left(1-3 \lambda_{1}+3 \lambda_{1}^{2}\right)^{2}}$ for Model Ot, $b_{0}=7 / 6$ for Model I and $b_{0}=19 / 15$ for Model $I I^{3}$.

Remark. It should be noted, that for Model 0, I and II the bias does not depend on the break fraction $\lambda$. But, this is not the case for Model 0 t. For Models 0 and II the bias is the same as for models without a structural break considered in KT. In addition, if $\lambda_{1}$ approaches 0 or 1 the bias term for Model Ot approaches $19 / 15$, which corresponds to the models with no break. However, the bias for Model I is the same for any value of $\lambda_{1}$, except for $\lambda_{1}=0$ and $\lambda_{1}=1$ (the case of structural break absence). The explanation for this discontinuity arises from the fact that in the proof we use the condition of Assumption 2, that $\gamma_{T_{1}}$ and $\gamma_{T-T_{1}}$ are $o(1)$. However, as $\lambda_{1} \rightarrow 0$ and $\lambda_{1} \rightarrow 1$ autocovariances $\gamma_{T_{1}}$ and $\gamma_{T-T_{1}}$ approach $\gamma_{0} \neq 0$, so that the bias parameter $b_{0}$ will approach the case without break.

Parameter $\gamma_{0}$ can be constructed by 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 the previous section is based on the assumption that the timing of the structural break is known. However, in many cases it cannot be established. In such cases it is possible to replace a known break fraction with its superconsistent estimate. Then, the limiting distribution of the test statistics remains the same. The superconsistent estimator of the break fraction $\lambda_{1}=T_{1} / T$ can be obtained by 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 $I(0)$ case $^{4}$ (see, e.g., Perron and Zhu (2005)).

The alternative procedure of searching an unknown break date was proposed by Carrion-iSilvestre et al. (2009), using preliminary (quasi) GLS-detrending of $y_{t}$. Let us estimate the following regression:

$$
\begin{equation*}
\mathbf{y}=\mathbf{X}\left(\lambda_{1}\right) \beta+\mathbf{u}, \tag{10}
\end{equation*}
$$

where vector $\mathbf{y}=\left[y_{1}, y_{2} \ldots, y_{T}\right]^{\prime}$, matrix $\mathbf{X}\left(\lambda_{1}\right)$ includes all regressors, and $\beta$ denotes the corresponding parameters vector. Then, the GLS-estimate $\hat{\beta}$ of vector $\beta$ is the OLS-estimate of coefficient vector in equation

$$
\begin{equation*}
\mathbf{y}^{\bar{\rho}}=\mathbf{X}^{\bar{\rho}}\left(\lambda_{1}\right) \beta+\mathbf{u}^{\bar{\rho}}, \tag{11}
\end{equation*}
$$

where

$$
\mathbf{y}^{\bar{\rho}}=\left[y_{1},(1-\bar{\rho} L) y_{2}, \ldots,(1-\bar{\rho} L) y_{T}\right]^{\prime},
$$

[^3]$$
\mathbf{X}^{\bar{\rho}}\left(\lambda_{1}\right)=\left[x_{1},(1-\bar{\rho} L) x_{2}, \ldots,(1-\bar{\rho} L) x_{T}\right]^{\prime}
$$

Carrion-i-Silvestre et al. (2009) suggested choosing $\bar{c}$ in $\bar{\rho}=1+\bar{c} / T$ depending on timing of break, because it is not known whether the series $u_{t}$ is $I(0)$ or $I(1)$. Then, the estimator of break date is:

$$
\begin{equation*}
\hat{\lambda}_{1}^{\bar{\rho}}=\arg \min _{\lambda_{1} \in \Lambda(e)} S\left(\bar{\rho}, \lambda_{1}\right), \tag{12}
\end{equation*}
$$

where $S\left(\bar{\rho}, \lambda_{1}\right)$ is the sum of squared residuals in regression (11). The obtained estimate of the break fraction will be superconsistent for Models I and II under both the $I(0)$ and $I(1)$ cases.

Harvey and Leybourne (2013) proposed the modification of this break date estimator for Models I and II using additional information, if the series is $I(1)$. In this case, the estimate of break fraction obtained for the GLS detrended series will also be superconsistent, but the estimate obtained for the first-differenced series will be efficient. This study proposes the use of a hybrid estimator:

$$
\begin{equation*}
\hat{\lambda}_{1}^{D_{m}}=\arg \min _{\lambda_{1} \in \Lambda(e), \bar{\rho} \in D_{m}} S\left(\bar{\rho}, \lambda_{1}\right), \tag{13}
\end{equation*}
$$

where $D_{m}=\left\{\rho_{1}^{\prime}, \rho_{2}^{\prime}, \ldots, \rho_{m-1}^{\prime}, 1\right\}$ is the $m$ element set, where $\left|\rho_{i}^{\prime}\right|<1$ for all $i$ and, without loss of generality, $-1<\rho_{1}^{\prime}<\rho_{2}^{\prime}<\cdots<\rho_{m-1}^{\prime}<1$.

Asymptotic results show that it is necessary to set $\rho_{m-1}^{\prime}$ close enough to unit, so that the estimator (13) has desired asymptotic properties (if the true value of $\rho>\rho_{m-1}^{\prime}$, then the break fraction estimator will be inefficient). As set $D_{m}$ Harvey and Leybourne (2013) suggest using $D_{m}=\{0,0.2,0.4,0.6,0.8,0.9,0.95,0.975,1\}$, reasoning that a negative serial correlation is not typically observed in practice, and also the serial correlation can oftenly be strongly positive. Therefore, inclusion of 0.975 value allows a small enough interval $0.975<\rho<1$ for adequate asymptotic choice.

