A re-examination of inflation persistence dynamics in OECD countries: A new approach.*

Gabriel Zsurkis[a], João Nicolau[b] and Paulo M. M. Rodrigues[c]

[a] Banco de Portugal and ISEG-Universidade de Lisboa

[b] ISEG-Universidade de Lisboa and CEMAPRE

[c] Banco de Portugal and Nova School of Business and Economics

September 21, 2020

Abstract

This paper introduces a simple and easy to implement procedure to test for changes in persistence. The time-varying parameter that characterizes persistence changes under the alternative hypothesis is approximated by a parsimonious cosine function. The new test is the minimum of a t-statistic computed from a test regression that considers a set of reasonable values for a frequency term that is used to evaluate the time varying properties of persistence. The asymptotic distributions of the new tests are derived and critical values provided. An in-depth Monte Carlo analysis shows that the new procedure has important power gains when compared to the local GLS de-trended Dickey-Fuller type tests introduced by Elliott et al. (1996) under various data generating processes with persistence changes. Moreover, an empirical application to OECD countries' inflation shows that for most series analyzed persistence was high in the first half of the sample and subsequently decreased. These results conform with modern macroeconomic theories that point to changes in inflation dynamics in the early 1980s and also with recent empirical evidence against the $I(1)$-$I(0)$ dichotomy.

Keywords: Nonstationarity, unit roots, inflation, CPI, persistence changes.
JEL classification: C12 (Hypothesis Testing), C22 (Time-Series Models)

*We are grateful to Editor Professor Anindya Banerjee and three anonymous referees for their helpful and constructive comments. Financial support from the Portuguese Science Foundation (FCT) through projects UID/MULTI/00491/2019, PTDC/EGE-ECO/28924/2017, and (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, LISBOA-01-0145-FEDER-022209), POR Lisboa (LISBOA-01-0145-FEDER-007722, LISBOA-01-0145-FEDER-022209) and POR Norte (LISBOA-01-0145-FEDER-022209) is gratefully acknowledged.

1 Introduction

Structural breaks in time series result from the occurrence of exogenous shocks, such as crises or policy decisions, which may have permanent effects on the variables' dynamics. For instance, Perron (1989) showed that the Great Crash of 1929 caused a dramatic decrease in the mean of most aggregate variables of the US economy and that the 1973 oil price shock possibly caused a slope change in the output's trend function which was responsible for the subsequent slowdown of its growth rate.

The persistence of time series is frequently assessed through its order of integration, which is determined from the results of a unit root test. However, since the finite sample power performance of traditional unit root tests is not satisfactory when structural breaks are present in the data, this has led to the development of statistics that allow for changes in the deterministic kernel of the data generation process (Perron, 2005, provides an interesting survey). However, most procedures available consider that structural breaks occur instantaneously, which may not be consistent with the fact that shifts in economic aggregates are influenced by changes in the behavior of a very large number of agents that may not react simultaneously to a given shock.

Enders and Lee (2012) and Rodrigues and Taylor (2012) proposed tests that do not require assumptions about the number of breaks and their exact form. To this end Fourier terms have been used to approximate structural changes of unknown functional form in the deterministic component, reducing the specification problem to the selection of the appropriate frequency components of the Fourier approximation. Since from a spectral frequency perspective structural breaks are usually associated to the zero frequency, low-frequency Fourier terms are commonly employed.

In recent literature, the simple stationary/nonstationary (I(0)/I(1)) dichotomy which is typically considered has been questioned and it has been suggested that certain macroeconomic and financial time series may display changes in persistence over time, i.e., changes from low to high persistence or vice versa. For example, Alogoskoufis and Smith (1991) report that inflation persistence has changed over time in the US and that it seems to be a positive function of the degree of monetary and exchange-rate accommodation;

and DeLong and Summers (1988) found that shocks to real output series of the US and European countries were less persistent in the post-World War II years.

Several approaches to test for changes in persistence have recently been developed. Most procedures are related either to the residual-based tests for stationarity introduced by Kim (2000) or to the sub-sample augmented Dickey-Fuller type tests of Banerjee et al. (1992). The first are based on the ratio of two partial sum processes of the residuals from regressions of the time series of interest on deterministic components before and after a given break date (Busetti and Taylor; 2004 and Harvey et al.; 2006). Since the break date is typically unknown, statistics based on the ratios for all possible break dates, such as the maximum Chow-type test, are considered. Regarding the second class of tests, these are based on the minimum of a sequence of ADF type statistics computed by recursive least squares across changing sub-samples of the data. For instance, Leybourne et al. (2003) extended the work of Banerjee et al. (1992) to allow for local GLS de-trended ADF tests of the null hypothesis of a stable unit root process versus a switch from I(0) to I(1) or vice versa. However, Leybourne et al. (2007) point out that tests for a single change in persistence may not be consistent in processes with multiple changes. To overcome this drawback, they propose doubly-recursive sequences of ADF-type unit root statistics which are valid in the presence of multiple shifts in persistence regardless of the direction of change.[1]

Moreover, a change in persistence, as is typically considered, also originates shifts in the deterministic component of the process (see Section 2). The tests proposed in this paper take this effect into account by local GLS de-trending the data with a time-varying autoregressive parameter. Specifically, the tests introduced approximate parameter changes using a single cosine function, but unlike Leybourne et al. (2007) do not provide estimates for the break dates. However, as the complete sample is used for estimation and not fractions of the sample as with recursive tests, it is expected that these approaches have better power performance than the latter when shifts in the autoregressive parameter are the

[1]Phillips et al. (2011) introduced a closely related methodology based on recursive right tailed ADF tests for the detection of speculative bubbles.

only cause for changes in the trend function.[2] Work on testing for multiple changes in persistence is still scant. This paper looks to contribute to this literature by proposing a simple and easy to implement procedure which allows testing for (multiple) persistence changes.

The remainder of the paper is organized as follows. Section 2 introduces the test procedures and derives their asymptotic distributions under the null and local alternative hypotheses. Section 3 investigates the finite sample properties of the statistics through Monte Carlo simulations. Specifically, the impact of conditional heteroskedasticity, breaks in the innovation variance and serially correlated errors are examined. Section 4 presents an in-depth analysis of inflation data of seven OECD countries: Canada, France, Germany, Italy, Japan, the UK and the US. Section 5 concludes. An online Appendix collects the proofs of all results presented throughout the paper as well as additional Monte Carlo results and figures.

2 Motivation and Proposed Statistic

Consider a model for persistence changes in line with Harvey et al. (2006), i.e.,

$$y_t = \mathbf{x}_t^{'}\boldsymbol{\beta} + u_t \tag{1}$$

$$u_t = \rho_t u_{t-1} + \varepsilon_t, \tag{2}$$

where $\mathbf{x}_t$ is a deterministic kernel (i.e. $\mathbf{x}_t := 1$ or $\mathbf{x}_t := [1, t]^{'}$), $\boldsymbol{\beta}$ is the corresponding vector of parameters and (2) describes the stochastic behavior of y_t. ε_t follows a mean zero process satisfying the α-mixing conditions of Phillips and Perron (1988, p.336) with strictly positive and bounded long-run variance $\omega^2 \equiv \lim_{T\to\infty} \mathrm{E}(\sum_{t=1}^{T} \varepsilon_t)^2$.

[2] Monte Carlo results to support this statement are presented in Tables S.11 and S.12 of the online appendix.

2.1 The Test Procedure

To implement the persistence change test, the time series of interest, y_t, is first locally GLS de-trended using $\tilde{\rho}_t := 1 + \frac{\tilde{c}}{T}\cos(k,t)$, where $\tilde{c}$ is fixed and non-positive, k is fixed[3] and

$$\cos(k,t) := \frac{1 + \cos(2\pi kt/T)}{2} = \cos^2(\pi kt/T). \tag{3}$$

A simple cosine function with a single frequency is used in order to mimic the pattern of the unknown shifts in the autoregressive parameter ρ_t. The de-trended variable is computed as $\hat{u}_{\tilde{c},t} = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{\tilde{\mathbf{c}}}$, where $\mathbf{x}_t = 1$ (demeaned case) or $\mathbf{x}_t = [1, t]'$ (linear trend case), and $\hat{\boldsymbol{\beta}}_{\tilde{c}} = \left(\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}_{\tilde{c},t}{}'\right)^{-1} \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}y_{\tilde{c},t}$, with $y_{\tilde{c},1} = y_1$, $\mathbf{x}_{\tilde{c},1} = \mathbf{x}_1$, and $y_{\tilde{c},t} := y_t - \widetilde{\rho}_t y_{t-1}$ and $\mathbf{x}_{\tilde{c},t} := \mathbf{x}_t - \widetilde{\rho}_t\mathbf{x}_{t-1}$, for $t > 1$. As $\cos(k, t)$ takes values between 0 and 1, this approach can be seen as local GLS de-trending with time-varying weights.

In the second step, the presence of a unit root in $\hat{u}_{\tilde{c},t}$ is investigated considering, for k known and fixed, the t-statistic on ϕ computed from the test regression

$$\Delta\hat{u}_{\tilde{c},t} = \phi\cos(k,t)\hat{u}_{\tilde{c},t-1} + \sum_{j=1}^{p}\delta_j\Delta\hat{u}_{\tilde{c},t-j} + \varepsilon_t, \tag{4}$$

where under the null hypothesis of a unit root, $H_0 : \phi = 0$, and under the alternative $H_A : \phi < 0$.[4] The non-centrality parameter $\tilde{c}$ will assume different values depending on the deterministic component and also on the frequency parameter k considered (see Table 1 in Section 3)[5]. In (4), p denotes the lag truncation order chosen to account for any weak dependence in $\{\varepsilon_t\}$. More generally, when ε_t is a linear process satisfying standard summability and moment conditions, p needs to be such that $1/p + p^3/T \to \infty$ as $T \to \infty$; Said and Dickey (1984) and Chang and Park (2002). Since the difference between the proposed test and DF^{GLS} is only due to the presence of some deterministic terms when k is known, the results of Chang and Park (2002) remain valid in this context.

Changes in persistence impact the conditional and unconditional means as well as the

[3]Note that when $k = 0$, $\tilde{\rho}_t$ corresponds to the representation used by Elliott et al. (1996).
[4]Note that under the null hypothesis no structural breaks in the deterministic kernel are allowed.
[5]Elliott et al. (1996) showed that there is no uniformly most powerful unit root test and proposed choosing $\tilde{c}$ as the value at which the test is tangent to the power envelope at 50%.

unconditional variance. Hence, both de-trending and testing steps are influenced by the autoregressive parameter ρ_t in (2). Moreover, the assumption of parameter constancy ($\rho_t = \rho$), when invalid, seems to favor the null hypothesis of a unit root. Thus, it is important to propose tests that allow for changes in persistence under the alternative hypothesis. Since the number of breaks and its functional form are typically unknown in practice, trigonometric functions have been used to accommodate these features. For instance, Fourier series which are linear combinations of sine and cosine functions are widely applied in this context; Enders and Lee (2012) and Rodrigues and Taylor (2012).

In our framework, we use a single factor since the increase in flexibility of the functions employed to describe parameter changes has been associated with a deterioration in the power performance of the tests as a consequence of over-fitting the data; Enders and Lee (2012).

The $\cos(k,t)$ function in (3) is crucial to properly approximate, but not to identify persistence change dates. Its shape is entirely determined by the frequency parameter k. Most empirical work using Fourier terms has only considered integer values for k, which imply that the starting and ending values of $\cos(k,t)$ are the same. For instance, considering $k = 1$, the resulting function may be useful in cases with two breaks, where an increase in ρ_t somewhere in the middle of the sample is followed by a decrease of similar magnitude later. However, when there is an increase in persistence at an unknown point in time and the parameter does not return to its initial value, a fractional frequency needs to be considered (see Figure S.1 in the online appendix for illustration).

To test the null hypothesis, $H_0 : \phi = 0$, in equation (4), when k is unknown (which is the empirically relevant case) the following test statistic is considered,

$$\mathcal{T}_{\hat{k}}^{GLS} := \min_{k \in K} \hat{t}_k^{GLS} = \min_{k \in K} \frac{\sum_{t=2}^{T} \Delta \hat{u}_{\bar{c},t} \cos(k,t) \hat{u}_{\bar{c},t-1}}{\left(\hat{\sigma}_k^2 \sum_{t=2}^{T} \cos^2(k,t) \hat{u}_{\bar{c},t-1}^2\right)^{1/2}}, \tag{5}$$

where $K = \{0.5, 1, 1.5, 2, 2.5, 3\}$ and $\hat{\sigma}_k^2$ is the least-squares estimate of $E(\varepsilon_t^2)$ obtained from (4) under a fixed k. Note that although a time series is not weakly stationary when $k \neq 0$, it does follow an intrinsically mean-reverting process with some exceptional periods during which ρ_t is close to unity.

The approach considered in this paper accommodates possible structural breaks in the deterministic kernel and in the autoregressive parameter of the time series when testing for nonstationarity, however it does not allow for nonlinearity under the alternative. With respect to the latter, there have been significant developments in the unit root testing context in recent years with the introduction of nonlinear unit root tests (see e.g. Kapetanios et al.; 2003; Sollis; 2004; Sollis; 2009; Kruse; 2011 and Caner and Hansen; 2001). Christopoulos and León-Ledesma (2010) propose a unit root test that combines a Fourier kernel and a smooth transition autoregressive nonlinearity component. Our procedure can be straightforwardly generalized to accommodate this type of nonlinearity. For instance, to consider the nonlinear framework of Christopoulos and León-Ledesma (2010) we may replace (4) by one of the following test regressions:

$$\Delta\hat{u}_{\tilde{c},t} = \phi_1 \cos(k, \mathrm{t})\hat{u}_{\tilde{c},t-1}[1 - exp(-\theta\Delta\hat{u}^2_{\tilde{c},t-i})] + \sum_{j=1}^{p}\delta_{1j}\Delta\hat{u}_{\tilde{c},t-j} + \varepsilon_{1t}, \tag{6}$$

where $\theta > 0$ corresponds to the speed of transition between two extreme regimes, $i = 1, ..., L$ and ε_{1t} is a white noise term; and

$$\Delta\hat{u}_{\tilde{c},t} = \phi_2 \cos(k,t)\hat{u}^3_{\tilde{c},t-1} + \sum_{j=1}^{p}\delta_{2j}\Delta\hat{u}_{\tilde{c},t-j} + \varepsilon_{2t}, \tag{7}$$

where ε_{2t} is are white noise innovations. Hence, in line with Christopoulos and León-Ledesma (2010), (6) and (7) in addition to allowing for possible persistence change also allow the adjustment speed to be nonlinear and to follow an exponential smooth transition autoregressive (ESTAR) process, respectively. To compute a test statistics as in (5) from (6) minimization needs to be done over $k \in K$ and $\theta \in \Theta$ where $\Theta = [\underline{\theta}, \overline{\theta}]$ and $0 < \underline{\theta} < \theta < \overline{\theta}$; whereas to compute it from (7) it is only necessary to replace $\cos(k,t)\hat{u}_{\tilde{c},t-1}$ by $\cos(k,t)\hat{u}^3_{\tilde{c},t-1}$ in the numerator and denominator of (5) . Given the relevance of nonlinearity in itself, this merits a detailed analysis, which is the scope of a different paper and will be left for future work.

2.2 Unconditional and Conditional Heteroskedasticity

It is important to examine how the proposed tests perform under breaks in the unconditional or the conditional variance of the error process $\{\varepsilon_t\}$.[6]

Simultaneous increases in persistence and in the innovation variance (two reinforcing effects that cause an increase in σ_y^2) may impact the finite sample performance of the proposed test given that it may be hard to distinguish whether the increase in the unconditional variance of y_t is caused by a true change in persistence or by an exogenous shift in the innovations' variance. Moreover, a large increase in the unconditional variance may cause the process to be confounded more often with a unit root process.

To accommodate (conditional or unconditional) heteroskedasticity, heteroskedasticity-consistent standard errors, as proposed by Eicker-White (EW), are typically employed; Demetrescu (2008) and Phillips (1987). Thus, the proposed test statistic with EW robust standard errors, considering fixed k and no short-run dependence in ε_t is,

$$\hat{t}_{k,EW}^{GLS} := \frac{\sum_{t=2}^{T} \Delta \hat{u}_{\tilde{c},t} \cos(k,t) \hat{u}_{\tilde{c},t-1}}{\left(\sum_{t=2}^{T} \cos^2(k,\mathrm{t}) \hat{u}_{\tilde{c},t-1}^2 \hat{\varepsilon}_t^2\right)^{1/2}}. \tag{8}$$

Proposition 1 *Considering data generated from (1) and (2), and given a fixed k, as $T \to \infty$ it follows: i) under the null hypothesis, $H_0 : \phi = 0$, and considering Assumptions 1 and 2 in Demetrescu (2008), that $\hat{t}_{k,EW}^{GLS} - \hat{t}_k^{GLS} \xrightarrow{p} 0$; and ii) under the alternative hypothesis, $H_1 : \phi = \frac{c}{T}$, for any fixed non-positive c and $\tilde{c}$, that $\hat{t}_{k,EW}^{GLS} - \hat{t}_k^{GLS} \xrightarrow{p} 0$; where " $\xrightarrow{p}$ " represents convergence in probability.*

An alternative to the EW approach used in (8) also widely applied in the literature to deal with, among other things, (unconditional and conditional) heteroskedasticity of unknown form is the Wild bootstrap; Gonçalves and Kilian (2004) and Cavaliere and Taylor (2008). The approach consists of using the residuals $\hat{\varepsilon}_t$ computed from (4) and generating a new unit root process as $\hat{u}_t^b = \hat{u}_{t-1}^b + v_t^b$, where $v_t^b := e_t \hat{\varepsilon}_t$ and

[6] Hamori and Tokihisa (1997) and Kim et al. (2002) showed that a permanent variance shift causes size distortions in the ADF tests; and although conditional heteroskedasticity does not affect the asymptotic distribution of ADF type test statistics (Phillips; 1987), the presence of ARCH effects does cause size distortions in finite samples; Kim and Schmidt (1993).

e_t is such that any heteroskedasticity in $\hat{\varepsilon}_t$ is preserved in the newly created residuals v_t^b. We consider $e_t \sim$i.i.d. $N(0,1)$, but the Rademacher distribution is also frequently used. Next, B bootstrap series of $\hat{u}_t^b$ are generated and in each iteration the bootstrap statistic $\mathcal{T}_k^{GLS^*}$ is computed based on the auxiliary regression $\Delta\hat{u}_t^b = \phi^b\cos(k,t)\hat{u}_{t-1}^b + \sum_{j=1}^{p}\delta_j\Delta\hat{u}_{t-j}^b + \eta_t$ where η_t is an error term. The bootstrap p-value is computed as $P_b\left(\mathcal{T}_k^{GLS}\right) := \frac{1}{B}\sum_{n=1}^{B} I\left(\mathcal{T}_k^{GLS^*} > \mathcal{T}_k^{GLS}\right)$, where B is the number of bootstrap iterations and $I(.)$ is the indicator function (for more details see e.g. Davidson and Flachaire; 2008).

2.3 Asymptotic Distributions

In this section the asymptotic distributions of the proposed tests are derived under the null hypothesis of a unit root and the alternative of local breaks in persistence. Moreover, the test statistics employed in the construction of the asymptotic local power envelope and their asymptotic distributions are also presented.

Theorem 1 *Under the null hypothesis of a unit root, $H_0 : \phi = 0$ ($c = 0$), as $T \to \infty$ the limit distribution of the proposed test statistic in (5) is,*

$$\mathcal{T}_{\hat{k}}^{GLS_{\mathfrak{v}}} \Rightarrow \min_{k\in K}\left(\frac{\cos(k,1)W_{\mathfrak{v}}^2(1) + \frac{1}{2}(2\pi k)^2\int_0^1 \cos(2\pi kr)W_{\mathfrak{v}}^2(r)dr - 1}{2\left(\int_0^1 \cos^2(k,r)W_{\mathfrak{v}}^2(r)dr\right)^{1/2}}\right), \mathfrak{v} = \mu, \tau \qquad (9)$$

where $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ and $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ are computed from local GLS demeaned and local GLS de-trended data, respectively, k is fixed, $W_\mu(r) = W(r)$ is a standard Brownian motion, $\cos(k,r)$ is as defined in (3) with $r := t/T$ and

$$\begin{aligned} W_\tau(r) \quad := \quad & \sigma W(r) - \sigma r\frac{(1-\tilde{c}\cos(k,r))W(1) + \tilde{c}^2\int_0^1 r\cos^2(k,r)W(r)dr}{\int_0^1\left[1 - 2\tilde{c}r\cos^2(k,r) + r^2\tilde{c}^2\cos(k,r)\right]dr} + \\ & + \sigma r\frac{\tilde{c}k\pi\int_0^1 r\sin(2\pi kr)W(r)dr}{\int_0^1\left[1 - 2\tilde{c}r\cos^2(k,r) + r^2\tilde{c}^2\cos(k,r)\right]dr}. \end{aligned}$$

As in the traditional unit root testing context, local GLS demeaning has no effect on the proposed test's asymptotic distribution (see online appendix for details) and therefore, the asymptotic distribution of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ is equivalent to that of a test statistic computed from a test regression with no deterministics.

Theorem 2 *Under the local alternative hypothesis, $H_A : \phi = \frac{c}{T} < 0$, the limit distribution of the proposed statistic is*

$$\mathcal{T}_{\hat{k}}^{GLS_{\mathfrak{v}}} \Rightarrow \min_{k \in K} \left(\frac{\cos(k,1) J_{\mathfrak{v},c}^2(1) + \frac{1}{2}(2\pi k)^2 \int_0^1 \cos(2\pi k r) J_{\mathfrak{v},c}^2(r) dr - 1}{2 \left(\int_0^1 \cos^2(k,r) J_{\mathfrak{v},c}^2(r) dr \right)^{1/2}} \right), \mathfrak{v} = \mu, \tau \qquad (10)$$

where $k \in K$, $r \in (0,1)$, $J_{\mu,c}(r) = J_c(r)$ is a standard Ornstein-Uhlenbeck [OU] process and $J_{\tau,c}(r)$ is a local GLS de-trended OU process.

In order to construct the asymptotic power envelope, an asymptotically equivalent test to the infeasible most powerful invariant likelihood ratio statistic proposed by Elliott et al. (1996) will be used. Specifically, for a given $\tilde{c}$ and k it has the form $P_{\tilde{c}} := \frac{\sum_{t=1}^{T} \hat{\varepsilon}_{\tilde{c},t}^2 - \tilde{\rho}_t \sum_{t=1}^{T} \hat{\varepsilon}_{0,t}^2}{\hat{\sigma}^2}$ where $\hat{\sigma}^2 := T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_{\tilde{c},t}^2$, $\hat{\varepsilon}_{0,t}$ and $\hat{\varepsilon}_{\tilde{c},t}$ are the residuals of the model defined by (1) and (2) considering $\tilde{c} = 0$ and $\tilde{c} < 0$, respectively, and $\tilde{\rho}_t := 1 + \frac{\tilde{c}}{T} \cos(k,t)$.

Theorem 3 *Under $H_0 : \phi = 0$ $(c = 0)$ and i.i.d. innovations the asymptotic distribution of the $P_{\tilde{c}}$ test statistic is $P_{\tilde{c}}^{\mu} \Rightarrow \tilde{c}^2 \int_0^1 \cos^2(r,k) J_c^2(r) - \tilde{c} \cos(T,k) J_c^2(1)$ and $P_{\tilde{c}}^{\tau} \Rightarrow \tilde{c}^2 \int_0^1 \cos^2(r,k) J_{\tau,c}^2(r) + (1 - \tilde{c} \cos(T,k)) J_{\tau,c}^2(1)$, where $P_{\tilde{c}}^{\mu}$ and $P_{\tilde{c}}^{\tau}$ are demeaned and de-trended test statistics, respectively, $J_c(r)$ is a standard OU process and $J_{\tau,c}(r)$ is a local GLS de-trended OU process.*

3 Monte Carlo Analysis

This section investigates the finite sample properties of the tests previously introduced under the null and alternative hypotheses. All simulations are performed in Gauss 10.

Table 1 presents values for $\tilde{c}$, given a specific k, for which the power of the test is tangent to the power envelope at 50%, and Table 2 provides the necessary critical values for $\mathcal{T}_{\hat{k}}^{GLS_{\mu}}$ and $\mathcal{T}_{\hat{k}}^{GLS_{\tau}}$ considering $T \in \{150, 250, 500, 1000\}$ and $k \in \{0.5, 1, 1.5, 2, 2.5, 3\}$.

[Please insert Tables 1 and 2 about here]

In what follows, the finite sample performance of the proposed tests will be compared to that of DF^{GLS_ς}, with $\varsigma = \mu, \tau$. We investigate how the tests perform under iid innovations, in the presence of autocorrelation and under unconditional and conditional heteroskedasticity. For the Wild bootstrap approach 1000 Monte Carlo and bootstrap replications were used, whereas for all other simulations 10,000 Monte Carlo replications were considered.

3.1 IID Innovations

Two data generation processes (DGPs) are considered: i) the first (henceforth DGP1) is, $y_t = \rho_t y_{t-1} + \varepsilon_t$ with $\rho_t = 1 + \phi \cos(k,t)$, $\cos(k,t) := (1 + \cos(2\pi kt/T))/2$ and $\phi \in \{-0.1, -0.2, 0\}$; and ii) the second (DGP2) is,

$$y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ \rho_2 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ \rho_3 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T, \end{cases} \tag{11}$$

where $\tau_1 \leq \tau_2$, $\tau_1 \in \{0.3, 0.4, 0.6, 0.8\}$, $\tau_2 \in \{0.3, 0.6, 0.7, 0.8\}$, $\rho_1 \in \{0.8,\ 0.9\}$, $\rho_2 \in \{0.99,\ 1\}$ and $\rho_3 \in \{0.8, 0.9, 0.99,\ 1\}$ are used to investigate the finite sample power properties, and $\rho_1 = \rho_2 = \rho_3 = 1$ ($\phi = 0$) for the finite sample size of the tests. In both cases, $\varepsilon_t \sim N(0,1)$ and $y_1 = \varepsilon_1 \sim N(0,1)$ is used.

DGP1 implies that the transition from regimes with $\phi < 0$ ($\rho_t < 1$) to regimes with $\phi = 0$ ($\rho_t = 1$) is smooth, since ρ_t is a function of ϕ and the time-varying weights defined by $\cos(k,t)$ in (3). This cosine function is approximately 1 for the first values of t, so that ρ_t is smaller at the beginning of the sample. Table 3 presents the rejection rates of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$. The results show that the empirical size is close to the nominal 5% significance level for all cases considered. In contrast to DF^{GLS_μ} and when there are breaks in persistence, $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ provides power gains for almost all of the simulation parameters used (the only exception is when $k = 0.5$). As expected, the proposed test displays more difficulties in rejecting the null hypothesis when $\phi = -0.1$, since ρ_t is already large before persistence increases. However, there are also significant power gains in this case which further increase with the sample size (the empirical power of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ is close to 100% for $T = 500$).

For $\phi = -0.1$, $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ only presents power gains for all values of k when $T = 500$. Nonetheless, there are some relevant positive differences relatively to the DF^{GLS_τ} for $k > 0.5$ even when $T = 250$. For $\phi = -0.2$ relevant power gains are observed even when $T = 150$ and $k > 0.5$ (see Table S.4 in the online appendix).

In DGP2, to save space, we only allow for a maximum of two abrupt changes which result in a single period of higher persistence. When $\rho_2 = \rho_3$ the break divides the process into two regimes, with a larger ρ_t in the last sub-period. If $\rho_1 = \rho_3$ the sub-period of higher persistence occurs in the middle of the sample and ρ_t returns to the value assumed at the beginning of the process. The results in Table 4 show, as expected, that the power of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ and DF^{GLS_μ} is lower when $\rho_2 = 1$ in a larger percentage of the sample. For instance, if a persistence change occurs at $\tau_1 = \tau_2 = 0.6$ the time series behaves as a random walk over the last 40% of the sample. When the sample size is moderate ($T = 250$) and a change occurs, $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ displays significant power gains even when $\rho_1 = 0.9$ and $\rho_2 = \rho_3 = 1$. When $\rho_2 = \rho_3$, $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ only provides power gains relative to DF^{GLS_μ} for $T \geq 250$. The differences between the rejection rates of these two tests are maximized when the process exhibits higher persistence for a relevant portion of the sample ($\tau_1 = \tau_2 = 0.6$). The same holds when $\rho_1 = \rho_3$, where power gains are observed when $T = 150$ and $\tau_1 = 0.3$ and $\tau_2 = 0.7$.

When the statistics are de-trended, and the changes in persistence are abrupt and $\rho_2 = \rho_3$, the advantages of using $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ only become clear when the sample size is large. On the other hand, when $\rho_1 = \rho_3$ there are positive differences relative to the DF^{GLS_τ} for all cases (see Table S.5 in the online appendix).

Note that the trigonometric function used considers that persistence is always lower at the beginning of the sample. The symmetric cases can be investigated using the time series in reverse chronological order. This transformation alters the asymptotic distributions of the statistics (see (14)). Since the critical values are very close to those obtained using the normal chronological order, reversing the time series when the process starts with a period of higher persistence leads to results similar to those discussed in this Section and are therefore omitted.

[Please insert Tables 3 to 4 about here]

3.2 Serially Correlated Errors

To investigate the finite sample properties of the tests in the presence of autocorrelation it is assumed that the error process of DGP1 is an ARMA, such as, $\varepsilon_t = \delta\varepsilon_{t-1} + \theta e_{t-1} + e_t$ with $e_t \sim N(0,1)$, $y_1 = \varepsilon_1 \sim N(0,1)$, $\delta \in \{0, 0.3, 0.6\}$ and $\theta \in \{-0.8, -0.4, 0, 0.4, 0.8\}$. The lags of the augmented test regression are chosen using the MAIC information criteria proposed by Ng and Perron (2001).

Table S.1 in the online appendix summarizes the finite sample results for the demeaned tests. Although we also performed simulations for $T = 150$ and $T = 500$, for the sake of space only the results for $T = 250$ are reported as the conclusions are qualitatively the same. $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ displays good finite sample performance. Its empirical size only exceeds the nominal 5% significance level when $\theta < 0$, but even in these cases the results are close to or slightly lower than those of the DF^{GLS_μ}.

To evaluate the power performance we considered $\phi = -0.1$ and that the persistence change is smooth and approximated by cosine functions. The sign of θ also affects the power properties of the proposed test. When $\theta \geq 0$ and $k > 0.5$ there are relevant power gains even when $T = 150$. For larger samples, the power differences relative to DF^{GLS_μ} are more pronounced and occur for all values of k. All tests have more difficulties in rejecting the null hypothesis when $\theta = -0.8$ and $\delta = 0$.

Similarly to the demeaned case, also the empirical size of $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ exceeds the nominal 5% level less than the DF^{GLS_τ} when $\theta < 0$, and the two tests are more undersized when $\theta > 0$ (see Table S.6). This problem is attenuated as the sample size increases. The results show that the power of the tests is low when a linear trend is used and ρ_t is large before the increase in persistence. For $T = 500$ the results improve and the superiority of the proposed test becomes evident especially when $\theta \geq 0$.

3.3 Heteroskedasticity

3.3.1 Unconditional Heteroskedasticity

To evaluate the impact of unconditional heteroscedasticity on the performance of the tests we consider DGP1 with the following specification for the innovation process,

$$\varepsilon_t \sim \begin{cases} N(0,\sigma_1^2) & for \quad t = 1, ..., \lfloor \bar{\tau}_1 T \rfloor \\ N(0,\sigma_2^2) & for \quad t = \lfloor \bar{\tau}_1 T \rfloor + 1, ..., \lfloor \bar{\tau}_2 T \rfloor \\ N(0,\sigma_3^2) & for \quad t = \lfloor \bar{\tau}_2 T \rfloor + 1, ..., T, \end{cases} \tag{12}$$

and $y_1 \sim N(0, \sigma_1^2)$.

It is important to infer how changes in the innovation variance may affect the finite sample properties of the proposed test. To save space, only results based on DGP1 with $\phi = -0.1$ are reported. Two cases were considered: in the first, a single break is allowed which either increases or decreases the innovation variance; and in the second, two breaks in variance are permitted, the first causing an increase in the innovation variance and the second a reduction to the value it assumed before the occurrence of any break. The values considered in the simulations allow for large changes in the unconditional variance.

Size results for demeaned data are reported in Table S.2. Results show that $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ reveals in some cases finite sample size distortions. However, the rejection rates are, overall, not too far from those of the DF^{GLS_μ} test. Over-rejections are more severe when the innovation variance faces a large increase (e.g. $\sigma_2^2 = 4$), even if a negative variation of the same magnitude occurs later. The largest empirical size distortions are observed when breaks occur closer to the end of the sample. For decreases in volatility, there are also small over-rejections of the null, especially when these occur at the beginning of the sample ($\bar{\tau}_1 = 0.3$). Table S.7 presents the rejection rates for $\mathcal{T}_{\hat{k}}^{GLS_\tau}$. Although the empirical size of $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ displays the same patterns reported for $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ the rejection rates under the null are larger.

Table 5 presents the results for the tests when the Wild bootstrap is used and $T = 250$. The empirical size of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ is close to the nominal 5% significance level for all cases investigated. Regarding the power properties, comparing the results in Table 5 with those in Table 3, where the innovation variance is constant, we observe that the power loss is greater for an increase in volatility, e.g. $(\sigma_1, \sigma_2, \sigma_3) = (1, 2, 2)$, with $k = 0.5$ and $\bar{\tau}_1 = 0.7$. When a decrease in the unconditional variance is observed, the power loss is greater for $k > 0.5$ and $\bar{\tau}_1 = 0.3$. Finally, the occurrence of two breaks in the unconditional variance cause more severe power losses when $k = 1$ and the magnitude of the changes is larger ($\sigma_2 = 2$).

The superior power performance of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ relatively to DF^{GLS_μ} also depends on k, $\bar{\tau}_1$ and σ. For instance, when the innovation variance increases, the rejection rates of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ are greater for $k > 0.5$ and the difference relative to DF^{GLS_μ} reaches its maximum for $\bar{\tau}_1 = 0.3$. When the variance decreases, $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ shows better power properties for all k and the superiority relative to DF^{GLS_μ} is more prominent for $\bar{\tau}_1 = 0.7$. On the other hand, if two breaks in the unconditional variance occur, the positive difference in the percentage of rejection relative to the DF^{GLS_μ} test is greater when $k = 1$ and the change in the innovation variance is smaller ($\sigma_2 = 1.5$).

Table S.8 presents the results for $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ with $T = 250$. Power losses (when compared to Table S.4) due to breaks in the variance are similar to those reported for the demeaned case. Regarding the comparison with the DF^{GLS_τ} test, we see that power gains are small when $\phi = -0.1$. In this case, the superior performance of $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ only becomes clear in larger samples ($T = 500$).