We will use this estimator in the next section for Models I and II. However, for Models 0 and $0 t$, the estimator of the break fraction, obtained by minimizing the sum of squared residuals for series in level, should be used, as the break date estimators discussed above may not be consistent in $I(1)$ case. It should be noted, that as an estimator of break fraction is superconsistent under the null hypothesis of $I(0)$, according Kurozumi (2002) the critical values in the known break date case (obtained by Busetti and Harvey (2001) or Kurozumi (2002)) are asymptotically valid to control the size of the test. Also, the test is consistent because under the alternative hypothesis of $I(1)$, the test statistics diverge, irrespective of the value of break fraction, although it would not be consistent. However, some power loss is expected in comparison to the case with a known break date. This effect is investigated by simulations.

We will briefly discuss another important issue that arises in estimating the dating of structural changes. When the break date is estimated, it may differ from the actual date. This leads to the fact that the expression (6) will not be correct. In the case of an unknown break date, this expression
can be rewritten as

$$
\begin{aligned}
\hat{u}_{t} & =u_{t}+d_{t}^{\prime}\left(\lambda_{1}\right) \beta\left(\lambda_{1}\right)-d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\left(\sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right) d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\right)^{-1} \sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right)\left[u_{t}+d_{t}^{\prime}\left(\lambda_{1}\right) \beta\left(\lambda_{1}\right)\right] \\
& =\left[u_{t}-d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\left(\sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right) d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\right)^{-1} \sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right) u_{t}\right] \\
& +\left[d_{t}^{\prime}\left(\lambda_{1}\right)-d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\left(\sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right) d_{t}^{\prime}\left(\hat{\lambda}_{1}\right)\right)^{-1} \sum_{t=1}^{T} d_{t}\left(\hat{\lambda}_{1}\right) d_{t}^{\prime}\left(\lambda_{1}\right)\right] \beta\left(\lambda_{1}\right)
\end{aligned}
$$

This leads to an additional bias of order $1 / T$ in the leading term of the test statistic (7), although it does not introduce bias into the term $R_{1}-R_{2}$ (for example under the $T$ consistent break date estimator). However, correction of the test statistic to this new bias becomes more complicated than the correction undertaken in the previous section and we leave this problem open for future research. We investigate the effect of this bias on the power by simulations.

## 4 Finite sample properties

In this section we investigate the finite sample behavior of the tests. First, consider following DGP:

$$
\begin{equation*}
y_{t}=d_{t}^{\prime} \beta+u_{t}, u_{t}=\phi_{1} u_{t-1}+\phi_{2} u_{t-2}+\varepsilon_{t} \tag{14}
\end{equation*}
$$

i.e., we allow $\operatorname{AR}(2)$ errors in DGP. Here, we assume $\varepsilon_{t} \sim i . i . d . N(0,1)$ and set $\phi_{2}=0.3$ and -0.3 as in KT , and set $\phi_{1}$ such, that the value $\phi_{1}+\phi_{2}$ ranges from 0.5 to 1 . The break fraction $\lambda_{1}$ is assumed to be 0.5 and the parameters of deterministic component are $\mu_{0}=1, \mu_{1}=1, \beta_{0}=0.2$ and $\beta_{1}=0.02$. The sample size is $T=150$, the significance level is 0.05 and 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 a correction factor, obtained under structural change in Section 2 (marked BC in graphs). The second is the SPC test with AR (1) pre-whitening (marked as SPC). Also, for comparison, we consider a KPSS test without the correcting factor $\hat{b}_{T}$ (marked as NC), and a KPSS test with break, proposed by Kurozumi (2002) bandwidth:

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

with $k=$ boundary value (marked as K ). In all cases, the lag length is selected using BIC with maximum lag length $p_{\max }=\left[12(T / 100)^{1 / 4}\right]$, where $[\cdot]$ denotes the integer part.

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 various boundary values $1-c / \sqrt{T}=0.7,0.8$ and 0.9 . The horizontal axis corresponds to the value $\phi_{1}+\phi_{2}$. Specific boundary value implies that if it is lower than $\phi_{1}+\phi_{2}$, then the null hypothesis of $I(0)$ would tend to be rejected, while for $\phi_{1}+\phi_{2}<1-c / \sqrt{T}$ the size of test will be close to the nominal. In Figures 1(a)-(c) parameter $\phi_{2}=-0.3$, in Figures 1(d)-(f) parameter $\phi_{2}=0.3$. In all cases, the
test proposed by Kurozumi (2002) has serious size distortions (and these distortions have nonmonotonic behavior in the case $\phi_{1}=0.3$ ) to the left of the boundary value and much lower power in comparison with the other tests. For $\phi_{2}=-0.3$ and small boundary values ( 0.7 ) the tests BC , NC and SPC have similar properties, but the behavior of SPC and NC considerably deteriorate with increasing boundary value. For $\phi_{2}=0.3$, the bias-corrected test is slightly oversized around the boundary value, but is much more powerful than the SPC and NC. For large boundary values (0.9), the power of the SPC and NC become less than 0.05 , while the bias-corrected test has power close to 0.5 .

Figure 2 shows the size and power of tests for unknown break date, estimated by the Harvey and Leybourne (2013) procedure. In this case, the size of all tests becomes larger, but for SPC and NC the size distortions are more pronounced. Also, the power of all tests is lower than in the case of known break date.