[**Please insert Table 5 about here**]

3.3.2 Conditional Heteroskedasticity

Finally, to investigate the finite sample distortions caused by the existence of conditional heteroskedasticity we consider that the innovations, ε_t, in (11) follow a GARCH(1,1) process, such that $\varepsilon_t = e_t\sqrt{h_t}$ with $h_t = \omega + \zeta\varepsilon_{t-1}^2 + \xi h_{t-1}$, $y_1 = \varepsilon_1 \sim N(0, 1)$, $h_1 = 1$, $e_t \sim N(0, 1)$, $\zeta \in \{0.7, 0.8, 0.9\}$, $\xi \in \{0, 0.05, 0.1, 0.2\}$ and $\omega = 1 - \zeta - \xi$, implying an unconditional variance of unity.

Table S.3 reports the empirical size for $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ (results for $\mathcal{T}_{\hat{k}}^{GLS_\tau}$ are provided in Table S.9 of the online Appendix). Results show that this test suffers relevant size distortions in the presence of conditional heteroskedasticity, especially when ζ is large. Its empirical size is slightly larger than that of DF^{GLS_μ} and considerably exceeds the nominal 5% level for all parameter values considered. However, the results using the EW version of the test in (6) and the Wild bootstrap approach presented in Table 6 for $T = 250$ (see Table S.10 for results for $\mathcal{T}_{\hat{k}}^{GLS_\tau}$), show that the empirical sizes are close to the nominal 5% significance level for both demeaned and de-trended cases.

Under the alternative hypothesis relevant power gains are observed relatively to DF^{GLS} when EW is considered. When the Wild bootstrap technique is employed, the superiority relatively to DF^{GLS} is less pronounced for all parameter configurations considered.

[Please insert Table 6 about here]

4 Empirical Application

The strong persistence of inflation has been seen as a "stylized fact" of developed economies. However, in recent years, substantial evidence has emerged suggesting that this conclusion may result from changes in the time series properties of inflation that have not been taken into account. For instance, the inflation target and the willingness to stabilize inflation may vary over time causing inflation series not to return to a constant mean in a linear AR framework.

An important branch of literature uses AR model-based measures, such as, the largest autoregressive root (LAAR) and the sum of the autoregressive coefficients (SARC) to investigate inflation persistence; Levin and Piger (2003) and Taylor (2000). In general, these works conclude that, when the deterministic terms or the autoregressive parameters are allowed to change, inflation persistence is lower and far from that of a random walk process. There is some consensus that inflation persistence decreased since the 1980s (possibly due to preferences for price stability). For time-variation in the autoregressive parameters researchers have tended to rely on split samples or rolling regressions. An alternative approach also used is to estimate the path of the time-varying autoregressive parameter using a state-space model and the Kalman filter; Beechey and Österholm

(2012).

Several works have analysed the order of integration as a measure of inflation persistence and employed stationarity or unit root tests in the analysis. Testing procedures that allow the order of integration to endogenously change over time have also been considered in this context. For instance, Harvey et al. (2006) found evidence, using persistence change tests, that CPI inflation in the US suffered a shift in persistence from $I(1)$ to $I(0)$. Halunga et al. (2009) applied the same tests to UK and US inflation data and reported similar results. To circumvent the single change in persistence, which is a limitation of these tests, the sample is partitioned when a break is found and the test re-applied on each sub-sample. The findings achieved using this approach indicate a first change from $I(0)$ to $I(1)$ in the early 1970s and a subsequent reversion to $I(0)$ in the early 1980s, suggesting that the nonstationary dynamics of inflation lasted only about ten years.

Most of the recently proposed methodologies provide strong evidence against the statement that inflation dynamics is described by a pure $I(1)$ process. In this section, we apply the proposed tests, allowing for multiple changes in persistence, to inflation data for several developed economies. In the empirical analysis we allow for a maximum of three breaks, which seems adequate given the available sample sizes.

In our analysis, we consider inflation persistence as the speed at which inflation converges to equilibrium after a shock. Thus, the parameter $\phi_t := \phi\cos(k,t)$ in the test regression $\Delta\hat{\pi}_t = \phi_t\hat{\pi}_{t-1} + \sum_{j=1}^{p}\delta_j\Delta\hat{\pi}_{t-j} + \varepsilon_t$, where $\hat{\pi}_t$ is the locally GLS demeaned inflation time series, is a reasonable indicator of the persistence dynamics and its statistical significance is tested using the procedure introduced in Section 2, i.e.,

$$\mathcal{T}_{\hat{k}}^{GLS_\mu} = \min_{k\in K} \hat{t}_k^{GLS_\mu},\ K = \{0.5, 1, 1.5, 2, 2.5, 3\}\,. \tag{13}$$

However, as the timing of the breaks influences the power performance of the tests, we also considered the time series in reverse chronological order. The reversed test,

$$\mathcal{T}_{\hat{k},r}^{GLS_\mu} := \min_{k\in K} \hat{t}_{k,r}^{GLS} = \min_{k\in K} \frac{\sum_{t=2}^{T}\Delta\hat{\pi}_{t,r}\cos(k,t)\hat{\pi}_{t-1,r}}{\left[\hat{\sigma}_k^2\sum_{t=2}^{T}\cos^2(k,t)\hat{\pi}_{t-1,r}^2\right]^{1/2}},$$

where $\hat{\pi}_{1,r} = \hat{\pi}_T, \hat{\pi}_{2,r} = \hat{\pi}_{T-1}, ..., \hat{\pi}_{T,r} = \hat{\pi}_1$, has the following asymptotic distribution,

$$\mathcal{T}_{\hat{k},r}^{GLS_\mu} \Rightarrow \min_{k \in K} \frac{-\cos(k,0)W(1)^2 + \frac{1}{2}(2\pi k)^2 \int_0^1 \cos(2\pi k r) W(r)^2 dr - 1}{2\left(\int_0^1 \cos^2(k,r) W(r)^2 dr\right)^{1/2}}. \tag{14}$$

The values of $\tilde{c}$ for a given $k \in K = \{0.5, 1, 1.5, 2, 2.5, 3\}$ and the critical values of the $\mathcal{T}_{\hat{k},r}^{GLS_\mu}$ test are very close to those obtained using the normal chronological order (see Tables 1 and 2) and are omitted for the sake of space, but can be obtained from the authors.

Our sample consists of quarterly CPI data for the G7 countries obtained from OECD Statistics for the period from 1955Q1 to 2018Q2. The quarterly CPI is then used to compute the year-on-year and the quarterly growth rates of the CPI data.[7] Note that when the quarterly growth rate is stationary, the year-on-year growth of the CPI introduces a non-invertible moving average component in the resulting time series, which is responsible for a loss of power of the ADF tests. However, since the year-on-year growth of CPI is relevant for monetary policy decisions (inflation targeting regimes typically observe its evolution) we will also consider this definition.

Figure S.2 and Figure S.3 in the online Appendix display, respectively, the year-on-year and the quarterly growth rates of the CPI time series for the G7 countries. Simple visual inspection suggests that the 1970s was the period with the highest inflation rates for almost all countries. Over these years, which have as a milestone the collapse of Bretton Woods, the world economy faced a period of turbulence in which the option for a highly accommodating monetary policy seems to have triggered an unusual increase in inflation persistence only attenuated in the early 80's with the shift to a more restrictive monetary policy. This option, characterized by the introduction of inflation targeting as a framework for monetary policy, contributed to a long period of low and stable inflation in developed countries. Since the 2008 global financial crisis, price stability is once more a concern but this time because of the risk that inflation could remain too low for too

[7]According to Hassler and Demetrescu (2005), the power performance of the ADF test crucially depends on whether inflation is assumed to be equal to the year-on-year growth or to the quarterly growth of CPI.

long. In order to avoid deflation and to bring inflation back to a desirable level, central banks implemented an expansionary monetary policy, which may have had an impact on inflation persistence. Thus, for the sample period considered in this work, two breaks in persistence may have occurred associated with periods of turbulence and relevant changes in monetary policy.

In the analysis we also consider the unit root test proposed by Rodrigues and Taylor (2012) to infer if allowing for structural breaks in the intercept is sufficient to gather evidence against the null hypothesis. Moreover, Wild bootstrap p-values of the proposed test are also computed in order to prevent that the results are influenced by breaks in the unconditional variance (or even by the presence of GARCH effects).

Table 7 reports the results for the year-on-year growth rate of CPI. Considering the minimum between the proposed test computed in normal and reverse chronological order (the critical values used are in the note to Table 7), the null hypothesis of a unit root is rejected at the 5% significance level for all countries except Japan. For five of the seven countries, the stronger rejections occur when the reverse chronological order is considered. The estimated $\hat{k}$ parameter equals one for all cases, suggesting that the period of stronger persistence occurred somewhere around the middle of the sample. France and Germany are the two exceptions for which using the time series in normal chronological order leads to stronger rejections of the null hypothesis. Here, two breaks in persistence seem to have occurred. One before the middle of the sample and the other at the end of the sample, suggesting that the 2008/2009 financial crisis originated a statistically significant change in inflation persistence.

Table 7 also reports results for the quarterly growth rate of CPI. When the minimum between $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ and $\mathcal{T}_{\hat{k},r}^{GLS_\mu}$ is considered, there is statistical evidence against the unit root hypothesis for all countries at the 5% significance level. In terms of the frequency parameter, $\hat{k}$, associated with the smallest test statistic, there are some differences relative to the conclusions drawn above for the year-on-year growth rate. The selected $\hat{k}$ is the same for France, Italy and the US but differs for the other four countries. For instance, for Japan, the quarterly growth rate of CPI had three periods of stronger persistence, one of them at the beginning of the sample.

The proposed cosine term can only provide an approximation of the breaks in persistence and, for the chosen values of k, it suggests that the persistence parameter only assumes two values: one close to unity when the function reaches its minimum value and a lower persistence value when the function reaches its maximum. Furthermore, considering $\hat{k} = 3$ implies that there were a maximum of three periods of shorter duration with higher persistence in inflation. Given the sample length available this number may be large in some cases. Lastly, simulations show that the power gains achieved by the new test are maximized when k is smaller. Even with these limitations, we found relevant evidence against the unit root hypothesis in the inflation rate series for several industrial countries.

The proposed test provides statistical evidence of breaks in persistence for several countries considering both the year-on-year and the quarterly growth rate of CPI. When the year-on-year growth CPI is considered, $\hat{k} = 1$ is selected for most countries employing the minimum between the proposed test computed in normal and reverse chronological order, which seems to suggest that the difficulty in rejecting the $I(1)$ hypothesis is largely due to the Great Inflation of the 1970s (beyond the limitations of available unit root tests such as, for instance, low power in small samples). Regarding the last ten years, the plausible change in persistence after the crisis seems not to have such a large influence, since the proposed unit root test continues to provide strong rejections of the unit root hypothesis.

[**Please insert Table 7 about here**]

5 Conclusions

In this paper we propose a simple approach to detect potential persistence changes and allow for the possibility of the occurrence of up to three breaks in persistence (note that more breaks can be allowed for if deemed necessary). The unknown shape and timing of the breaks are approximated using a cosine term. The test procedure is based on a one-sided t-statistic where this statistic is minimized over a set of values chosen *a priori* for the frequency parameter k.

We find via Monte Carlo simulations that our proposed test has interesting power performance when compared to e.g. the local GLS de-trended unit root tests when breaks in persistence are present, as well as others.

In addition to the DGPs suggested by the specification of the alternative hypothesis, which implies smooth breaks in persistence, we also investigated the power properties of the proposed test when abrupt breaks in ρ_t in (2) occur and results remain favorable. The power gains relative to the DF^{GLS} test are even greater if increases in persistence induce temporary nonstationary behavior. Moreover, we also performed simulations to investigate the effects of unconditional and conditional heteroskedasticity. Although our proposed test shows some size distortions when the homoskedasticity assumption does not hold, this problem is attenuated using the Wild bootstrap which produces empirical sizes close to the nominal 5% level.

Applications of the proposed test to G7 countries' inflation data provided relevant statistical evidence of breaks in persistence. When the year-on-year growth of CPI in reverse chronological order is used, the null hypothesis of a unit root is rejected for all countries. Comparing these results with those obtained in normal chronological order suggests that the evidence of nonstationarity of the inflation series previously reported in the literature is possibly due to the occurrence of a period of stronger persistence in the first half of the sample.

This paper alerts to the consequences of ignoring the occurrence of breaks in persistence. Most of the work on unit root testing that employed Fourier series to approximate smooth structural breaks has focused on changes in the deterministic parameters only. However, changes in the behavior of economic and financial variables caused by, for instance, exogenous events, shifts in monetary policy or improvements in the available technology may have altered not only the equilibrium value but also the speed of reversion to equilibrium after a shock. As previously mentioned, changes in persistence affect the conditional mean, the unconditional mean and the unconditional variance of the process.

Thus, it is possible that some of the evidence in the literature regarding the occurrence of structural breaks in volatility may have been influenced by persistence changes; Sensier and van Dijk (2003) and McConnell and Perez-Quiros (2000).

References

Alogoskoufis, G. S. and Smith, R. (1991). The Phillips Curve, The persistence of Inflation, and the Lucas Critique: Evidence from Exchange-Rate Regimes, *The American Economic Review* **81**(5): 1254–1275.

Banerjee, A., Lumsdaine, R. L. and Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business & Economic Statistics* **10**(3): 271–287.

Beechey, M. and Österholm, P. (2012). The Rise and Fall of U.S. Inflation Persistence, *International Journal of Central Banking* **8**(3): 55–86.

Busetti, F. and Taylor, A. (2004). Tests of stationarity against a change in persistence, *Journal of Econometrics* **123**(1): 33 – 66.

Caner, M. and Hansen, B. E. (2001). Threshold autoregression with a unit root, *Econometrica* **69**(6): 1555–1596.

Cavaliere, G. and Taylor, A. R. (2008). Testing for a change in persistence in the presence of non-stationary volatility, *Journal of Econometrics* **147**(1): 84 – 98. Econometric modelling in finance and risk management: An overview.

Chang, Y. and Park, J. Y. (2002). On the asymptotics of adf tests for unit roots, *Econometric Reviews* **21**(4): 431–447.

Christopoulos, D. K. and León-Ledesma, M. A. (2010). Smooth breaks and non-linear mean reversion: Post-bretton woods real exchange rates, *Journal of International Money and Finance* **29**(6): 1076 – 1093.

Davidson, R. and Flachaire, E. (2008). The wild bootstrap, tamed at last, *Journal of Econometrics* **146**(1): 162 – 169.

DeLong, J. B. and Summers, L. (1988). How does macroeconomic policy affect output?, *Brookings Papers on Economic Activity* **19**(2): 433–494.

Demetrescu, M. (2008). On the Dickey-Fuller test with White standard errors, *Statistical Papers* **51**(1): 11.

Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996). Efficient tests for an autoregressive unit root, *Econometrica* **64**(4): 813–836.

Enders, W. and Lee, J. (2012). A Unit Root Test Using a Fourier Series to Approximate Smooth Breaks, *Oxford Bulletin of Economics and Statistics* **74**(4): 574–599.

Gonçalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form, *Journal of Econometrics* **123**(1): 89 – 120.

Halunga, A. G., Osborn, D. R. and Sensier, M. (2009). Changes in the order of integration of US and UK inflation, *Economics Letters* **102**(1): 30 – 32.

Hamori, S. and Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift, *Economics Letters* **57**(3): 245 – 253.

Harvey, D. I., Leybourne, S. J. and Taylor, A. R. (2006). Modified tests for a change in persistence, *Journal of Econometrics* **134**(2): 441 – 469.

Hassler, U. and Demetrescu, M. (2005). Spurious Persistence and Unit Roots due to Seasonal Differencing: The Case of Inflation Rates, *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)* **225**(4): 413–426.

Kapetanios, G., Shin, Y. and Snell, A. (2003). Testing for a unit root in the nonlinear star framework, *Journal of Econometrics* **112**(2): 359 – 379.

Kim, J.-Y. (2000). Detection of change in persistence of a linear time series, *Journal of Econometrics* **95**(1): 97 – 116.

Kim, K. and Schmidt, P. (1993). Unit root tests with conditional heteroskedasticity, *Journal of Econometrics* **59**(3): 287 – 300.

Kim, T.-H., Leybourne, S. and Newbold, P. (2002). Unit root tests with a break in innovation variance, *Journal of Econometrics* **109**(2): 365 – 387.

Kruse, R. (2011). A new unit root test against estar based on a class of modified statistics., *Statistical Papers* **52**(1): 71–85.

Levin, A. and Piger, J. (2003). Is inflation persistence intrinsic in industrial economies?, *Working Papers 2002-023*, Federal Reserve Bank of St. Louis.

Leybourne, S., Kim, T.-H., Smith, V. and Newbold, P. (2003). Tests for a change in persistence against the null of difference-stationarity, *The Econometrics Journal* **6**(2): 291–311.

Leybourne, S., Kim, T.-H. and Taylor, A. R. (2007). Detecting Multiple Changes in Persistence, *Studies in Nonlinear Dynamics & Econometrics* **11**(3): 1–34.

McConnell, M. M. and Perez-Quiros, G. (2000). Output Fluctuations in the United States: What Has Changed Since the Early 1980's?, *The American Economic Review* **90**(5): 1464–1476.

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power, *Econometrica* **69**(6): 1519–1554.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Econometrica* **57**(6): 1361–1401.

Phillips, P. C. B. (1987). Time Series Regression with a Unit Root, *Econometrica* **55**(2): 277–301.

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika* **75**(2): 335–346.