Now consider a situation when the errors $u_{t}$ follow the $\operatorname{ARMA}(1,1)$ process, i.e.

$$
\begin{equation*}
u_{t}=\alpha u_{t-1}+\varepsilon_{t}-\theta \varepsilon_{t-1}, \tag{15}
\end{equation*}
$$

where parameter $\theta$ in the MA component takes values $\{-0.8,-0.4,0.0,0.4,0.8\}$ and other parameters take the same values as above for the $\operatorname{AR}(2)$ process. If $\theta \neq 0$, then the measure of the strength of persistence $\phi_{\infty}(1)=\phi_{1}+\cdots+\phi_{p}+\ldots$ should be based on the autoregressive representation of the series. In this study we rewrite ARMA(1,1) DGP for $u_{t}$ as

$$
(1-\alpha L)(1-\theta L)^{-1} u_{t}=\varepsilon_{t} .
$$

Then, $(1-\alpha L)(1-\theta L)^{-1}=1-\phi(L)$, and the measure of the strength of persistence $\phi_{1}+$ $\cdots+\phi_{p}+\cdots=\phi_{\infty}(1)=(\alpha-\theta) /(1-\theta)$. Figure 3 (the case of known break date) and Figure 4 (the case of unknown break date) show the size and power of the BC, NC, SPC and K tests when the error term follows the $\operatorname{ARMA}(1,1)$ model. The horizontal axis corresponds to the value $\phi_{\infty}(1)=(\alpha-\theta) /(1-\theta)$. We set the boundary value equal to 0.8 and we expect that the null hypothesis of $I(0)$ would tend to be rejected if the strength of persistence will be greater than 0.8 . For all $\theta$, the K test is clearly dominated by all other tests. For $\theta=-0.8$, the size of BC is closer to the nominal one for all $\phi(1)$ and the power of BC, NC and SPC tests is almost the same. For $\theta=-0.4$ the BC test has a slightly higher size for $\phi(1)$ (close to 0.8 ) than the NC and SPC, however it is more powerful. For $\theta=0$, the size of SPC and NC is much lower than the nominal one. Hence, these tests lose power considerably, while the BC test is slightly oversized and has more power. For $\theta=0.8$ and $\theta=0.4$ all tests have serious size distortions, although for $\theta=0.4$ the bias-corrected test is more powerful than the other tests.

In the case of an unknown break date (Figure 4) the results are similar to the $\operatorname{AR}(2)$ case, i.e. tests NC and SPC are more oversized than BC, and all tests have lower power. It should be noted that in the case of a highly negative MA components $(\theta=0.8)$, the size and power of all tests is lower than in the case of a known break date.

We also observe that the proposed bias-corrected test over-corrects the bias term in some cases, so our test suffers from over-size distortion. However, these size distortions disappear with larger $T$, as in KT.

## 5 Conclusion

In this paper we considered the extension of the test proposed by Kurozumi and Tanaka (2010) in the case of a single structural break. Using the generalization of the boundary rule in SPC with $\operatorname{AR}(1)$ pre-whitening, we found the finite sample bias of the numerator in case of structural change occurring in the DGP. Results are similar to the case where the break is absent. In simulation analysis the case of unknown break date has been considered using the approach proposed by Harvey and Leybourne (2013). Finite sample results indicate superiority of the obtained modification in the presence of the break due to better size control allowed by this test. Therefore, continued use of the bias-corrected KPSS test should be considered in empirical applications.

## Appendix

Proof of Theorem $\mathbf{1}^{5}$. Consider the most general case with $d_{t}^{\prime}=\left(1, t, D U_{t}, D T_{t}\right)$. As in KT , we decompose $R_{1}$ into three terms:

$$
R_{1}=R_{11}-R_{12}+R_{13},
$$

where

$$
\begin{aligned}
& R_{11}=\frac{1}{T^{2}} \sum_{t=1}^{T}\left(\sum_{s=1}^{t} \Delta v_{s}\right)^{2} \\
& R_{12}=\frac{2}{T^{2}}\left(\sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} \Delta v_{t} \\
& R_{13}=\frac{1}{T^{2}} \sum_{t=1}^{T} \Delta v_{t} d_{t}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{T} \sum_{s=1}^{t} d_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} \Delta v_{t}
\end{aligned}
$$

Notice that $\sum_{s=1}^{t} \Delta v_{s}=v_{t}-v_{0}$, then (see KT)

$$
\begin{equation*}
\mathbb{E}\left[R_{11}\right]=\frac{2}{T} \gamma_{0}+O\left(T^{-2}\right) \tag{16}
\end{equation*}
$$

The second term is expressed as:

$$
\begin{equation*}
\mathbb{E}\left[R_{12}\right]=\frac{2}{T^{2}} \operatorname{tr}\left\{\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \mathbb{E}\left[\left(\sum_{t=1}^{T} d_{t} \Delta v_{t}\right)\left(\sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\right]\right\} \tag{17}
\end{equation*}
$$

For proof we use the following results:

$$
\sum_{t=1}^{T} d_{t} \Delta v_{t}=\left[\begin{array}{c}
\sum_{t=1}^{T} \Delta v_{t}  \tag{18}\\
\sum_{t=1}^{T} t \Delta v_{t} \\
\sum_{t=1}^{T} D U_{t} \Delta v_{t} \\
\sum_{t=1}^{T} D T_{t} \Delta v_{t}
\end{array}\right]=\left(\begin{array}{c}
v_{T}-v_{0} \\
(T+1) v_{T}-v_{0}-\sum_{t=1}^{T} v_{t} \\
v_{T}-v_{T_{1}} \\
\left(T+1-T_{1}\right) v_{T}-v_{T_{1}}-\sum_{t=T_{1}+1}^{T} v_{t}
\end{array}\right),
$$