Phillips, P. C. B., Wu, Y. and Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?, *International Economic Review* **52**(1): 201–226.

Rodrigues, P. M. M. and Taylor, A. M. R. (2012). The Flexible Fourier Form and Local Generalised Least Squares De-trended Unit Root Tests, *Oxford Bulletin of Economics and Statistics* **74**(5): 736–759.

Sensier, M. and van Dijk, D. (2003). Testing for Volatility Changes in US Macroeconomic Time Series, *Centre for Growth and Business Cycle Research Discussion Paper Series 36*, Economics, The Univeristy of Manchester.

Sollis, R. (2004). Asymmetric adjustment and smooth transitions: a combination of some unit root tests., *Journal of time series analysis* **25**(3): 409–417.

Sollis, R. (2009). A simple unit root test against asymmetric star nonlinearity with an application to real exchange rates in nordic countries, *Economic Modelling* **26**(1): 118–125.

Taylor, J. B. (2000). Low inflation, pass-through, and the pricing power of firms, *European Economic Review* **44**(7): 1389 – 1408.

Table 1: Values for the non-centrality parameter $\tilde{c}$ for different values of k

k	0	0.5	1	1.5	2	2.5	3
$\mathbf{x}_t = 1$	-7.0	-15.6	-11.8	-12.7	-10.7	-11.2	-10.2
$\mathbf{x}_t = [1, t]'$	-13.5	-25.4	-25.8	-26.1	-22.2	-23.3	-20.2

Note: Values computed based on 100,000 replications and T=1000.

Table 2: Critical values

	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$			$\mathcal{T}_{\hat{k}}^{GLS_\tau}$		
T	1%	5%	10%	1%	5%	10%
150	-3.266	-2.695	-2.403	-4.092	-3.589	-3.336
250	-3.192	-2.629	-2.346	-4.008	-3.517	-3.268
500	-3.152	-2.592	-2.303	-3.958	-3.467	-3.215
1000	-3.133	-2.574	-2.285	-3.935	-3.438	-3.189

Notes: The reported critical values are based on 100,000 simulations.

Table 3: Empirical size and power with *iid* errors

DGP: $y_t = y_{t-1} + \phi((1 + \cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0, 1)$

	DF^{GLS_μ}			$\mathcal{T}_{\hat{k}}^{GLS_\mu}$			$DF^{GLS_\mu^*}$			$\mathcal{T}_{\hat{k}}^{GLS_\mu^*}$		
						$T = 150$						
k/ϕ	0	-0.1	-0.2	0	-0.1	-0.2	0	-0.1	-0.2	0	-0.1	-0.2
0	0.052	0.925	0.999	0.051	0.707	0.962	0.053	0.918	0.996	0.048	0.679	0.952
0.5	0.052	0.317	0.523	0.051	0.286	0.654	0.053	0.330	0.547	0.048	0.268	0.640
1.0	0.052	0.356	0.601	0.051	0.375	0.769	0.053	0.349	0.619	0.048	0.366	0.748
1.5	0.052	0.336	0.574	0.051	0.354	0.734	0.053	0.336	0.574	0.048	0.337	0.725
2.0	0.052	0.377	0.658	0.051	0.428	0.811	0.053	0.377	0.659	0.048	0.419	0.803
2.5	0.052	0.361	0.635	0.051	0.420	0.796	0.053	0.365	0.624	0.048	0.400	0.762
3.0	0.052	0.405	0.725	0.051	0.461	0.843	0.053	0.423	0.726	0.048	0.465	0.820
						$T = 250$						
0	0.052	0.996	1.000	0.052	0.942	0.996	0.040	0.935	0.998	0.040	0.935	0.998
0.5	0.052	0.484	0.674	0.052	0.557	0.942	0.040	0.477	0.663	0.040	0.532	0.930
1.0	0.052	0.555	0.765	0.052	0.689	0.967	0.040	0.549	0.762	0.040	0.685	0.961
1.5	0.052	0.524	0.753	0.052	0.657	0.958	0.040	0.526	0.754	0.040	0.628	0.958
2.0	0.052	0.601	0.854	0.052	0.740	0.974	0.040	0.578	0.851	0.040	0.720	0.971
2.5	0.052	0.588	0.839	0.052	0.717	0.967	0.040	0.577	0.832	0.040	0.703	0.955
3.0	0.052	0.666	0.918	0.052	0.774	0.976	0.040	0.661	0.919	0.040	0.752	0.970
						$T = 500$						
0	0.050	1.000	1.000	0.052	0.999	1.000	0.041	1.000	1.000	0.041	1.000	1.000
0.5	0.050	0.683	0.818	0.052	0.961	0.999	0.041	0.684	0.818	0.041	0.950	0.998
1.0	0.050	0.774	0.911	0.052	0.982	1.000	0.041	0.758	0.909	0.041	0.978	1.000
1.5	0.050	0.761	0.913	0.052	0.977	1.000	0.041	0.778	0.923	0.041	0.966	1.000
2.0	0.050	0.859	0.976	0.052	0.988	1.000	0.041	0.852	0.968	0.041	0.984	0.999
2.5	0.050	0.848	0.968	0.052	0.983	1.000	0.041	0.844	0.972	0.041	0.972	1.000
3.0	0.050	0.922	0.995	0.052	0.989	1.000	0.041	0.915	0.992	0.041	0.985	0.999

Note: $DF^{GLS_\mu^*}$ and $\mathcal{T}_{\hat{k}}^{GLS_\mu^*}$ are Wild bootstrap based test statistics.

Table 4: Empirical power when the breaks in persistence are abrupt

$$\text{DGP:} \quad y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ \rho_2 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ \rho_3 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T, \end{cases} \qquad \varepsilon_t \sim N(0,1)$$

	DF^{GLS_μ}		$\mathcal{T}_{\hat{k}}^{GLS_\mu}$		DF^{GLS_μ}		$\mathcal{T}_{\hat{k}}^{GLS_\mu}$		DF^{GLS_μ}		$\mathcal{T}_{\hat{k}}^{GLS_\mu}$	
	CV	WB	CV	WB	CV	WB	CV	WB	CV	WB	CV	WB
$T = 150$												
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.233	0.252	0.287	0.287	0.515	0.541	0.724	0.714	0.783	0.792	0.908	0.889
$(0.8, 1, 1)$	0.148	0.165	0.205	0.211	0.408	0.418	0.661	0.655	0.718	0.723	0.897	0.876
$(0.9, 0.99, 0.99)$	0.194	0.207	0.188	0.191	0.398	0.419	0.373	0.384	0.629	0.650	0.553	0.526
$(0.9, 1, 1)$	0.124	0.135	0.133	0.133	0.314	0.308	0.315	0.317	0.561	0.576	0.525	0.502
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.752	0.768	0.880	0.866	0.743	0.746	0.920	0.896	0.607	0.623	0.865	0.838
$(0.8, 1, 0.8)$	0.667	0.676	0.876	0.863	0.658	0.675	0.924	0.905	0.501	0.509	0.873	0.850
$(0.9, 0.99, 0.9)$	0.569	0.578	0.538	0.520	0.568	0.573	0.574	0.544	0.454	0.457	0.500	0.473
$(0.9, 1, 0.9)$	0.485	0.487	0.534	0.517	0.491	0.498	0.581	0.561	0.362	0.360	0.513	0.490
$T = 250$												
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.342	0.345	0.475	0.452	0.640	0.628	0.940	0.933	0.863	0.873	0.992	0.992
$(0.8, 1, 1)$	0.170	0.157	0.286	0.286	0.449	0.448	0.878	0.856	0.763	0.777	0.989	0.989
$(0.9, 0.99, 0.99)$	0.305	0.316	0.332	0.327	0.569	0.557	0.686	0.644	0.798	0.809	0.857	0.842
$(0.9, 1, 1)$	0.149	0.125	0.189	0.180	0.392	0.402	0.573	0.531	0.687	0.698	0.834	0.815
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.856	0.853	0.982	0.981	0.857	0.855	0.990	0.990	0.745	0.732	0.981	0.977
$(0.8, 1, 0.8)$	0.733	0.729	0.981	0.980	0.737	0.730	0.991	0.987	0.570	0.577	0.983	0.979
$(0.9, 0.99, 0.9)$	0.773	0.769	0.831	0.813	0.775	0.777	0.868	0.852	0.659	0.655	0.804	0.793
$(0.9, 1, 0.9)$	0.635	0.638	0.822	0.810	0.633	0.631	0.879	0.864	0.481	0.475	0.820	0.804
$T = 500$												
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.578	0.569	0.752	0.753	0.824	0.816	0.998	0.998	0.948	0.934	1.000	1.000
$(0.8, 1, 1)$	0.177	0.161	0.369	0.355	0.476	0.472	0.959	0.957	0.801	0.800	1.000	1.000
$(0.9, 0.99, 0.99)$	0.550	0.549	0.644	0.657	0.793	0.789	0.978	0.973	0.928	0.921	0.997	0.997
$(0.9, 1, 1)$	0.165	0.147	0.291	0.269	0.440	0.440	0.881	0.877	0.761	0.756	0.995	0.993
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.961	0.965	1.000	1.000	0.962	0.955	1.000	1.000	0.910	0.909	1.000	1.000
$(0.8, 1, 0.8)$	0.784	0.770	1.000	1.000	0.792	0.785	1.000	1.000	0.615	0.615	1.000	1.000
$(0.9, 0.99, 0.9)$	0.941	0.945	0.995	0.994	0.943	0.935	0.997	0.996	0.878	0.866	0.992	0.990
$(0.9, 1, 0.9)$	0.727	0.723	0.992	0.988	0.733	0.730	0.998	0.997	0.560	0.554	0.993	0.993

Table 5: Empirical size and power of Wild bootstrap based statistics under unconditional variance breaks.

DGP: $y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t,$

$$\varepsilon_t \sim \begin{cases} N(0,\sigma_1^2) & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ N(0,\sigma_2^2) & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ N(0,\sigma_3^2) & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T, \end{cases}$$

	$DF^{GLS_{\mu^*}}$			$\mathcal{T}_{\hat{k}}^{GLS_\mu^*}$		
$T = 250$	$\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 2$					
$(\phi, k)/(\bar{\tau}_1,\ \bar{\tau}_2)$	$(0.3, 0.3)$	$(0.5, 0.5)$	$(0.7, 0.7)$	$(0.3, 0.3)$	$(0.5, 0.5)$	$(0.7, 0.7)$
$(0, 0)$	0.051	0.054	0.055	0.058	0.046	0.044
$(-0.1, 0.5)$	0.372	0.312	0.313	0.323	0.276	0.276
$(-0.1, 1.0)$	0.428	0.575	0.650	0.590	0.654	0.676
$(-0.1, 1.5)$	0.519	0.552	0.463	0.581	0.571	0.423
$(-0.1, 2.0)$	0.604	0.538	0.534	0.689	0.649	0.640
$(-0.1, 2.5)$	0.569	0.510	0.570	0.652	0.597	0.577
$(-0.1, 3.0)$	0.607	0.641	0.589	0.736	0.716	0.654
	$\sigma_1 = 2$, $\sigma_2 = 2$ and $\sigma_3 = 1$					
$(0, 0)$	0.057	0.053	0.041	0.052	0.037	0.049
$(-0.1, 0.5)$	0.624	0.683	0.651	0.683	0.699	0.671
$(-0.1, 1.0)$	0.582	0.452	0.463	0.606	0.552	0.576
$(-0.1, 1.5)$	0.490	0.491	0.583	0.564	0.549	0.685
$(-0.1, 2.0)$	0.486	0.573	0.562	0.591	0.685	0.688
$(-0.1, 2.5)$	0.532	0.613	0.599	0.601	0.670	0.712
$(-0.1, 3.0)$	0.597	0.606	0.670	0.638	0.658	0.733
	$\sigma_1 = 1$, $\sigma_2 = 1.5$ and $\sigma_3 = 1$					
$(\phi, k)/\bar{\tau}$	$(0.3, 0.6)$	$(0.4, 0.7)$	$(0.3, 0.7)$	$(0.3, 0.6)$	$(0.4, 0.7)$	$(0.3, 0.7)$
$(0, 0)$	0.046	0.052	0.056	0.042	0.052	0.050
$(-0.1, 0.5)$	0.505	0.460	0.486	0.520	0.455	0.482
$(-0.1, 1.0)$	0.422	0.458	0.424	0.569	0.588	0.548
$(-0.1, 1.5)$	0.528	0.610	0.566	0.663	0.685	0.671
$(-0.1, 2.0)$	0.667	0.616	0.629	0.754	0.692	0.696
$(-0.1, 2.5)$	0.574	0.550	0.580	0.695	0.679	0.700
$(-0.1, 3.0)$	0.640	0.653	0.667	0.729	0.747	0.749
	$\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 1$					
$(0, 0)$	0.052	0.056	0.052	0.059	0.058	0.057
$(-0.1, 0.5)$	0.505	0.434	0.461	0.509	0.387	0.436
$(-0.1, 1.0)$	0.352	0.401	0.339	0.442	0.482	0.444
$(-0.1, 1.5)$	0.520	0.647	0.581	0.625	0.684	0.667
$(-0.1, 2.0)$	0.707	0.592	0.604	0.727	0.643	0.648
$(-0.1, 2.5)$	0.558	0.514	0.573	0.653	0.618	0.650
$(-0.1, 3.0)$	0.602	0.617	0.649	0.677	0.710	0.740

Table 6: Empirical size and power in the presence of GARCH effects using EW standard errors and Wild bootstrap based test statistics.

DGP:

$$y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t,$$
$$\varepsilon_t = e_t\sqrt{h_t},\ h_t = \omega + \zeta\varepsilon_{t-1}^2 + \xi h_{t-1}$$
$$\omega = 1-\zeta-\xi,\ e_t \sim N(0,1)$$

$T=250$		$DF^{GLS_\mu^{EW}}$			$\mathcal{T}_{\hat{k}}^{GLS_\mu^{EW}}$			$DF^{GLS_\mu^{*}}$			$\mathcal{T}_{\hat{k}}^{GLS_\mu^{*}}$		
	ξ/ζ	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9
$\phi=0$	0	0.051	0.049	0.049	0.058	0.054	0.052	0.046	0.055	0.056	0.06	0.064	0.067
	0.05	0.05	0.05	0.048	0.056	0.055	0.05	0.051	0.055	0.055	0.057	0.063	0.064
	0.1	0.05	0.048	-	0.057	0.052	-	0.052	0.051	-	0.058	0.068	-
	0.2	0.052	-	-	0.053	-	-	0.053	-	-	0.058	-	-
$\phi=-0.1$													
$k=0.5$	0	0.396	0.37	0.336	0.434	0.399	0.348	0.442	0.436	0.413	0.468	0.459	0.428
	0.05	0.386	0.352	0.299	0.422	0.379	0.306	0.442	0.427	0.405	0.473	0.453	0.4
	0.1	0.375	0.335	-	0.407	0.352	-	0.446	0.426	-	0.465	0.44	
	0.2	0.34	-	-	0.36	-	-	0.445	-	-	0.458		
$k=1$	0	0.454	0.422	0.382	0.571	0.531	0.483	0.531	0.518	0.491	0.556	0.528	0.487
	0.05	0.438	0.406	0.351	0.557	0.513	0.432	0.525	0.506	0.469	0.541	0.515	0.448
	0.1	0.426	0.385	-	0.543	0.486	-	0.516	0.499	-	0.533	0.492	-
	0.2	0.39	-	-	0.496	-	-	0.497	-	-	0.513	-	-
$k=1.5$	0	0.429	0.396	0.359	0.52	0.482	0.432	0.523	0.518	0.488	0.551	0.526	0.494
	0.05	0.415	0.38	0.325	0.505	0.46	0.391	0.523	0.514	0.469	0.544	0.515	0.447
	0.1	0.4	0.36	-	0.488	0.437	-	0.526	0.514	-	0.525	0.493	-
	0.2	0.366	-	-	0.449	-	-	0.524	-	-	0.485	-	-
$k=2$	0	0.494	0.459	0.4138	0.613	0.581	0.532	0.556	0.55	0.534	0.637	0.612	0.554
	0.05	0.482	0.44	0.3748	0.603	0.56	0.484	0.553	0.543	0.511	0.627	0.592	0.496
	0.1	0.467	0.418	-	0.589	0.536	-	0.544	0.532	-	0.619	0.551	-
	0.2	0.428	-	-	0.545	-	-	0.538	-	-	0.573	-	-
$k=2.5$	0	0.475	0.435	0.3921	0.603	0.566	0.514	0.544	0.542	0.52	0.621	0.599	0.555
	0.05	0.459	0.417	0.3579	0.589	0.546	0.462	0.546	0.535	0.495	0.618	0.587	0.516
	0.1	0.444	0.395	-	0.576	0.518	-	0.546	0.524	-	0.597	0.569	-
	0.2	0.404	-	-	0.531	-	-	0.519	-	-	0.574	-	-
$k=3$	0	0.533	0.491	0.441	0.66	0.625	0.569	0.613	0.585	0.56	0.684	0.643	0.575
	0.05	0.518	0.469	0.398	0.647	0.603	0.517	0.595	0.582	0.527	0.663	0.621	0.519
	0.1	0.498	0.443	-	0.631	0.577	-	0.596	0.562	-	0.65	0.598	-
	0.2	0.45	-	-	0.587	-	-	0.571	-	-	0.613	-	-

Note: $DF^{GLS_\mu^{EW}}$ and $\mathcal{T}_{\hat{k}}^{GLS_\mu^{EW}}$ are EW based test statistics and, $DF^{GLS_\mu^{*}}$ and $\mathcal{T}_{\hat{k}}^{GLS_\mu^{*}}$ are Wild bootstrap based test statistics.