[^4]\[

$$
\begin{gather*}
\left(\sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)=\left(\begin{array}{c}
\sum_{t=1}^{T} t\left(v_{t}-v_{0}\right) \\
\sum_{t=1}^{T}\left(v_{t}-v_{0}\right) \sum_{s=1}^{t} s \\
\sum_{t=1}^{T}\left(v_{t}-v_{0}\right) \sum_{s=1}^{t} D U_{s} \\
\sum_{t=1}^{T}\left(v_{t}-v_{0}\right) \sum_{s=1}^{t} D T_{s}
\end{array}\right)^{\prime},  \tag{19}\\
\sum_{t=1}^{T} \sum_{s=1}^{t} D U_{s}=\frac{T^{2}\left(1-\lambda_{1}\right)^{2}}{2}+o\left(T^{2}\right),  \tag{20}\\
\sum_{t=1}^{T} \sum_{s=1}^{t} D T_{s}=\frac{T^{3}\left(1-\lambda_{1}\right)^{3}}{3}+o\left(T^{3}\right)  \tag{21}\\
\sum_{t=1}^{T} t D U_{t}=\frac{\left(T+T_{1}+1\right)\left(T-T_{1}\right)}{2}  \tag{22}\\
\sum_{t=1}^{T} t D T_{t}=\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T+T_{1}+1\right)}{6}  \tag{23}\\
\sum_{t=1}^{T} D T_{t} D T_{t}=\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T-2 T_{1}+1\right)}{6} \tag{24}
\end{gather*}
$$
\]

Using (22)-(24) it can be shown that
$\sum_{t=1}^{T} d_{t} d_{t}^{\prime}=$

$$
\left(\begin{array}{cccc}
T & \frac{T(T+1)}{2} & T-T_{1} & \frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{2}  \tag{25}\\
\frac{T(T+1)}{2} & \frac{T(T+1)(2 T+1)}{6} & \frac{\left(T-T_{1}\right)\left(T+T_{1}+1\right)}{2} & \frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T+T_{1}+1\right)}{6} \\
T-T_{1} & \frac{\left(T-T_{1}\right)\left(T+T_{1}+1\right)}{2} & T-T_{1} & \frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{2} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{2} & \frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T+T_{1}+1\right)}{6} & \frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{2} & \frac{\left(2 T-2 T_{1}+1\right)\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{6}
\end{array}\right) .
$$

Then
$\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}=$

Consider the expectation of right hand side of (17) using (18) and (19). The expectation of $(1,1)$ element equal to (see KT)

$$
\begin{equation*}
\mathbb{E}[(1,1) \text { element }]=\frac{T^{2}}{2} \gamma_{0}+o\left(T^{2}\right) \tag{27}
\end{equation*}
$$

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

$$
\begin{equation*}
\mathbb{E}[(1,2) \text { element }]=\frac{T^{3}}{6} \gamma_{0}+o\left(T^{3}\right) \tag{28}
\end{equation*}
$$

Consider ( 1,3 ) element.

$$
\begin{align*}
\mathbb{E}[(1,3) \text { element }] & =\mathbb{E}\left[\left(v_{T}-v_{0}\right) \sum_{t=1}^{T}\left(v_{t}-v_{0}\right) \sum_{s=1}^{t} D U_{s}\right]  \tag{29}\\
& =\sum_{t=1}^{T} \gamma_{T-t} \sum_{s=1}^{t} D U_{s}-\gamma_{T} \sum_{t=1}^{T} \sum_{s=1}^{t} D U_{t}-\sum_{t=1}^{T} \gamma_{t} \sum_{s=1}^{t} D U_{t}+\gamma_{0} \sum_{t=1}^{T} \sum_{s=1}^{t} D U_{t} \\
& =\frac{\left.T^{2}\left(1-\lambda_{1}\right)^{2}\right)}{2} \gamma_{0}+o\left(T^{2}\right)
\end{align*}
$$

as $\gamma_{t}$ absolutely summable and $\gamma_{T}=o(1)$, and also according to equation (20). Similarly expectation of $(1,4)$ element (we use (21)):

$$
\begin{equation*}
\mathbb{E}[(1,4) \text { element }]=\frac{\left.T^{3}\left(1-\lambda_{1}\right)^{3}\right)}{6} \gamma_{0}+o\left(T^{3}\right) \tag{30}
\end{equation*}
$$

As in $\mathrm{KT}(2,1)$ and $(2,2)$ elements are $o\left(T^{3}\right)$ and $o\left(T^{4}\right)$, respectively. Clearly, $(2,3)$ and $(2,4)$ elements have the same orders as $o\left(T^{3}\right)$ and $o\left(T^{4}\right)$, respectively, because $D U_{t}$ and $D T_{t}$ have the same orders as the constant and the trend, respectively.

Elements of third row of the matrix differ, so that in corresponding expressions $\gamma_{0}$ replaces $\gamma_{T_{1}}$, but as $T_{1}>p$ then $\gamma_{T_{1}}=o(1)$. Then, elements of third row are $o\left(T^{2}\right), o\left(T^{3}\right), o\left(T^{2}\right), o\left(T^{3}\right)$. Elements of fourth row are equal to the corresponding elements of the second row, as they have the same order. Hence, we can write the expectation as

$$
\begin{align*}
& \mathbb{E}\left[\left(\sum_{t=1}^{T} d_{t} \Delta v_{t}\right)\left(\sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\right]= \\
& \left(\begin{array}{cccc}
\frac{T^{2}}{2} \gamma_{0}+o\left(T^{2}\right) & \frac{T^{3}}{6} \gamma_{0}+o\left(T^{3}\right) & \frac{T^{2}\left(1-\lambda_{1}\right)^{2}}{2} \gamma_{0}+o\left(T^{2}\right) & \frac{T^{3}\left(1-\lambda_{1}\right)^{3}}{6} \gamma_{0}+o\left(T^{3}\right) \\
o\left(T^{3}\right) & o\left(T^{4}\right) & o\left(T^{3}\right) & o\left(T^{4}\right) \\
o\left(T^{2}\right) & o\left(T^{3}\right) & o\left(T^{2}\right) & o\left(T^{3}\right) \\
o\left(T^{3}\right) & o\left(T^{4}\right) & o\left(T^{3}\right) & o\left(T^{4}\right)
\end{array}\right) \tag{31}
\end{align*}
$$