Table 7: Results for the year-on-year and quarterly growth of the CPI

Year-on-year growth						
Country	DF^{GLS_μ}	$\hat{k}$	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$	WB p-value	k	$t_\alpha^{ERS_f^\mu}$
Canada	−1.040	1.5	−2.900 [b]	0.028	1	−2.276
France	−1.348	1.5	−3.494 [a]	0.016	1	−2.724
Germany	−2.481 [b]	1.5	−3.692 [a]	0.002	1	−3.040 [c]
Italy	−1.719 [c]	1.5	−2.280	0.161	1	−3.109 [c]
Japan	−1.320	2	−2.002	0.230	1	−2.378 [c]
United Kingdom	−1.392	2	−2.446 [c]	0.111	1	−1.951
United States	−1.065	1.5	−2.031	0.235	1	−2.343
reverse chronological order						
Canada	−1.348	1	−3.748 [a]	0.000	1	−2.432
France	−1.109	1	−3.108 [b]	0.045	1	−2.381
Germany	−2.051 [b]	1	−3.583 [a]	0.007	1	−3.104 [c]
Italy	−1.432	1	−3.174 [b]	0.033	1	−3.114 [c]
Japan	−1.647	1	−2.586 [c]	0.091	1	−2.805
United Kingdom	−1.278	1	−2.894 [b]	0.054	1	−2.068
United States	−1.471	1	−4.096 [a]	0.002	1	−2.533
Quarterly growth						
Country	DF^{GLS^μ}	$\hat{k}$	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$	WB p-value	k	$t_\alpha^{ERS_f^\mu}$
Canada	−1.499	1.5	−3.021 [b]	0.016	1	−2.790
France	−1.303	1.5	−3.313 [a]	0.034	1	−2.713
Germany	−0.901	1.5	−1.706	0.325	1	−1.905
Italy	−1.375	2	−2.243	0.170	1	−2.487
Japan	−1.355	2	−1.967	0.219	1	−2.561
United Kingdom	−1.383	2	−3.178 [b]	0.029	1	−2.108
United States	−1.417	1.5	−2.528 [c]	0.235	1	−2.460
reverse chronological order						
Canada	−1.745 [b]	1	−2.697 [b]	0.044	1	−1.887
France	−1.005	1	−3.013 [b]	0.046	1	−2.046
Germany	−1.635	1	−3.125 [b]	0.017	1	−3.156 [b]
Italy	−1.591	1	−3.205 [a]	0.037	1	−2.447
Japan	−1.654	2.5	−3.016 [b]	0.036	1	−3.055 [c]
United Kingdom	−0.988	2.5	−2.516 [c]	0.124	1	−2.079
United States	−1.798 [b]	1	−3.411 [a]	0.007	1	−2.329

Notes: (1) a, b and c denote significance at the 1%, 5% and 10% significance levels, respectively; (2) $t_\alpha^{ERS_f^\mu}$ corresponds to the test proposed by Rodrigues and Taylor (2012); (3) the lags of the unit root tests were chosen using the MAIC information criterion; (4) the values of the critical values of $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ with reverse chronological order and T=250 are -3.198, -2.634 and -2.352 for 1%, 5% and 10% significance levels, respectively; (5) the critical values for the minimum between $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ with normal and reverse chronological order and T=250 are -3.382, -2.839 and -2.568 for 1%, 5% and 10% significance levels, respectively.

Supplementary Appendix:

A reexamination of inflation persistence dynamics in OECD countries: A new approach.

by

September 21, 2020

Proofs of main results

Preliminary Results

Consider first limit results for the local GLS demeaned and local GLS de-trended data.

Case A: Local GLS Demeaning

Under local GLS demeaning we estimate the parameter vector $\boldsymbol{\beta}$ in (1) using $\mathbf{x}_t = 1$. Hence, consider $y_{\tilde{c},1} := y_1$, $y_{\tilde{c},t} := y_t - \tilde{\rho}_t y_{t-1}$, $\mathbf{x}_{\tilde{c},1} := \mathbf{x}_1$, $\mathbf{x}_{\tilde{c},t} := \mathbf{x}_t - \tilde{\rho}_t \mathbf{x}_{t-1}$, and compute the OLS estimates as,

$$\hat{\boldsymbol{\beta}}_{\tilde{c}} = \left[\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}'_{\tilde{c},t},\right]^{-1} \left[\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} y_{\tilde{c},t}\right]. \tag{S.1}$$

Thus, the local GLS demeaned data is, $\hat{u}_{\tilde{c},t} = y_t - \mathbf{x}'_t \hat{\boldsymbol{\beta}}_{\tilde{c}} = u_t - \mathbf{x}'_t(\hat{\boldsymbol{\beta}}_{\tilde{c}} - \boldsymbol{\beta})$ or equivalently,

$$\hat{u}_{\tilde{c},t} = u_t - \mathbf{x}_t \left[\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}'_{\tilde{c},t},\right]^{-1} \left[\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} u_{\tilde{c},t}\right].$$

Since,

$$\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}'_{\tilde{c},t}, = 1 + \left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) = 1 + o(1).$$

and

$$\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} u_{\tilde{c},t} \quad = \quad u_1 - \left(\frac{\tilde{c}}{T}\right) \sum_{t=2}^{T} \cos(k,t)\Delta u_t + \left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \cos^2(k,t) u_{t-1}$$

we establish that,

$$\frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}'_{\tilde{c},t},\right]^{-1} \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} u_{\tilde{c},t} \to 0.$$

Since $\frac{1}{T}\sum_{t=2}^{T}\cos^2(k,t) \to \int_0^1 \cos^2(k,r)dr$ then $\left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \cos^2(k,t) = o(1)$.

Case B: Local GLS De-trending

For local GLS de-trending consider again (S.1) but with the denominator and numerator given as $D_T \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}_{\tilde{c},t}{}' D_T$ and $D_T \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} u_{\tilde{c},t}$, respectively, where $D_T :=$ diag $\left(1, T^{-1/2}\right)$.

Hence,

$$\begin{aligned} D_T \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}_{\tilde{c},t}{}' D_T &= D_T \mathbf{x}_1 \mathbf{x}_1{}' D_T + D_T \sum_{t=2}^{T} \mathbf{x}_{\tilde{c},t}\mathbf{x}_{\tilde{c},t}{}' D_T \\ &= \begin{bmatrix} 1 & T^{-1/2} \\ T^{-1/2} & T^{-1} \end{bmatrix} + D_T \begin{bmatrix} \Xi_1 & \Xi_2 \\ \Xi_2 & \Xi_3 \end{bmatrix} D_T \\ &\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & \int_0^1 \left[1 - 2\tilde{c} r \text{cos}^2(k,r) + r^2 \tilde{c}^2 \text{cos}^2(k,r)\right] dr \end{bmatrix}, \end{aligned} \qquad \text{(S.2)}$$

where, $\mathbf{x}_1 = (1,1)'$ and $\mathbf{x}_{\tilde{c},t} = (-\tilde{c}\text{cos}(k,t)T^{-1}, 1-(t-1)\tilde{c}\text{cos}(k,t)T^{-1})'$ for $t > 1$, with $\Xi_1 := \left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \text{cos}^2(k,t)$, $\Xi_2 := \left(1 - \frac{(t-1)\tilde{c}\text{cos}(k,t)}{T}\right)\frac{\tilde{c}}{T}\text{cos}(k,t)$, and $\Xi_3 := \left(1 - \frac{(t-1)\tilde{c}\text{cos}(k,t)}{T}\right)^2$.

Moreover, note that,

$$\begin{aligned} \frac{1}{T}\sum_{t=2}^{T} \text{cos}(k,t) &\Rightarrow \int_0^1 \text{cos}(k,r)dr; \\ \frac{1}{T^2}\sum_{t=2}^{T} t\text{cos}(k,t) &\Rightarrow \int_0^1 r\text{cos}(k,r)dr; \\ \frac{1}{T^3}\sum_{t=2}^{T} t^2\text{cos}^2(k,t) &\Rightarrow \int_0^1 r^2\text{cos}^2(k,r)dr; \\ T^{-5/2}\sum_{t=3}^{T} t\text{cos}(k,t)u_{t-1} &\Rightarrow \int_0^1 r\text{cos}(k,r)W(r), \qquad 0 \le r \le 1. \end{aligned}$$

Finally,

$$D_T \sum_{t=1}^{T} \mathbf{x}_{\tilde{c},t} u_{\tilde{c},t} = \begin{bmatrix} \Xi_4 \\ \Xi_5 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ \sigma W(1)(1 - \tilde{c}\text{cos}(k,T)) - \sigma\pi k\tilde{c}\int_0^1 r\text{sin}(2\pi kt/T)W(r)dr \\ +\sigma\tilde{c}^2 \int_0^1 r\text{cos}(k,r)W(r)dr \end{bmatrix}$$

where Ξ_4 is defined as in (S.2) and since $u_{\tilde{c},t} = \Delta u_{\tilde{c},t} - \tilde{c}\cos(k,t)T^{-1}$,

$$\Xi_5 = u_1 + u_T - u_1 - \frac{\tilde{c}}{T}\sum_{t=2}^{T}\cos(k,t)u_{t-1} - \frac{\tilde{c}}{T}\sum_{t=2}^{T}(t-1)\cos(k,t)\Delta u_t + \left(\frac{\tilde{c}}{T}\right)^2\sum_{t=2}^{T}(t-1)\cos(k,t)^2 u_{t-1}$$

$$= (1-\tilde{c}\cos(k,\, T))u_T + 2\tilde{c}\cos(k,\, 2)u_1 T^{-1} - \frac{\pi k}{T}\frac{\tilde{c}}{T}\sum_{t=3}^{T} t\sin(2\pi \mathrm{kt/T})u_{t-1} + \frac{\tilde{c}}{T}\sum_{t=3}^{T}\cos(k,\, t-1)u_{t-1} -$$

$$\frac{\tilde{c}}{T}\sum_{t=2}^{T}\cos(k,\, t)u_{t-1} + \left(\frac{\tilde{c}}{T}\right)^2\sum_{t=2}^{T} t\cos(k,\, t)^2 u_{t-1} - \left(\frac{\tilde{c}}{T}\right)^2\sum_{t=2}^{T}\cos(k,\, t)^2 u_{t-1} + \frac{\tilde{c}}{T}\sum_{t=2}^{T}\cos(k,\, t)\Delta u_t.$$

It follows from the FCLT and CMT that

$$T^{-1/2}\hat{u}_{[Tr]} = T^{-1/2}u_{[Tr]} - T^{-1/2}\mathbf{x}_{[Tr]}'\left[D_T\sum_{t=1}^{T}\mathbf{x}_{\tilde{c},t}\mathbf{x}_{\tilde{c},t}{}'D_T\right]^{-1}\left[D_T\sum_{t=1}^{T}\mathbf{x}_{\tilde{c},t}u_{\tilde{c},t}\right]$$

$$\begin{aligned} \Rightarrow\quad & \sigma W(r) - \sigma r\left[\frac{(1-\tilde{c}\cos(k,r))W(1) + \tilde{c}^2\int_0^1 r\cos^2(k,r)W(r)dr - \tilde{c}k\pi\int_0^1 r\sin(2\pi \mathrm{kr})W(r)dr}{\int_0^1\left[1 - 2\tilde{c}r\cos^2(k,r) + r^2\tilde{c}^2\cos(k,r)\right]dr}\right] \\ =:\quad & \sigma W_\tau(r) \end{aligned} \qquad \text{(S.3)}$$

Thus from the results for Case A and for Case B above, we can now state the limit results for the test statistic. The OLS t-statistics to test $H_0 : \phi = 0$, computed from a test regression as in (4) based on locally GLS demeaned (μ) or de-trended (τ) data, i.e., $\hat{u}_{\tilde{c},t}^{\mathfrak{v}} = y_t - \mathbf{x}_t'\boldsymbol{\beta}_{\mathfrak{v},\tilde{c}}$, $\mathfrak{v} = \mu$ or τ, is,

$$\hat{t}_k^{GLS_{\mathfrak{v}}} := \frac{\sum_{t=2}^{T}\Delta\hat{u}_{\tilde{c},t}^{\mathfrak{v}}\cos(k,t)\hat{u}_{\tilde{c},t-1}^{\mathfrak{v}}}{\left[\hat{\sigma}_k^2\sum_{t=2}^{T}\cos^2(k,\mathrm{t})\hat{u}_{\tilde{c},t-1}^{\mathfrak{v}2}\right]^{1/2}}, \text{with } \mathfrak{v} = \mu,\ \tau. \qquad \text{(S.4)}$$

Considering that $\hat{u}_{\tilde{c},t}^{\mathfrak{v}} = \Delta\hat{u}_{\tilde{c},t}^{\mathfrak{v}} + \hat{u}_{\tilde{c},t-1}^{\mathfrak{v}}$, squaring both sides and multiplying by $\cos(k,t)$ leads to $\cos(k,t)\hat{u}_{\tilde{c},t}^{\mathfrak{v}2} = \cos(k,t)\left[(\Delta\hat{u}_{\tilde{c},t})^{\mathfrak{v}2} + 2\Delta\hat{u}_{\tilde{c},t}^{\mathfrak{v}}\hat{u}_{\tilde{c},t-1}^{\mathfrak{v}} + \hat{u}_{\tilde{c},t-1}^{\mathfrak{v}2}\right]$. Summing over t and rearranging gives,

$$\sum_{t=2}^{T} \Delta \hat{u}^{\mathfrak{v}}_{\tilde{c},t} \cos(k,t) \hat{u}^{\mathfrak{v}}_{\tilde{c},t-1} = \frac{1}{2} \left[\sum_{t=2}^{T} \cos(k,t) \hat{u}^{\mathfrak{v}2}_{\tilde{c},t} - \sum_{t=2}^{T} \cos(k,t) \hat{u}^{\mathfrak{v}2}_{\tilde{c},t-1} - \sum_{t=2}^{T} \cos(k,t) (\Delta \hat{u}_{\tilde{c},t})^{\mathfrak{v}2} \right]. \tag{S.5}$$

Since under the null $\Delta \hat{u}^{\mathfrak{v}}_{\tilde{c},t} = \hat{\varepsilon}_t$, it follows that,

$$\sum_{t=2}^{T} \Delta \hat{u}^{\mathfrak{v}}_{\tilde{c},t} \cos(k,t) \hat{u}^{\mathfrak{v}}_{\tilde{c},t-1} = \frac{1}{2} \left[\cos(k,\mathit{T}) \hat{u}^{\mathfrak{v}2}_{\tilde{c},T} - \cos(k,\mathit{2}) \hat{u}^{\mathfrak{v}2}_{\tilde{c},1} - \sum_{t=3}^{T} \Delta\cos(k,t) \hat{u}^{\mathfrak{v}2}_{\tilde{c},t-1} - \sum_{t=2}^{T} \cos(k,t) \hat{\varepsilon}_t^2 \right]. \tag{S.6}$$

In what follows the following limit results will prove useful. In specific, as $T \to \infty$,

$$\cos(k,\mathit{T}) \frac{1}{T} \hat{u}^2_{\tilde{c},T} \Rightarrow \sigma^2 \cos(k,\mathit{1}) W_{\mathfrak{v}}(1)^2; \tag{S.7}$$

$$\cos(k,\mathit{2}) \frac{1}{T} \hat{u}^2_{\tilde{c},1} \Rightarrow \sigma^2 \cos(k,\mathit{0}) W_{\mathfrak{v}}(0)^2 = 0; \tag{S.8}$$

$$\frac{1}{T} \sum_{t=3}^{T} \Delta\cos(k,t) \hat{u}^2_{\tilde{c},t-1} \Rightarrow \frac{\sigma^2}{2} (2\pi k)^2 \int_0^1 \cos(2\pi k r) W_{\mathfrak{v}}(r)^2 dr; \tag{S.9}$$

$$\frac{1}{T} \sum_{t=2}^{T} \cos(k,t) (\Delta \hat{u}_{\tilde{c},t-1})^2 \to \sigma^2 \int_0^1 \cos(k,r) dr = \frac{\sigma^2}{2}. \tag{S.10}$$

The result in (S.9) is obtained given that $\Delta\cos(k,t) = -\frac{1}{2}(2\pi k/T)\sin(2\pi kt/T) + o(1)$ (see Enders and Lee;2012 and Lemma A.1 in Bierens;1997 Recall that $\cos(k,t) := \frac{1}{2}(1 + \cos(2\pi kt/T))$. Hence, for the numerator of (S.4) we establish as $T \to \infty$ that,

$$\frac{1}{T} \sum_{t=2}^{T} \Delta \hat{u}^{\mathfrak{v}}_{\tilde{c},t} \cos(k,t) \hat{u}^{\mathfrak{v}}_{\tilde{c},t-1} \Rightarrow \frac{\sigma^2}{2} \left\{ \cos(k,1) W_{\mathfrak{v}}(1)^2 + \frac{1}{2} (2\pi k)^2 \int_0^1 \cos(2\pi k r) W_{\mathfrak{v}}(r)^2 dr - 1 \right\} \tag{S.11}$$

and for the denominator it follows from the continuous mapping theorem that,

$$\frac{1}{T^2} \sum_{t=2}^{T} \cos^2(k,t) \hat{u}^{\mathfrak{v}2}_{\tilde{c},t-1} \Rightarrow \sigma^2 \int_0^1 \cos^2(k,r) W_{\mathfrak{v}}(r)^2 dr. \tag{S.12}$$