Then, using (26) and (31) we obtain:

$$
\begin{equation*}
\mathbb{E}\left[R_{12}\right]=2 \frac{-\lambda_{1}^{2} \gamma_{0}+3 \lambda_{1}^{3} \gamma_{0}-3 \lambda_{1}^{4} \gamma_{0}+\lambda_{1}^{5} \gamma_{0}}{T\left(\lambda_{1}-1\right)^{3} \lambda_{1}^{2}}+o\left(T^{-1}\right)=\frac{2}{T} \gamma_{0}+o\left(T^{-1}\right) \tag{32}
\end{equation*}
$$

Consider the third term of $R_{13}$ :

$$
\begin{equation*}
\mathbb{E}\left[R_{13}\right]=\frac{1}{T^{2}} \operatorname{tr}\left\{\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{T} \sum_{s=1}^{t} d_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \mathbb{E}\left[\sum_{t=1}^{T} d_{t} \Delta v_{t} \sum_{t=1}^{T} \Delta v_{t} d_{t}^{\prime}\right]\right\} \tag{33}
\end{equation*}
$$

It can be shown, using (18) and the fact that $\gamma_{T_{1}}$ and $\gamma_{T-T_{1}}$ are $o(1)$, that

$$
\begin{align*}
& \mathbb{E}\left[\sum_{t=1}^{T} d_{t} \Delta v_{t} \sum_{t=1}^{T} \Delta v_{t} d_{t}^{\prime}\right]= \\
& \left(\begin{array}{cccc}
2 \gamma_{0}+o(1) & T \gamma_{0}+o(T) & \gamma_{0}+o(1) & T\left(1-\lambda_{1}\right) \gamma_{0}+o(T) \\
T \gamma_{0}+o(T) & T^{2} \gamma_{0}+o\left(T^{2}\right) & T \gamma_{0}+o(T) & T^{2}\left(1-\lambda_{1}\right) \gamma_{0}+o\left(T^{2}\right) \\
\gamma_{0}+o(1) & T \gamma_{0}+o(T) & 2 \gamma_{0}+o(1) & T\left(1-\lambda_{1}\right) \gamma_{0}+o(T) \\
T\left(1-\lambda_{1}\right) \gamma_{0}+o(T) & T^{2}\left(1-\lambda_{1}\right) \gamma_{0}+o\left(T^{2}\right) & T\left(1-\lambda_{1}\right) \gamma_{0}+o(T) & T^{2}\left(1-\lambda_{1}\right)^{2} \gamma_{0}+o\left(T^{2}\right)
\end{array}\right) . \tag{34}
\end{align*}
$$

Consider the component $D D:=\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{T} \sum_{s=1}^{t} d_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}$. We can show that

$$
\left(\sum_{t=1}^{T} \sum_{s=1}^{t} d_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)=\left(\xi_{1}, \xi_{2}, \xi_{3}, \xi_{4}\right)
$$

where

$$
\begin{aligned}
& \xi_{1}=\left(\begin{array}{c}
\frac{T(T+1)(2 T+1)}{6} \\
\frac{T(T+1)(T+2)(3 T+1)}{24} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T+T_{1}+1\right)}{6} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)\left(3 T+T_{1}+1\right)}{24}
\end{array}\right), \\
& \xi_{2}=\left(\begin{array}{c}
\frac{T(T+1)(T+2)(3 T+1)}{24} \\
\frac{T(T+1)(T+2)\left(3 T^{2}+6 T+1\right)}{60} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(3 T^{2}+2 T T_{1}+7 T+T_{1}^{2}+3 T_{1}+2\right)}{24} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)\left(6 T^{2}+3 T T_{1}+12 T+T_{1}^{2}+3 T_{1}+2\right)}{120}
\end{array}\right), \\
& \xi_{3}=\left(\begin{array}{c}
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(2 T+T_{1}+1\right)}{6} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(3 T^{2}+2 T T_{1}+7 T+T_{1}^{2}+3 T_{1}+2\right)}{24} \\
\frac{\left(2 T-2 T_{1}+1\right)\left(T-T_{1}\right)\left(T-T_{1}+1\right)}{6} \\
\frac{\left(3 T-3 T_{1}+1\right)\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)}{24}
\end{array}\right), \\
& \xi_{4}=\left(\begin{array}{c}
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)\left(3 T+T_{1}+1\right)}{24} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)\left(6 T^{2}+3 T T_{1}+12 T+T_{1}^{2}+3 T_{1}+2\right)}{120} \\
\frac{\left(3 T-3 T_{1}+1\right)\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)}{24} \\
\frac{\left(T-T_{1}\right)\left(T-T_{1}+1\right)\left(T-T_{1}+2\right)\left(3 T^{2}-6 T T_{1}+6 T+3 T_{1}^{2}-6 T_{1}+1\right)}{60}
\end{array}\right) .
\end{aligned}
$$

Thus using (26),

$$
\begin{equation*}
D D=\left(\kappa_{1}, \kappa_{2}, \kappa_{3}, \kappa_{4}\right) \tag{35}
\end{equation*}
$$

where

$$
\kappa_{1}=\left(\begin{array}{c}
\frac{15 T T_{1}^{2}-15 T T_{1}+2 T_{1}^{3}+22 T_{1}^{2}-8 T_{1}+2}{15\left(T_{1}-1\right) T_{1}} \\
-\frac{11 T_{1}^{2}-5 T_{1}+6}{10\left(T_{1}-1\right) T_{1}} \\
-\frac{\left(T_{1}-2\right)\left(T_{1}+2\right)}{30 T_{1}} \\
\frac{\left(T_{1}+2\right)\left(T_{1}+3\right)}{10\left(T_{1}-1\right) T_{1}}
\end{array}\right)
$$