Thus, from (S.11) and (S.12) it follows, under joint convergence, that,

$$\hat{t}_k^{GLS_{\mathfrak{v}}} \Rightarrow \frac{\cos(k,1) W_{\mathfrak{v}}(1)^2 + \frac{1}{2}(2\pi k)^2 \int_0^1 \cos(2\pi k r) W_{\mathfrak{v}}(r)^2 dr - 1}{2 \left(\int_0^1 \cos^2(k,r) W_{\mathfrak{v}}(r)^2 dr \right)^{1/2}},$$

where k is a fixed value and $\mathfrak{v} = \mu$ or τ depending on whether local GLS demeaning or local GLS de-trending is used, respectively. ■

Proof of Theorem 2.3

An extension of the FCLT to near integrated process, $\rho_t = 1 - \frac{c}{T}$, states that, $\frac{1}{\sqrt{T}} u_{[Tr]} \Rightarrow \sigma J_c(r),\ 0 \leq r \leq 1$, where J_c is a standard OU process (see Phillips;1987). The $P_{\tilde{c}}$ test statistic is given by

$$P_{\tilde{c}} = \frac{\sum_{t=1}^{T} \hat{\varepsilon}_{\tilde{c},t}^2 - \left[1 + \frac{\tilde{c}}{T}\cos(k,t)\right] \sum_{t=1}^{T} \hat{\varepsilon}_{0,t}^2}{\hat{\sigma}^2},$$

where $\hat{\varepsilon}_{0,t}$ is the residual term under $H_0 : \tilde{\rho}_t = 0$ and $\hat{\varepsilon}_{\tilde{c},t}$ is the residual term under $H_1 : \tilde{\rho}_t := 1 + \frac{\tilde{c}}{T}$ for a given $\tilde{c}$. The null hypothesis is rejected for small values of this statistic. Note that, in the case of demeaning,

$$\hat{\varepsilon}_{\tilde{c},t} = y_t - \left(1 + \frac{\tilde{c}}{T}\mathrm{cos(k,t)}\right) y_{t-1} - \beta_1 \left(1 - \left(1 + \frac{\tilde{c}}{T}\mathrm{cos(k,t)}\right)\right) \to \Delta u_t - \frac{\tilde{c}}{T}\mathrm{cos(k,t)} u_{t-1},$$

$$\hat{\varepsilon}_{\tilde{c},t}^2 \to \Delta u_t^2 - \frac{2\tilde{c}}{T}\Delta u_t \mathrm{cos(k,t)} u_{t-1} + \left(\frac{\tilde{c}}{T}\right)^2 \cos^2(k,t) u_{t-1}^2,$$

$$\hat{\varepsilon}_{0,t}^2 = (\Delta u_t)^2,$$

Putting these results together we have that,

$$\sum_{t=1}^{T} \hat{\varepsilon}_{\tilde{c},t}^2 - \left[1 + \frac{\tilde{c}}{T}\mathrm{cos(k,t)}\right] \sum_{t=1}^{T} \hat{\varepsilon}_{0,t}^2 = -\frac{2\tilde{c}}{T} \sum_{t=2}^{T} \Delta u_t \cos(k,t) u_{t-1}$$

$$-\frac{\tilde{c}}{T} \sum_{t=2}^{T} \cos(k,t) \hat{\varepsilon}_{0,t}^2 + \left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \cos^2(k,t) u_{t-1}^2,$$

and given that

$$-\frac{2\tilde{c}}{T} \sum_{t=2}^{T} \Delta u_t \mathrm{cos(k,t)} u_{t-1} \Rightarrow \tilde{c} \left[\sigma^2 \int_0^1 \mathrm{cos(k,r)} dr - \sigma^2 \mathrm{cos(k,T)} J_c^2(1)\right],$$

$$\left(\frac{\tilde{c}}{T}\right)^2 \sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) u_{t-1}^2 \Rightarrow \tilde{c}^2\sigma^2 \int_0^1 \cos^2(\mathrm{k},\mathrm{r}) J_c^2(r),$$

$$\frac{\tilde{c}}{T} \sum_{t=2}^{T} \cos_j(k,t) \hat{\varepsilon}_{0,t}^2 \Rightarrow \tilde{c}\sigma^2 \int_0^1 \cos(k,r) dr,$$

the asymptotic distribution of $P_{\tilde{c}}$ is $P_{\tilde{c}} \Rightarrow \tilde{c}^2 \int_0^1 \cos^2(\mathrm{k},\mathrm{r}) J_c^2(r) - \tilde{c}\cos(\mathrm{k},\mathrm{T}) J_c^2(1)$. Finally, when de-trending is considered $\hat{\varepsilon}_{\tilde{c},t} = y_t - \rho_t y_{t-1} - \beta_1(1-\rho_t) - \beta_2(t - \rho_t\,(t-1))$. Thus, using the FCLT result presented previously it follows that,

$$P_{\tilde{c}} \Rightarrow \tilde{c}^2 \int_0^1 \cos^2(\mathrm{k},\mathrm{r}) \left[J_c^\tau(r)\right]^2 + (1 - \tilde{c}\cos(\mathrm{k},\mathrm{T}) \left[J_c^\tau(1)\right]^2,$$

where J_c^τ is the local GLS de-trended OU process. ■

Proof of Proposition 1

The proposed test statistic with White standard errors is defined as,

$$\hat{t}_{k,W}^{GLS} := \frac{\sum_{t=2}^{T} \Delta\hat{u}_t \cos(k,t) \hat{u}_{t-1}}{\left(\sum_{t=2}^{T} \cos^2(k,\mathrm{t}) \hat{u}_{t-1}^2 \hat{\varepsilon}_t^2\right)^{1/2}}.$$

As the numerator is the same as in equation (5), we only need to examine the denominator. Hence, considering

$$\frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) \hat{u}_{t-1}^2 \hat{\varepsilon}_t^2 = \frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) \hat{u}_{t-1}^2 \sigma^2 + \frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right) \quad \text{(S.13)}$$

and noting that, $\sigma^2 \frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) u_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 \cos^2(\mathrm{k},\mathrm{r}) W(r)^2$, we only need to prove that the second term in (S.13) is $o_p(1)$. Thus, from the result in Demetrescu (2008) $\frac{1}{T^2}\sum_{t=2}^{T} \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right) \xrightarrow{p} 0$, it follows that, $\frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right) \leq \frac{1}{T^2}\sum_{t=2}^{T} \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right)$, and therefore $\frac{1}{T^2}\sum_{t=2}^{T} \cos^2(\mathrm{k},\mathrm{t}) \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right)$ is also $o_p(1)$ since $\cos^2(k,t) \leq 1$ for fixed $k > 0$.

We also need to show that, $\frac{1}{T^2}\sum_{t=2}^{T} \hat{u}_{t-1}^2 \left(\hat{\varepsilon}_t^2 - \sigma^2\right) \xrightarrow{p} 0$ is still true when the process is near integrated. That is, when $u_t = (1 - c\cos(k,t)/T)u_{t-1} + \varepsilon_t$. Since, $y_0 = 0$, we have

$$u_t = \sum_{i=0}^{t-1} \left(1 - \frac{c}{T}\cos(k,\mathrm{i})\right)^i \varepsilon_{t-i};$$

$$\left(1 - \frac{c}{T}\cos(k,i)\right)^i = 1 - \frac{c}{T}\, i\cos(k,i) + O(T^{-1}),$$

and

$$u_t = \sum_{i=0}^{t-1} \varepsilon_{t-i} - \frac{c}{T}\sum_{i=0}^{t-1} i\cos(k,\mathrm{i})\,\varepsilon_{t-i} + O(T^{-0.5}).$$

Since $i/T = O(1)$, the result can be derived in the same way as in the proof of Proposition 2. ■

References

Bierens, H. J. (1997). Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the us price level and interest rate, Journal of Econometrics 81(1): 29 - 64.

Demetrescu, M. (2008). On the Dickey-Fuller test with White standard errors, Statistical Papers 51(1): 11.

Enders, W. and Lee, J. (2012). A Unit Root Test Using a Fourier Series to Approximate Smooth Breaks, Oxford Bulletin of Economics and Statistics 74(4): 574-599.

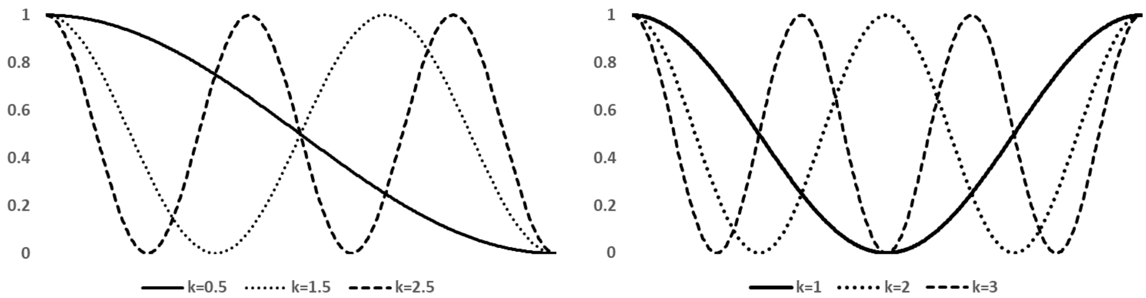
Leybourne, S., Kim, T.-H. and Taylor, A. R. (2007). Detecting Multiple Changes in Persistence, Studies in Nonlinear Dynamics & Econometrics 11(3): 1 - 34.

Phillips, P. C. B. (1987). Time Series Regression with a Unit Root, Econometrica 55(2): 277-301.

Additional Figure

Figure S.1 illustrates the shape of $\cos(k, t)$ for $k = (0.5, 1, 1.5, 2, 2.5, 3)$ and shows the usefulness of fractional frequency values for the approximation of periods of higher persistence at the end of the sample.

Figure S.1: The cosine function for non-integer and integer values of k



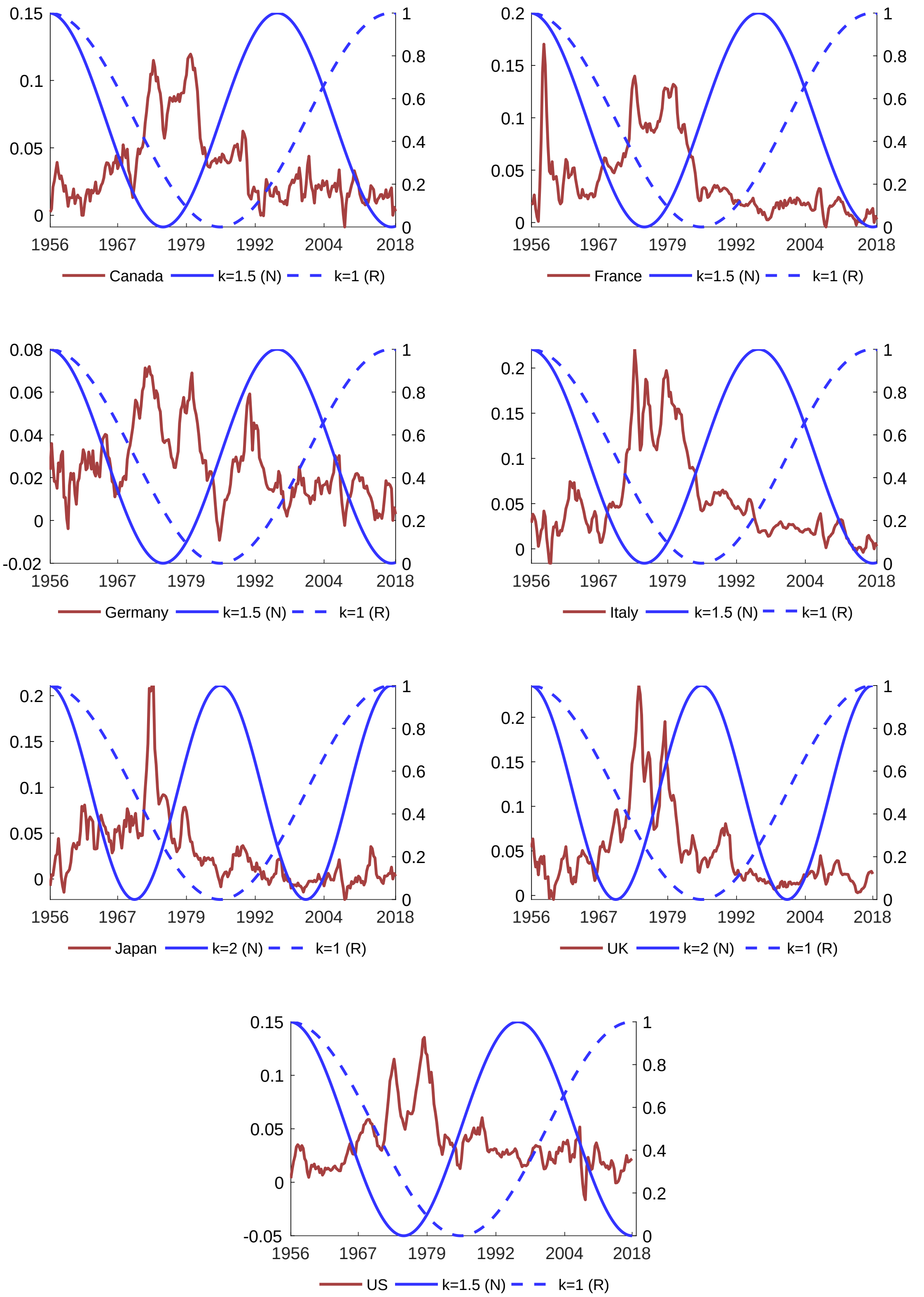


Figure S.2: Year-on-year growth of the CPI for the G7 countries and, superimposed, the chosen cosine function defined in (3) for the time series with normal (N) and reverse chronological order (R).

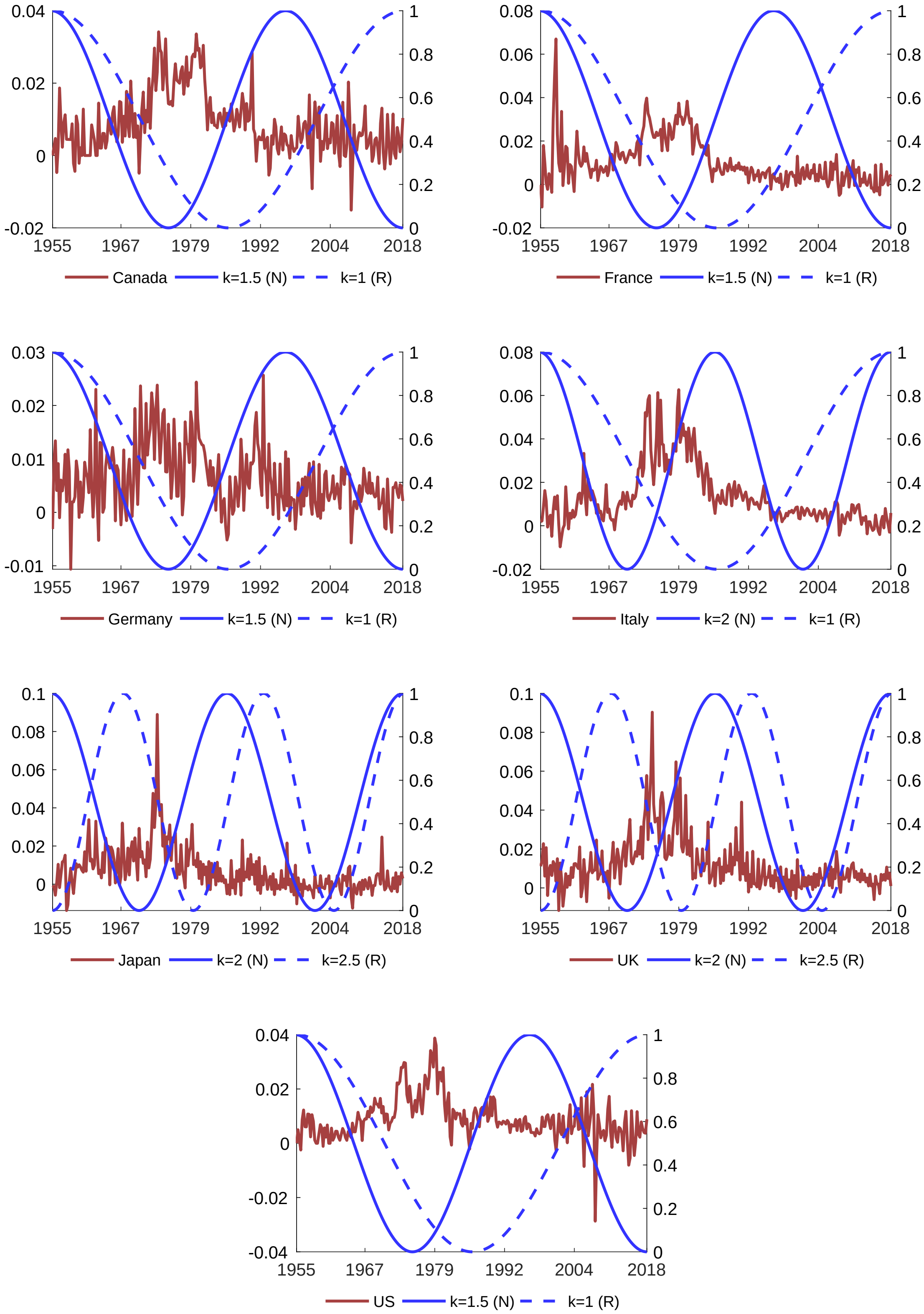


Figure S.3: Quarterly growth of the CPI for the G7 countries and, superimposed, the chosen cosine function defined in (3) for the time series with normal (N) and reverse chronological order (R).

Additional Tables

Table S.1: Finite sample size and power with ε_t autocorrelated

DGP: $y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t$
$\varepsilon_t = \delta\varepsilon_{t-1} + \theta e_{t-1} + e_t$, $e_t \sim N(0,1)$

$T = 250$		DF^{GLS_μ}			$\mathcal{T}_{\hat{k}}^{GLS_\mu}$		
(ϕ, k)	θ / δ	0	0.3	0.6	0	0.3	0.6
$(0, 0)$	−0.8	0.078	0.091	0.104	0.063	0.074	0.091
	−0.4	0.052	0.053	0.029	0.055	0.056	0.034
	0	0.044	0.044	0.043	0.045	0.046	0.044
	0.4	0.043	0.040	0.039	0.046	0.043	0.044
	0.8	0.039	0.039	0.040	0.045	0.044	0.046
$(-0.1, 0.5)$	−0.8	0.235	0.317	0.426	0.164	0.266	0.433
	−0.4	0.361	0.395	0.289	0.378	0.452	0.282
	0	0.392	0.375	0.351	0.439	0.426	0.411
	0.4	0.356	0.335	0.312	0.403	0.379	0.343
	0.8	0.289	0.283	0.262	0.331	0.322	0.301
$(-0.1, 1.0)$	−0.8	0.290	0.389	0.512	0.226	0.350	0.550
	−0.4	0.434	0.480	0.366	0.518	0.604	0.496
	0	0.473	0.463	0.445	0.605	0.603	0.606
	0.4	0.437	0.420	0.400	0.588	0.576	0.553
	0.8	0.362	0.359	0.342	0.529	0.524	0.509
$(-0.1, 1.5)$	−0.8	0.275	0.364	0.492	0.204	0.330	0.523
	−0.4	0.401	0.441	0.316	0.483	0.570	0.436
	0	0.436	0.418	0.393	0.562	0.561	0.545
	0.4	0.392	0.373	0.344	0.533	0.517	0.486
	0.8	0.321	0.313	0.293	0.481	0.469	0.449
$(-0.1, 2.0)$	−0.8	0.323	0.432	0.576	0.249	0.380	0.592
	−0.4	0.487	0.536	0.405	0.582	0.672	0.584
	0	0.518	0.509	0.498	0.668	0.677	0.676
	0.4	0.489	0.474	0.452	0.658	0.648	0.634
	0.8	0.423	0.419	0.410	0.608	0.605	0.591
$(-0.1, 2.5)$	−0.8	0.308	0.411	0.553	0.243	0.378	0.584
	−0.4	0.466	0.512	0.364	0.557	0.645	0.540
	0	0.493	0.478	0.462	0.641	0.641	0.637
	0.4	0.455	0.439	0.415	0.623	0.610	0.587
	0.8	0.388	0.384	0.369	0.570	0.567	0.549
$(-0.1, 3)$	−0.8	0.355	0.476	0.642	0.269	0.415	0.624
	−0.4	0.547	0.608	0.464	0.618	0.704	0.635
	0	0.585	0.585	0.589	0.706	0.717	0.714
	0.4	0.560	0.553	0.550	0.696	0.689	0.675
	0.8	0.508	0.512	0.515	0.651	0.648	0.633

Table S.2: Empirical size under unconditional variance breaks - constant case

DGP: $y_t = y_{t-1} + \varepsilon_t,$

$$\varepsilon_t \sim \begin{cases} N(0,\sigma_1^2) & for \quad t = 1,...,\lfloor \tau_1 T \rfloor \\ N(0,\sigma_2^2) & for \quad t = \lfloor \tau_1 T \rfloor + 1,...,\lfloor \tau_2 T \rfloor \\ N(0,\sigma_3^2) & for \quad t = \lfloor \tau_2 T \rfloor + 1,...,T, \end{cases}$$

			$T=150$			
	DF^{GLS_μ}	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$	DF^{GLS_μ}	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$	DF^{GLS_μ}	$\mathcal{T}_{\hat{k}}^{GLS_\mu}$
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.3)$		$(\tau_1=0.5,\tau_2=0.5)$		$(\tau_1=0.7,\tau_2=0.7)$	
$(1,2,2)$	0.068	0.077	0.070	0.082	0.069	0.083
$(2,1,1)$	0.059	0.071	0.062	0.073	0.057	0.052
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.6)$		$(\tau_1=0.4,\tau_2=0.7)$		$(\tau_1=0.3,\tau_2=0.7)$	
$(1,1.5,1)$	0.058	0.059	0.058	0.062	0.060	0.066
$(1,2,1)$	0.068	0.078	0.065	0.079	0.071	0.084
			$T=250$			
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.3)$		$(\tau_1=0.5,\tau_2=0.5)$		$(\tau_1=0.7,\tau_2=0.7)$	
$(1,2,2)$	0.069	0.077	0.069	0.080	0.072	0.088
$(2,1,1)$	0.063	0.070	0.061	0.069	0.056	0.058
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.6)$		$(\tau_1=0.4,\tau_2=0.7)$		$(\tau_1=0.3,\tau_2=0.7)$	
$(1,1.5,1)$	0.056	0.059	0.059	0.058	0.060	0.063
$(1,2,1)$	0.069	0.072	0.069	0.078	0.069	0.080
			$T=500$			
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.3)$		$(\tau_1=0.5,\tau_2=0.5)$		$(\tau_1=0.7,\tau_2=0.7)$	
$(1,2,2)$	0.065	0.075	0.071	0.083	0.068	0.089
$(2,1,1)$	0.058	0.066	0.057	0.068	0.057	0.060
$(\sigma_1,\sigma_2,\sigma_3)$	$(\tau_1=0.3,\tau_2=0.6)$		$(\tau_1=0.4,\tau_2=0.7)$		$(\tau_1=0.3,\tau_2=0.7)$	
$(1,1.5,1)$	0.051	0.058	0.058	0.061	0.054	0.059
$(1,2,1)$	0.061	0.072	0.070	0.078	0.063	0.073

Table S.3: Empirical size assuming *iid* errors in the presence of GARCH effects

DGP:

$$y_t = y_{t-1} + \varepsilon_t$$
$$\varepsilon_t = e_t\sqrt{h_t},\ h_t = \omega + \zeta\varepsilon_{t-1}^2 + \xi h_{t-1}$$
$$\omega = 1 - \zeta - \xi,\ e_t \sim N(0,1)$$

	DF^{GLS_μ}			$\mathcal{T}_{\hat{k}}^{GLS_\mu}$		
			$T = 150$			
ξ/ζ	0.7	0.8	0.9	0.7	0.8	0.9
0	0.063	0.065	0.071	0.083	0.093	0.106
0.05	0.065	0.071	0.079	0.086	0.099	0.114
0.1	0.068	0.075	-	0.089	0.102	-
0.2	0.077	-	-	0.102	-	-
			$T = 250$			
0	0.066	0.073	0.079	0.085	0.094	0.109
0.05	0.068	0.074	0.079	0.088	0.099	0.114
0.1	0.069	0.076	-	0.092	0.104	-
0.2	0.076	-	-	0.102	-	-
			$T = 500$			
0	0.061	0.067	0.071	0.071	0.085	0.096
0.05	0.064	0.068	0.074	0.078	0.091	0.104
0.1	0.066	0.071	-	0.086	0.098	-
0.2	0.069	-	-	0.096	-	-

Table S.4: Empirical size and power with iid errors - linear trend case

DGP: $y_t = y_{t-1} + \phi((1 + \cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0, 1)$

	DF^{GLS_τ}			$\mathcal{T}_{\hat{k}}^{GLS_\tau}$			$DF^{GLS_\tau^*}$			$\mathcal{T}_{\hat{k}}^{GLS_\tau^*}$		
						$T = 150$						
k/ϕ	0	-0.1	-0.2	0	-0.1	-0.2	0	-0.1	-0.2	0	-0.1	-0.2
0	0.050	0.594	0.991	0.048	0.364	0.859	0.053	0.584	0.988	0.046	0.340	0.839
0.5	0.050	0.184	0.400	0.048	0.153	0.373	0.053	0.195	0.415	0.046	0.146	0.351
1.0	0.050	0.119	0.280	0.048	0.134	0.396	0.053	0.120	0.274	0.046	0.110	0.348
1.5	0.050	0.126	0.298	0.048	0.130	0.376	0.053	0.132	0.297	0.046	0.125	0.338
2.0	0.050	0.134	0.314	0.048	0.147	0.439	0.053	0.133	0.307	0.046	0.142	0.406
2.5	0.050	0.130	0.307	0.048	0.147	0.434	0.053	0.129	0.324	0.046	0.130	0.415
3.0	0.050	0.146	0.348	0.048	0.171	0.481	0.053	0.149	0.343	0.046	0.155	0.446
						$T = 250$						
0	0.048	0.960	1.000	0.051	0.725	0.994	0.048	0.969	1.000	0.056	0.706	0.994
0.5	0.048	0.333	0.604	0.051	0.290	0.759	0.048	0.356	0.629	0.056	0.309	0.737
1.0	0.048	0.230	0.478	0.051	0.298	0.819	0.048	0.223	0.456	0.056	0.300	0.787
1.5	0.048	0.246	0.504	0.051	0.280	0.797	0.048	0.259	0.517	0.056	0.264	0.789
2.0	0.048	0.253	0.537	0.051	0.336	0.850	0.048	0.264	0.544	0.056	0.325	0.824
2.5	0.048	0.244	0.542	0.051	0.325	0.833	0.048	0.243	0.527	0.056	0.309	0.818
3.0	0.048	0.282	0.598	0.051	0.369	0.872	0.048	0.275	0.602	0.056	0.361	0.844
						$T = 500$						
0	0.050	1.000	1.000	0.048	0.995	1.000	0.049	1.000	1.000	0.046	0.987	1.000
0.5	0.050	0.607	0.806	0.048	0.764	0.999	0.049	0.612	0.808	0.046	0.749	0.997
1.0	0.050	0.486	0.728	0.048	0.827	0.999	0.049	0.490	0.727	0.046	0.807	1.000
1.5	0.050	0.509	0.758	0.048	0.806	0.999	0.049	0.517	0.767	0.046	0.782	0.998
2.0	0.050	0.535	0.811	0.048	0.859	1.000	0.049	0.540	0.793	0.046	0.851	1.000
2.5	0.050	0.540	0.828	0.048	0.847	1.000	0.049	0.550	0.817	0.046	0.837	1.000
3.0	0.050	0.595	0.884	0.048	0.879	1.000	0.049	0.597	0.882	0.046	0.871	0.999

Note: $DF^{GLS_\tau^*}$ and $\mathcal{T}_{\hat{k}}^{GLS_\tau^*}$ are Wild bootstrap based test statistics.

Table S.5: Empirical power when the breaks in persistence are abrupt - linear trend case

$$\text{DGP:}\;\; y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ \rho_2 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ \rho_3 y_{t-1} + \varepsilon_t & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T, \end{cases} \qquad \varepsilon_t \sim N(0,1)$$

	DF^{GLS_τ}		$\mathcal{T}_{\hat{k}}^{GLS_\tau}$		DF^{GLS_τ}		$\mathcal{T}_{\hat{k}}^{GLS_\tau}$		DF^{GLS_τ}		$\mathcal{T}_{\hat{k}}^{GLS_\tau}$	
	CV	WB	CV	WB	CV	WB	CV	WB	CV	WB	CV	WB
						$T = 150$						
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.185	0.206	0.211	0.199	0.442	0.460	0.440	0.429	0.712	0.728	0.661	0.658
$(0.8, 1, 1)$	0.143	0.156	0.168	0.154	0.360	0.369	0.386	0.372	0.646	0.660	0.637	0.620
$(0.9, 0.99, 0.99)$	0.129	0.147	0.124	0.100	0.244	0.253	0.186	0.176	0.375	0.381	0.259	0.234
$(0.9, 1, 1)$	0.105	0.119	0.103	0.091	0.198	0.203	0.157	0.155	0.337	0.346	0.241	0.215
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.449	0.450	0.615	0.584	0.461	0.475	0.657	0.604	0.304	0.320	0.531	0.494
$(0.8, 1, 0.8)$	0.376	0.380	0.610	0.575	0.393	0.390	0.651	0.598	0.241	0.248	0.530	0.468
$(0.9, 0.99, 0.9)$	0.228	0.222	0.232	0.208	0.240	0.230	0.230	0.196	0.166	0.175	0.181	0.153
$(0.9, 1, 0.9)$	0.187	0.184	0.231	0.208	0.198	0.185	0.214	0.181	0.125	0.131	0.172	0.135
						$T = 250$						
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.256	0.236	0.383	0.372	0.596	0.614	0.810	0.791	0.846	0.851	0.956	0.952
$(0.8, 1, 1)$	0.174	0.164	0.274	0.257	0.445	0.457	0.713	0.696	0.749	0.760	0.946	0.942
$(0.9, 0.99, 0.99)$	0.195	0.171	0.208	0.215	0.436	0.449	0.381	0.391	0.673	0.694	0.545	0.535
$(0.9, 1, 1)$	0.134	0.132	0.151	0.148	0.313	0.340	0.300	0.321	0.576	0.597	0.506	0.500
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.617	0.611	0.929	0.906	0.632	0.621	0.964	0.951	0.456	0.437	0.902	0.868
$(0.8, 1, 0.8)$	0.483	0.463	0.920	0.900	0.499	0.496	0.964	0.943	0.317	0.297	0.897	0.857
$(0.9, 0.99, 0.9)$	0.433	0.416	0.487	0.480	0.452	0.451	0.516	0.495	0.315	0.287	0.414	0.410
$(0.9, 1, 0.9)$	0.326	0.311	0.480	0.463	0.345	0.326	0.506	0.492	0.211	0.203	0.410	0.393
						$T = 500$						
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.3)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.6, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.8, 0.8)$			
$(0.8, 0.99, 0.99)$	0.412	0.396	0.680	0.678	0.799	0.791	0.991	0.989	0.951	0.938	1.000	1.000
$(0.8, 1, 1)$	0.200	0.192	0.413	0.404	0.509	0.517	0.920	0.912	0.812	0.815	1.000	1.000
$(0.9, 0.99, 0.99)$	0.360	0.339	0.474	0.466	0.723	0.720	0.863	0.855	0.910	0.904	0.968	0.958
$(0.9, 1, 1)$	0.169	0.164	0.267	0.256	0.434	0.440	0.697	0.675	0.749	0.750	0.951	0.940
(ρ_1, ρ_2, ρ_3)	$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.6)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.4, 0.7)$				$(\bar{\tau}_1, \bar{\tau}_2) = (0.3, 0.7)$			
$(0.8, 0.99, 0.8)$	0.819	0.813	0.999	0.999	0.841	0.833	1.000	1.000	0.699	0.685	0.999	0.999
$(0.8, 1, 0.8)$	0.563	0.580	0.998	0.997	0.582	0.572	1.000	1.000	0.386	0.381	0.997	0.992
$(0.9, 0.99, 0.9)$	0.740	0.732	0.941	0.937	0.757	0.745	0.963	0.950	0.607	0.608	0.913	0.904
$(0.9, 1, 0.9)$	0.479	0.489	0.924	0.917	0.497	0.494	0.966	0.945	0.313	0.309	0.901	0.878

Table S.6: Finite sample sizes and power with ε_t autocorrelated - linear trend case

DGP: $$y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t$$
$$\varepsilon_t = \delta\varepsilon_{t-1} + \theta e_{t-1} + e_t,\ e_t \sim N(0,1)$$

$T=250$		DF^{GLS_τ}			$\mathcal{T}_{\hat{k}}^{GLS_\tau}$		
(ϕ,k)	θ / δ	0	0.3	0.6	0	0.3	0.6
$(0,0)$	−0.8	0.066	0.085	0.107	0.029	0.047	0.093
	−0.4	0.046	0.046	0.013	0.050	0.054	0.022
	0	0.033	0.032	0.034	0.039	0.040	0.043
	0.4	0.030	0.027	0.026	0.037	0.033	0.033
	0.8	0.022	0.022	0.023	0.035	0.033	0.034
$(-0.1,0.5)$	−0.8	0.169	0.229	0.315	0.050	0.093	0.227
	−0.4	0.218	0.254	0.104	0.170	0.228	0.076
	0	0.231	0.204	0.186	0.196	0.185	0.176
	0.4	0.188	0.161	0.138	0.162	0.136	0.116
	0.8	0.118	0.109	0.099	0.112	0.107	0.102
$(-0.1,1.0)$	−0.8	0.117	0.162	0.235	0.054	0.106	0.247
	−0.4	0.151	0.172	0.067	0.201	0.254	0.123
	0	0.155	0.138	0.132	0.227	0.218	0.211
	0.4	0.128	0.112	0.097	0.201	0.179	0.160
	0.8	0.078	0.075	0.069	0.159	0.154	0.145
$(-0.1,1.5)$	−0.8	0.133	0.181	0.245	0.055	0.106	0.240
	−0.4	0.160	0.182	0.071	0.190	0.234	0.101
	0	0.167	0.146	0.133	0.206	0.195	0.184
	0.4	0.130	0.111	0.092	0.174	0.152	0.138
	0.8	0.079	0.073	0.065	0.137	0.134	0.126
$(-0.1,2.0)$	−0.8	0.148	0.201	0.278	0.068	0.128	0.278
	−0.4	0.175	0.201	0.078	0.242	0.295	0.168
	0	0.181	0.161	0.151	0.268	0.258	0.248
	0.4	0.144	0.127	0.109	0.241	0.219	0.196
	0.8	0.094	0.088	0.082	0.201	0.192	0.178
$(-0.1,2.5)$	−0.8	0.151	0.203	0.275	0.072	0.137	0.278
	−0.4	0.172	0.194	0.068	0.228	0.287	0.157
	0	0.17	0.151	0.141	0.252	0.246	0.236
	0.4	0.135	0.117	0.099	0.230	0.211	0.190
	0.8	0.089	0.081	0.070	0.187	0.183	0.171
$(-0.1,3)$	−0.8	0.177	0.238	0.327	0.080	0.151	0.313
	−0.4	0.209	0.232	0.088	0.277	0.334	0.207
	0	0.202	0.186	0.179	0.305	0.293	0.289
	0.4	0.168	0.146	0.134	0.276	0.257	0.237
	0.8	0.113	0.108	0.105	0.234	0.226	0.213