$$
\begin{aligned}
& \kappa_{2}=\left(\begin{array}{c}
-\frac{11 T_{1}^{2}-5 T_{1}+6}{10\left(T_{1}-1\right) T_{1}} \\
\frac{6\left(T_{1}^{2}+1\right)}{5\left(T_{1}-1\right) T_{1}\left(T_{1}+1\right)} \\
-\frac{\left(T_{1}-3\right)\left(T_{1}-2\right)}{10 T_{1}\left(T_{1}+1\right)} \\
-\frac{6\left(T_{1}^{2}+1\right)}{5\left(T_{1}-1\right) T_{1}\left(T_{1}+1\right)}
\end{array}\right), \\
& \kappa_{3}=\left(\begin{array}{c}
-\frac{\left(T_{1}-2\right)\left(T_{1}+2\right)}{30 T_{1}} \\
-\frac{\left(T_{1}-3\right)\left(T_{1}-2\right)}{10 T_{1}\left(T_{1}+1\right)} \\
\frac{2 T\left(T^{2} T_{1}^{2}+T^{2} T_{1}-2 T T_{1}^{3}-3 T T_{1}^{2}+7 T T_{1}-T+T_{1}^{4}+2 T_{1}^{3}-7 T_{1}^{2}+2 T_{1}+1\right)}{15 T_{1}\left(T_{1}+1\right)\left(T-T_{1}-1\right)\left(T-T_{1}\right)} \\
-\frac{3 T\left(T T_{1}-T-T_{1}^{2}+2 T_{1}+1\right)}{5 T_{1}\left(T_{1}+1\right)\left(T-T_{1}-1\right)\left(T-T_{1}\right)}
\end{array}\right), \\
& \kappa_{4}=\left(\begin{array}{c}
\frac{\left(T_{1}+2\right)\left(T_{1}+3\right)}{10\left(T_{1}-1\right) T_{1}} \\
-\frac{6\left(T_{1}^{2}+1\right)}{5\left(T_{1}-1\right) T_{1}\left(T_{1}+1\right)} \\
-\frac{3 T\left(T T_{1}-T-T_{1}^{2}+2 T_{1}+1\right)}{5 T_{1}\left(T_{1}+1\right)\left(T-T_{1}-1\right)\left(T-T_{1}\right)} \\
\frac{6 T\left(T^{2} T_{1}^{2}+T^{2}-2 T T_{1}^{3}-4 T T_{1}+T_{1}^{4}+4 T_{1}^{2}-1\right)}{5\left(T_{1}-1\right) T_{1}\left(T_{1}+1\right)\left(T-T_{1}-1\right)\left(T-T_{1}\right)\left(T-T_{1}+1\right)}
\end{array}\right) .
\end{aligned}
$$

We then obtain that

$$
\begin{equation*}
\mathbb{E}\left[R_{13}\right]=\frac{19}{15 T} \gamma_{0}+o\left(T^{-1}\right) \tag{36}
\end{equation*}
$$

Combining (16), (32) and (36) we obtain

$$
\begin{equation*}
\mathbb{E}\left[R_{1}\right]=\frac{19}{15 T} \gamma_{0}+o\left(T^{-1}\right) \tag{37}
\end{equation*}
$$

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

$$
\begin{equation*}
R_{2}=R_{21}-R_{22}-R_{23}+R_{24} \tag{38}
\end{equation*}
$$

where

$$
\begin{aligned}
& R_{21}=\sum_{t=1}^{t} \sum_{s=1}^{T} e_{s} \sum_{s=1}^{t} \Delta v_{s} \\
& R_{22}=\sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} e_{t} \\
& R_{23}=\sum_{t=1}^{T} \sum_{s=1}^{t} e_{s} \sum_{s=1}^{t} d_{s}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} \Delta v_{t} \\
& R_{24}=\sum_{t=1}^{T} e_{t} d_{t}^{\prime}\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{T} \sum_{s=1}^{t} d_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \sum_{t=1}^{T} d_{t} \Delta v_{t}
\end{aligned}
$$

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

$$
\begin{equation*}
\mathbb{E}\left[R_{21}\right]=\sigma_{e}^{2} T \sum_{t=0}^{T-1}\left(1-\frac{t}{T}\right) \tilde{\psi}_{t} \tag{39}
\end{equation*}
$$

Consider the second term:

$$
\begin{equation*}
\mathbb{E}\left[R_{22}\right]=\operatorname{tr}\left\{\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \mathbb{E}\left[\sum_{t=1}^{T} d_{t} e_{t} \cdot \sum_{t=1}^{T} \sum_{s=1}^{t} \Delta v_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right]\right\} \tag{40}
\end{equation*}
$$

Next, we use the Lemma provided in KT with some generalizations for dummy variables:
Lemma 2 Let $f_{t}$ and $g_{t}$ be deterministic sequences for $t=1, \ldots, T$. Then

$$
\begin{gather*}
E\left[\left(\sum_{t=1}^{T} f_{t} e_{t}\right)\left(\sum_{t=1}^{T} g_{t} v_{t}\right)\right]=\sigma_{e}^{2} \sum_{t=0}^{T-1}\left(\sum_{s=1}^{T-t} f_{s} g_{s+t}\right) \tilde{\psi}_{t}  \tag{41}\\
E\left[\left(\sum_{t=1}^{T} f_{t} e_{t}\right)\left(\sum_{t=1}^{T} g_{t} \Delta v_{t}\right)\right]=\sigma_{e}^{2} \sum_{t=0}^{T-1}\left(\sum_{s=1}^{T-t} f_{s} g_{s+t}-\sum_{s=1}^{T-t-1} f_{s} g_{s+t+1}\right) \tilde{\psi}_{t}  \tag{42}\\
\sum_{t=1}^{T} f_{t} \sum_{s=1}^{t} e_{s}=\sum_{t=1}^{T}\left(\sum_{s=t}^{T} f_{s}\right) e_{t} \tag{43}
\end{gather*}
$$

Also, because $f_{t}$ can be dummy variables, being zero up to the moment $T_{1}$, therefore for convenience it is necessary to transform (43) as follows:

$$
\begin{equation*}
\sum_{t=1}^{T} f_{t} \sum_{s=1}^{t} e_{s}=\sum_{t=1}^{T}\left(\sum_{s=T_{1}+1}^{T} f_{s}\right) e_{t}-\sum_{t=1}^{T}\left(\sum_{s=1}^{t} f_{s-1}^{b}\right) e_{t} \tag{44}
\end{equation*}
$$

where the expression in the 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 (41) (see KT), we obtain the following elements of the expectation of the right hand side of (40):

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

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

Using (26) we obtain the final expression of expectation of the second term:

$$
\begin{equation*}
\mathbb{E}\left[R_{22}\right]=\sigma_{e}^{2} T \sum_{t=0}^{T-1} \frac{1}{2}\left(1+f_{1} \frac{t}{T}+f_{2} \frac{t^{2}}{T^{2}}+f_{3} \frac{t^{4}}{T^{4}}\right) \tilde{\psi}_{t}+O(1) \tag{45}
\end{equation*}
$$

where $f_{1}, f_{2}$ and $f_{3}$ are some functions which may depend only on $\lambda_{1}$.
Consider the $R_{23}$ term:

$$
\begin{equation*}
\mathbb{E}\left[R_{23}\right]=\operatorname{tr}\left\{\left(\sum_{t=1}^{T} d_{t} d_{t}^{\prime}\right)^{-1} \mathbb{E}\left[\left(\sum_{t=1}^{T} d_{t} \Delta v_{t}\right)\left(\sum_{t=1}^{T} \sum_{s=1}^{t} e_{s} \sum_{s=1}^{t} d_{s}^{\prime}\right)\right]\right\} \tag{46}
\end{equation*}
$$

Using (43) and (44) we can show that:

$$
\begin{align*}
& \sum_{t=1}^{T} \sum_{s=1}^{t} e_{s} \sum_{s=1}^{t} d_{s}^{\prime}= {\left[\sum_{t=1}^{T}\left(\frac{T^{2}-t^{2}}{2}+O(T)\right) e_{t},\right.} \\
& \sum_{t=1}^{T}\left(\frac{T^{3}-t^{3}}{6}+O\left(T^{2}\right)\right) e_{t}, \sum_{t=1}^{T}\left(\frac{T^{2}\left(1-\lambda_{1}\right)^{2}}{2}+O(T)\right) e_{t}-\sum_{t=1}^{T}\left(\frac{\left(t-\lambda_{1} T\right)^{2}}{2}+O(T)\right) e_{t}, \\
&\left.\sum_{t=1}^{T}\left(\frac{T^{3}\left(1-\lambda_{1}\right)^{3}}{6}+O\left(T^{2}\right)\right) e_{t}-\sum_{t=1}^{T}\left(\frac{\left(t-\lambda_{1} T\right)^{3}}{6}+O\left(T^{2}\right)\right) e_{t},\right] . \tag{47}
\end{align*}
$$

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 the interior sum in (41) and (42) is performed given this fact.

Then, using (42) we obtain the following elements of the expectation of the right hand side of (45):

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

$\mathbb{E}[(1,3)$ element $]=\sigma_{e}^{2} T \sum_{t=0}^{T-1}\left(-\frac{t^{2}}{2}+T t-T \lambda_{1} t\right) \tilde{\psi}_{t}+O(T)$
$\mathbb{E}[(1,4)$ element $]=\sigma_{e}^{2} T \sum_{t=0}^{T-1}\left(\frac{t^{3}}{6}-\frac{T t^{2}}{2}+\frac{1}{2} T \lambda_{1} t^{2}+\frac{T^{2} t}{2}+\frac{1}{2} T^{2} \lambda_{1}^{2} t-T^{2} \lambda_{1} t\right) \tilde{\psi}_{t}+O\left(T^{2}\right)$

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

Thus, using (26) we obtain the final expression for $R_{23}$ :

$$
\begin{equation*}
\mathbb{E}\left[R_{23}\right]=\sigma_{e}^{2} T \sum_{t=0}^{T-1} \frac{1}{2}\left(1+f_{4} \frac{t}{T}+f_{5} \frac{t^{2}}{T^{2}}+f_{6} \frac{t^{4}}{T^{4}}\right) \tilde{\psi}_{t}+O(1) \tag{48}
\end{equation*}
$$

where $f_{4}, f_{5}$ and $f_{6}$ are some functions determining as for (45).
Similarly, we obtain an expression for $R_{24}$ :

$$
\begin{equation*}
\mathbb{E}\left[R_{24}\right]=\operatorname{tr}\left\{D D \times \mathbb{E}\left[\left(\sum_{t=1}^{T} d_{t} \Delta v_{t}\right)\left(\sum_{t=1}^{T} e_{t} d_{t}^{\prime}\right)\right]\right\} \tag{49}
\end{equation*}
$$

where the expression for $D D$ has been obtained in (35). Considering expectation in the right hand side and using (42) we obtain:

$$
\begin{equation*}
\mathbb{E}\left[\left(\sum_{t=1}^{T} d_{t} \Delta v_{t}\right)\left(\sum_{t=1}^{T} e_{t} d_{t}^{\prime}\right)\right]=\sigma_{e}^{2}\left(\eta_{1}, \eta_{2}, \eta_{3}, \eta_{4}\right) \tag{50}
\end{equation*}
$$