Table S.7: Empirical size under unconditional variance breaks - linear trend case

$$\text{DGP:} \quad y_t = y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \begin{cases} N(0,\sigma_1^2) & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ N(0,\sigma_2^2) & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ N(0,\sigma_3^2) & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T, \end{cases}$$

	DF^{GLS_τ}	$\mathcal{T}_{\hat{k}}^{GLS_\tau}$	DF^{GLS_τ}	$\mathcal{T}_{\hat{k}}^{GLS_\tau}$	DF^{GLS_τ}	$\mathcal{T}_{\hat{k}}^{GLS_\tau}$
			$T = 150$			
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.3)$		$(\tau_1 = 0.5, \tau_2 = 0.5)$		$(\tau_1 = 0.7, \tau_2 = 0.7)$	
$(1, 2, 2)$	0.074	0.074	0.077	0.088	0.081	0.107
$(2, 1, 1)$	0.074	0.083	0.076	0.079	0.070	0.066
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.6)$		$(\tau_1 = 0.4, \tau_2 = 0.7)$		$(\tau_1 = 0.3, \tau_2 = 0.7)$	
$(1, 1.5, 1)$	0.061	0.063	0.059	0.065	0.058	0.065
$(1, 2, 1)$	0.074	0.096	0.070	0.102	0.069	0.095
			$T = 250$			
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.3)$		$(\tau_1 = 0.5, \tau_2 = 0.5)$		$(\tau_1 = 0.7, \tau_2 = 0.7)$	
$(1, 2, 2)$	0.063	0.073	0.065	0.089	0.068	0.108
$(2, 1, 1)$	0.059	0.077	0.058	0.075	0.052	0.063
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.6)$		$(\tau_1 = 0.4, \tau_2 = 0.7)$		$(\tau_1 = 0.3, \tau_2 = 0.7)$	
$(1, 1.5, 1)$	0.058	0.063	0.054	0.062	0.058	0.062
$(1, 2, 1)$	0.069	0.091	0.065	0.101	0.071	0.093
			$T = 500$			
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.3)$		$(\tau_1 = 0.5, \tau_2 = 0.5)$		$(\tau_1 = 0.7, \tau_2 = 0.7)$	
$(1, 2, 2)$	0.065	0.078	0.066	0.088	0.067	0.105
$(2, 1, 1)$	0.059	0.075	0.056	0.070	0.054	0.062
$(\sigma_1, \sigma_2, \sigma_3)$	$(\tau_1 = 0.3, \tau_2 = 0.6)$		$(\tau_1 = 0.4, \tau_2 = 0.7)$		$(\tau_1 = 0.3, \tau_2 = 0.7)$	
$(1, 1.5, 1)$	0.058	0.062	0.062	0.069	0.059	0.064
$(1, 2, 1)$	0.071	0.089	0.077	0.103	0.074	0.093

Table S.8: Empirical size and power of Wild bootstrap based statistics under unconditional variance breaks - linear trend case.

$$\text{DGP:} \quad y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t,$$
$$\varepsilon_t \sim \begin{cases} N(0,\sigma_1^2) & for \quad t = 1,...,\lfloor \tau_1 T \rfloor \\ N(0,\sigma_2^2) & for \quad t = \lfloor \tau_1 T \rfloor + 1,...,\lfloor \tau_2 T \rfloor \\ N(0,\sigma_3^2) & for \quad t = \lfloor \tau_2 T \rfloor + 1,...,T, \end{cases}$$

	DF_τ^{GLS*}			$\mathcal{T}_{\hat{k}}^{GLS_\tau^*}$		
$T = 250$	$\sigma_1 = 1,\ \sigma_2 = 2$ and $\sigma_3 = 2$					
$(\phi, k)/(\bar{\tau}_1,\ \bar{\tau}_2)$	$(0.3, 0.3)$	$(0.5, 0.5)$	$(0.7, 0.7)$	$(0.3, 0.3)$	$(0.5, 0.5)$	$(0.7, 0.7)$
$(0, 0)$	0.045	0.054	0.051	0.050	0.050	0.060
$(-0.1, 0.5)$	0.259	0.230	0.200	0.188	0.178	0.176
$(-0.1, 1.0)$	0.175	0.262	0.332	0.234	0.285	0.341
$(-0.1, 1.5)$	0.260	0.317	0.244	0.254	0.287	0.222
$(-0.1, 2.0)$	0.294	0.258	0.236	0.279	0.263	0.256
$(-0.1, 2.5)$	0.276	0.241	0.322	0.274	0.242	0.265
$(-0.1, 3.0)$	0.266	0.315	0.273	0.311	0.319	0.269
	$\sigma_1 = 2,\ \sigma_2 = 2$ and $\sigma_3 = 1$					
$(0, 0)$	0.044	0.044	0.051	0.057	0.049	0.046
$(-0.1, 0.5)$	0.470	0.520	0.493	0.391	0.386	0.352
$(-0.1, 1.0)$	0.264	0.202	0.171	0.302	0.228	0.203
$(-0.1, 1.5)$	0.245	0.199	0.259	0.234	0.192	0.290
$(-0.1, 2.0)$	0.213	0.250	0.255	0.218	0.262	0.264
$(-0.1, 2.5)$	0.223	0.269	0.226	0.235	0.279	0.244
$(-0.1, 3.0)$	0.245	0.230	0.288	0.279	0.262	0.304
	$\sigma_1 = 1,\ \sigma_2 = 1.5$ and $\sigma_3 = 1$					
$(\phi, k)/\bar{\tau}$	$(0.3, 0.6)$	$(0.4, 0.7)$	$(0.3, 0.7)$	$(0.3, 0.6)$	$(0.4, 0.7)$	$(0.3, 0.7)$
$(0, 0)$	0.051	0.049	0.049	0.053	0.051	0.050
$(-0.1, 0.5)$	0.365	0.317	0.326	0.290	0.254	0.262
$(-0.1, 1.0)$	0.163	0.180	0.156	0.184	0.208	0.181
$(-0.1, 1.5)$	0.231	0.270	0.249	0.249	0.284	0.278
$(-0.1, 2.0)$	0.312	0.268	0.266	0.320	0.289	0.282
$(-0.1, 2.5)$	0.238	0.218	0.235	0.300	0.247	0.262
$(-0.1, 3.0)$	0.261	0.263	0.283	0.327	0.313	0.353
	$\sigma_1 = 1,\ \sigma_2 = 2$ and $\sigma_3 = 1$					
$(0, 0)$	0.053	0.053	0.054	0.043	0.050	0.049
$(-0.1, 0.5)$	0.361	0.302	0.317	0.267	0.208	0.245
$(-0.1, 1.0)$	0.126	0.146	0.124	0.126	0.128	0.124
$(-0.1, 1.5)$	0.220	0.278	0.254	0.226	0.278	0.252
$(-0.1, 2.0)$	0.321	0.267	0.253	0.302	0.279	0.255
$(-0.1, 2.5)$	0.228	0.210	0.226	0.276	0.194	0.237
$(-0.1, 3.0)$	0.240	0.253	0.278	0.276	0.270	0.316

Table S.9: Empirical size assuming *iid* errors in the presence of GARCH effects - linear trend case

DGP:
$$y_t = y_{t-1} + \varepsilon_t$$
$$\varepsilon_t = e_t\sqrt{h_t},\ h_t = \omega + \zeta\varepsilon_{t-1}^2 + \xi h_{t-1}$$
$$\omega = 1 - \zeta - \xi,\ e_t \sim N(0,1)$$

	DF^{GLS_τ}			$\mathcal{T}_{\hat{k}}^{GLS_\tau}$		
			$T = 150$			
ξ/ζ	0.7	0.8	0.9	0.7	0.8	0.9
0	0.079	0.091	0.102	0.127	0.146	0.168
0.05	0.084	0.097	0.108	0.130	0.153	0.183
0.1	0.087	0.102	-	0.138	0.164	-
0.2	0.098	-	-	0.157	-	-
			$T = 250$			
0	0.075	0.084	0.099	0.120	0.142	0.168
0.05	0.076	0.089	0.106	0.126	0.149	0.178
0.1	0.081	0.097	-	0.133	0.161	-
0.2	0.093	-	-	0.157	-	-
			$T = 500$			
0	0.070	0.078	0.090	0.103	0.126	0.158
0.05	0.072	0.084	0.100	0.110	0.138	0.173
0.1	0.076	0.091	-	0.120	0.153	-
0.2	0.089	-	-	0.147	-	-

Table S.10: Empirical size and power in the presence of GARCH effects using EW standard errors and Wild bootstrap based test statistics - linear trend case.

DGP:

$$y_t = y_{t-1} + \phi((1+\cos(2\pi kt/T))/2)y_{t-1} + \varepsilon_t,$$
$$\varepsilon_t = e_t\sqrt{h_t},\ h_t = \omega + \zeta\varepsilon_{t-1}^2 + \xi h_{t-1}$$
$$\omega = 1-\zeta-\xi,\ e_t \sim N(0,1)$$

$T=250$		$DF^{GLS_\tau^{EW}}$			$\mathcal{T}_{\hat{k}}^{GLS_\tau^{EW}}$			$DF^{GLS_\tau^*}$			$\mathcal{T}_{\hat{k}}^{GLS_\tau^*}$		
	ξ/ζ	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9
$\phi=0$	0	0.041	0.037	0.035	0.049	0.044	0.040	0.049	0.051	0.057	0.064	0.069	0.074
	0.05	0.039	0.035	0.033	0.049	0.043	0.040	0.049	0.055	0.061	0.068	0.068	0.076
	0.1	0.038	0.033	-	0.047	0.042	-	0.058	0.056	-	0.068	0.075	-
	0.2	0.027	-	-	0.044	-	-	0.058	-	-	0.070	-	-
$\phi=-0.1$													
$k=0.5$	0	0.229	0.200	0.170	0.225	0.201	0.175	0.321	0.323	0.311	0.297	0.290	0.275
	0.05	0.219	0.186	0.150	0.216	0.191	0.156	0.323	0.318	0.306	0.296	0.283	0.26
	0.1	0.207	0.174	-	0.207	0.180	-	0.322	0.307	-	0.295	0.282	-
	0.2	0.169	-	-	0.184	-	-	0.308	-	-	0.281	-	-
$k=1$	0	0.168	0.147	0.126	0.227	0.210	0.188	0.217	0.214	0.219	0.251	0.234	0.224
	0.05	0.159	0.139	0.114	0.219	0.201	0.174	0.218	0.22	0.216	0.245	0.227	0.208
	0.1	0.153	0.127	-	0.213	0.193	-	0.219	0.218	-	0.237	0.22	-
	0.2	0.125	-	-	0.196	-	-	0.226	-	-	0.219	-	-
$k=1.5$	0	0.169	0.148	0.126	0.215	0.196	0.176	0.238	0.240	0.241	0.246	0.24	0.232
	0.05	0.161	0.138	0.114	0.208	0.187	0.162	0.249	0.243	0.238	0.252	0.242	0.218
	0.1	0.151	0.131	-	0.200	0.181	-	0.246	0.242	-	0.247	0.237	-
	0.2	0.124	-	-	0.185	-	-	0.243	-	-	0.239	-	-
$k=2$	0	0.189	0.165	0.139	0.283	0.262	0.244	0.250	0.249	0.251	0.265	0.259	0.246
	0.05	0.180	0.155	0.128	0.276	0.252	0.227	0.249	0.249	0.249	0.274	0.258	0.254
	0.1	0.168	0.145	-	0.265	0.242	-	0.257	0.257	-	0.277	0.259	-
	0.2	0.138	-	-	0.247	-	-	0.262	-	-	0.260	-	-
$k=2.5$	0	0.186	0.164	0.142	0.282	0.261	0.242	0.250	0.247	0.251	0.286	0.267	0.246
	0.05	0.177	0.153	0.127	0.274	0.252	0.226	0.246	0.249	0.245	0.279	0.257	0.228
	0.1	0.167	0.142	-	0.264	0.242	-	0.250	0.257	-	0.267	0.243	-
	0.2	0.137	-	-	0.245	-	-	0.255	-	-	0.259	-	-
$k=3$	0	0.203	0.182	0.156	0.332	0.310	0.289	0.281	0.284	0.279	0.319	0.297	0.285
	0.05	0.193	0.169	0.139	0.324	0.299	0.268	0.285	0.289	0.285	0.313	0.295	0.260
	0.1	0.184	0.157	-	0.315	0.287	-	0.287	0.291	-	0.306	0.283	-
	0.2	0.148	-	-	0.288	-	-	0.287	-	-	0.293	-	-

Note: $DF^{GLS_\tau^{EW}}$ and $\mathcal{T}_{\hat{k}}^{GLS_\tau^{EW}}$ are EW based test statistics and, $DF^{GLS_\tau^*}$ and $\mathcal{T}_{\hat{k}}^{GLS_\tau^*}$ are Wild bootstrap based test statistics.

Table S.11: Comparison with the M test of Leybourne et al. (2007) - empirical size

DGP:	$y_t = y_{t-1} + u_t$ $u_t = \phi u_{t-1} + e_t,\ e_t \sim N(0,1)$			
	$T = 200$		$T = 400$	
ϕ	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$
0.4	0.105	0.077	0.080	0.044
0.6	0.105	0.077	0.082	0.044

Note: $\mathcal{T}_{\hat{k}}^{GLS_\mu,nr}$ is the minimum between $\mathcal{T}_{\hat{k}}^{GLS_\mu}$ with normal and reverse chronological order.

Table S.12: Comparison with the M test of Leybourne et al. (2007) - empirical power

DGP:

$$\begin{cases} y_t - d\,y_{\lfloor \tau_0 T \rfloor} = \rho_1 (y_{t-1} - d\,y_{\lfloor \tau_0 T \rfloor}) + e_t & for \quad t = 1, ..., \lfloor \tau_1 T \rfloor \\ y_t - d\,y_{\lfloor \tau_1 T \rfloor} = \rho_2 (y_{t-1} - d\,y_{\lfloor \tau_1 T \rfloor}) + e_t & for \quad t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ y_t - d\,y_{\lfloor \tau_2 T \rfloor} = \rho_3 (y_{t-1} - d\,y_{\lfloor \tau_2 T \rfloor}) + e_t & for \quad t = \lfloor \tau_2 T \rfloor + 1, ..., T. \end{cases}$$

				$T = 200$							
				$d = 1$				$d = 0$			
				$\alpha = 0.8$		$\alpha = 0.9$		$\alpha = 0.8$		$\alpha = 0.9$	
(τ_1, τ_2)	ρ_1	ρ_2	ρ_3	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$	M	$\mathcal{T}_{\hat{k}}^{GLS_\tau,nr}$
(0.30, 0.00)	1	α	-	0.801	0.773	0.254	0.394	0.687	0.886	0.185	0.586
(0.50, 0.00)	1	α	-	0.499	0.458	0.157	0.221	0.334	0.589	0.096	0.384
(0.30, 0.00)	α	1	-	0.168	0.187	0.076	0.119	0.168	0.187	0.076	0.119
(0.50, 0.00)	α	1	-	0.450	0.504	0.137	0.240	0.450	0.504	0.137	0.240
(0.25, 0.75)	1	α	1	0.597	0.233	0.190	0.164	0.462	0.462	0.140	0.310
(0.35, 0.75)	1	α	1	0.438	0.167	0.149	0.122	0.297	0.328	0.097	0.225
(0.25, 0.50)	α	1	α	0.399	0.384	0.142	0.231	0.296	0.919	0.100	0.576
(0.35, 0.50)	α	1	α	0.629	0.589	0.211	0.367	0.636	0.986	0.196	0.741
				$T = 400$							
(0.30, 0.00)	1	α	-	1.000	0.937	0.813	0.713	1.000	0.978	0.687	0.877
(0.50, 0.00)	1	α	-	0.991	0.663	0.511	0.415	0.963	0.724	0.334	0.599
(0.30, 0.00)	α	1	-	0.634	0.276	0.179	0.191	0.634	0.276	0.179	0.191
(0.50, 0.00)	α	1	-	0.988	0.715	0.468	0.505	0.988	0.715	0.468	0.505
(0.25, 0.75)	1	α	1	0.995	0.257	0.601	0.206	0.984	0.715	0.468	0.472
(0.35, 0.75)	1	α	1	0.957	0.518	0.451	0.148	0.886	0.423	0.294	0.340
(0.25, 0.50)	α	1	α	0.998	0.746	0.404	0.377	0.896	0.997	0.302	0.933
(0.35, 0.50)	α	1	α	0.813	0.713	0.637	0.570	0.995	1.000	0.642	0.992

Note: $\lfloor \tau_i T \rfloor$ denotes the integer part of $\tau_i T$ and $\tau_0 = 0$; d is a dummy variable that equals 1 for the data generating processes considered in Leybourne et al. (2007), and 0 otherwise.