where

$$
\begin{gathered}
\eta_{1}=\left(\begin{array}{c}
\sum_{t=0}^{T-1} \tilde{\psi}_{t} \\
\sum_{t=0}^{T-1} t \overline{\tilde{\psi}}_{t}+O(1) \\
0 \\
0
\end{array}\right) \\
\eta_{2}=\left(\begin{array}{c}
\sum_{t=0}^{T-1}(T-t) \tilde{\psi}_{t} \\
\sum_{t=1}^{T-1} \frac{T^{2}-t^{2}}{2} \tilde{\psi}_{t}+O(T) \\
\sum_{t=0}^{T-1}\left(T-\lambda_{1} T\right) \tilde{\psi}_{t}+O(T) \\
\sum_{t=0}^{T-1} \frac{\left(\lambda_{1}-1\right)^{2} T^{2}}{2} \tilde{\psi}_{t}+O(T)
\end{array}\right),
\end{gathered}
$$

$$
\begin{gathered}
\eta_{3}=\left(\begin{array}{c}
\sum_{t=0}^{T-1} \tilde{\psi}_{t} \\
\sum_{t=0}^{T-1}\left(t+\lambda_{1} T\right) \tilde{\psi}_{t}+O(1) \\
\sum_{t=0}^{T-1} \tilde{\psi}_{t} \\
\sum_{t=0}^{T-1} t \overline{\tilde{\psi}}_{t}+O(1)
\end{array}\right), \\
\eta_{4}=\left(\begin{array}{c}
\sum_{t=0}^{T-1}\left(T-\lambda_{1} T-t\right) \tilde{\psi}_{t}+O(1) \\
\sum_{t=0}^{T-1} \frac{T^{2}\left(1+\lambda_{1}\right)^{2}-t^{2}}{} \tilde{\psi}_{t} \\
\sum_{t=0}^{T-1}\left(T-\lambda_{1} T-t\right) \tilde{\psi}_{t}+O(1) \\
\sum_{t=0}^{T-1} \overline{T^{2}\left(1-\lambda_{1}\right)^{2}-t^{2}} \tilde{\psi}_{t}+O(T)
\end{array}\right) .
\end{gathered}
$$

Using obtained expression of $D D$ matrix it can be shown that:

$$
\begin{equation*}
\mathbb{E}\left[R_{24}\right]=\sigma_{e}^{2} T \sum_{t=0}^{T-1} \frac{1}{2}\left(\frac{19}{15}+f_{7} \frac{t}{T}+f_{8} \frac{t^{2}}{T^{2}}\right) \tilde{\psi}_{t}+O(1) \tag{51}
\end{equation*}
$$

where $f_{7}$ and $f_{8}$ are defined as before.
Further, from (39), (45), (48) and (51) we obtain:

$$
\begin{equation*}
\mathbb{E}\left[R_{2}\right]=\frac{2 \sigma_{e}^{2} \psi(1)}{T} \sum_{t=0}^{T-1}\left(\frac{19}{30}+f_{6} \frac{t}{T}+f_{7} \frac{t^{2}}{T^{2}}+f_{8} \frac{t^{4}}{T^{4}}\right) \tilde{\psi}_{t}+o\left(T^{-1}\right) . \tag{52}
\end{equation*}
$$

As $\sum_{j=0}^{\infty}\left|\tilde{\psi}_{j}\right|<\infty$ the sum in (52) converges to $\sum_{j=0}^{\infty}(19 / 30) \tilde{\psi}_{j}$. Also, notice that $\psi(1)=$ $1 / \phi(1)$ and $\sum_{j=0}^{\infty} \tilde{\psi}_{t}=\psi^{\prime}(1)=\left(\frac{1}{\phi(1)}\right)^{\prime}=-\frac{\phi^{\prime}(1)}{\phi^{2}(1)}$ we obtain:

$$
\begin{equation*}
\mathbb{E}\left[R_{2}\right]=-\frac{1}{T} \frac{19}{15} \frac{\sigma_{e}^{2} \phi^{\prime}(1)}{\phi^{3}(1)} \tag{53}
\end{equation*}
$$

## 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-i-Rosselló, A. J. (2005). The KPSS Test with Two Structural Breaks. Spanish Economic Review, 9, 105-127.

Carrion-i-Silvestre, J.L. and Sansó-i-Rosselló, 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. (2013). Break Date Estimation for Models with Deterministic Structural Change. Oxford Bulletin of Economics and Statistics, forthcoming.

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., Phillips, P.C.B., 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., Huang, C.J., and Shin, Y. (1997). On Stationary Tests in the Presence of Structural Breaks. Economics Letters, 55, 165-172.

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 in Palgrave Handbook of Econometrics. 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., Phillips, P.C.B., 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.


Figure 1. Size and power of tests, known break date, $\operatorname{AR}(2)$ case


Figure 2. Size and power of tests, unknown break date, $\operatorname{AR}(2)$ case


Figure 3. Size and power of tests, known break date, $\operatorname{ARMA}(1,1)$ case


Figure 4. Size and power of tests, unknown break date, $\operatorname{ARMA}(1,1)$ case


[^0]:    *We thank Eiji Kurozumi for providing his GAUSS code and for his helpful comments. We also thank two anonymous referees, Vladimir Nosko and Marina Turuntseva for their helpful comments and recommendations to improve earlier version of this paper.
    ${ }^{\dagger}$ E-mail: antonskrobotov@gmail.com

[^1]:    ${ }^{1}$ This may not happen if the local asymptotic is considered.

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

[^3]:    ${ }^{3} \mathrm{KT}$ showed that in case of only a constant, $b_{0}=5 / 3$ and in case of constant and trend, $b_{0}=19 / 15$.
    ${ }^{4}$ Perron and Zhu (2005) show that for Models I and II the estimator of break fraction obtained by minimizing the sum of squared residuals, is only $\sqrt{T}$ consistent under $I(1)$ case. In addition, the estimate of the break fraction for Models 0 and 0 t may not be consistent under $I(1)$ case.

[^4]:    ${ }^{5}$ Some calculations (multiplication, inversion) are carried out using Wolfram Mathematica 8.

