Testing for a break in persistence under long-range dependencies

Philipp Sibbertsen[1] and Robinson Kruse

Institut für Statistik, Wirtschaftswissenschaftliche Fakultät,
Leibniz Universität Hannover, D-30167 Hannover, Germany

Abstract

We show that tests for a break in the persistence of a time series in the classical $I(0)$ - $I(1)$ framework have serious size distortions when the actual data generating process exhibits long-range dependencies. We prove that the limiting distribution of a CUSUM of squares based test depends on the true memory parameter if the DGP exhibits long memory. We propose adjusted critical values for the test and give finite sample response curves which allow the practitioner to easily implement the test and to compute the relevant critical values. We furthermore prove consistency of the test and prove consistency for a simple break point estimator also under long memory. We show that the test has satisfying power properties when the correct critical values are used.

JEL-numbers: C12, C22.

Keywords: break in pesistence, long memory, CUSUM of squares based test.

1 Introduction

For a practitioner it is of big importance in terms of model building and forecasting to know whether a given time series has a certain kind of persistence, either stationary $I(d)$ with $0 \leq d < 1/2$ or non-stationary $I(d)$ with $1/2 < d < 3/2$ or whether the persistence breaks from stationary to non - stationary persistence or vice versa. Recently, a number of tests for a break

[1]Corresponding author:
Phone: +49- 511- 762-3783, Fax: +49- 511- 762-3923
e-mail: sibbertsen@statistik.uni-hannover.de
We are grateful to Jörg Breitung, Uwe Hassler and the participants of the workshop ”Persistence in financial time series” in Frankfurt am Main for inspiring discussion and helpful comments.

in the persistence have been proposed in the classical $I(0)$ - $I(1)$ framework. Kim (2000), Kim et al. (2002) and Busetti and Taylor (2004) propose tests for the null hypothesis that the data generating process is $I(0)$ throughout against the alternative of a break to $I(1)$. Contrary to these tests Banerjee et al. (1992) and Leybourne and Newbold (2003) propose tests for the opposite null of $I(1)$ throughout against the alternative of a break to $I(0)$. All these tests have problems when the data generating process does not exhibit a break in persistence, yet when the null is false as well. Therefore, Leybourne and Kim (2007) proposed a CUSUM of squares based test to overcome this problem. The Leybourne and Kim (2007) test is basically the ratio of two CUSUM of squares statistics based on the forward and reverse evaluation of the time series. Although the test is constructed for the null hypothesis that the data generating process is $I(1)$ throughout against a break in persistence to $I(0)$, Leybourne et al. show that it has also power against the alternative of a break from $I(0)$ to $I(1)$ and that it behaves well if the process is $I(0)$ throughout.

However, all of these tests stay in the classical $I(0)$ - $I(1)$ framework. One exception is Beran and Terrin (1996) who consider a test for constancy of the long-memory parameter against a change of it. Their test is based on a functional central limit theorem for quadratic forms. By now it is broadly accepted that many economic variables exhibit long - range dependencies which cannot be covered by the classical framework. Also in the more flexible $I(d)$ framework, $0 \leq d \leq 3/2$ it is crucial to know whether the memory parameter is in the stationary region or in the non - stationary region throughout or whether there is a change in the persistence. It turns out that the Leybourne and Kim (2007) test has serious size distortions, that means the test is conservative, if the data generating process has long memory and therefore the test has a lack of power in this model. This indicates that new critical values depending on the memory parameter are necessary in the $I(d)$ framework. In this paper we investigate the asymptotic behaviour of the Leybourne and Kim (2007) test under long - range dependencies. We derive the limiting distribution under the null that the data generating process exhibits non - stationary long memory. We furthermore show that the breakpoint estimator proposed by Leybourne and Kim (2007) is also consistent under long memory though with a slower rate of convergence depending on d. In a Monte Carlo study we show that the test has satisfying size and power properties when the adjusted critical values are used. Finally the test is applied to monthly US inflation data.

The paper is organized at follows. After introducing the model and the test in section 2, section 3 derives the asymptotic properties of the test. Section 4 contains an intensive Monte Carlo study showing the finite sample properties of the test as well as the power properties and gives response curves to easily compute critical values. Section 5 contains an empirical application to a monthly US inflation time series and section 6 concludes. All proofs are given in the appendix.

2 Model and Test

We assume that the data generating process follows an ARFIMA(p,d,q) - process as proposed by Granger (1980):

$$\Phi(B)(1-B)^d X_t = \Psi(B)\varepsilon_t,$$

where ε_t are iid random variables with mean zero and variance σ^2. The AR- and MA- polynomials $\Phi(B)$ and $\Psi(B)$ are assumed to have all roots outside the unit circle. The degree of integration of X_t is therefore solely determined by the memory parameter d. The test against a change in the persistence as proposed by Leybourne and Kim (2007) uses the statistic

$$R = \frac{\inf_{\tau\in\Lambda} K^f(\tau)}{\inf_{\tau\in\Lambda} K^r(\tau)}, \qquad (1)$$

where $K^f(\tau)$ and $K^r(\tau)$ are CUSUM of squared based statistics based on the forward and reversed residuals of the data generating process as given below. Here τ is the relative breakpoint where we assume that $\tau \in \Lambda$ and that $\Lambda \subset (0,1)$ is symmetric around 0.5. For now we assume τ to be fixed though unknown. However, as τ is usually unknown in practice we study the properties of a simple estimator for the breakpoint in the following section. In detail CUSUM of squared based statistics are defined by

$$K^f(\tau) = [\tau T]^{-2} \sum_{t=1}^{[\tau T]} \hat{v}_{t,\tau}^2$$

and

$$K^r(\tau) = (T - [\tau T])^{-2} \sum_{t=1}^{T-[\tau T]} \tilde{v}_{t,\tau}^2.$$

We denote by $[x]$ the biggest integer smaller than x. Here, $\hat{v}_{t,\tau}$ is the residual from the OLS regression of X_t on a constant $z_t = 1\ \forall t$ based on the observations up to $[\tau T]$. This is

$$\hat{v}_{t,\tau} = X_t - \bar{X}(\tau)$$

with $\bar{X}(\tau) = [\tau T]^{-1} \sum_{t=1}^{[\tau T]} X_t$. Similarly $\tilde{v}_{t,\tau}$ is defined for the reversed series $y_t = X_{T-t+1}$. Thus, it is given by

$$\tilde{v}_{t,\tau} = y_t - \bar{y}(1-\tau)$$

with $\bar{y}(1-\tau) = (T - [\tau T])^{-1} \sum_{t=1}^{T-[\tau T]} y_t$. The case of $z_t = [1,\ t]'$, which corresponds to linear de-trending, is considered later on as well.

Remark: It should be mentioned that the quantities $K^f(\tau)$ and $K^r(\tau)$ are originally defined by including an estimator of the long - run variance of the data generating process. As the behaviour of the test statistic R is independent of the long - run variance we omit this here to keep the notation and proofs simple.

3 Asymptotic properties

In this section we derive the asymptotic properties of the test statistic (1) when the data generating process is $I(d)$. In the following we denote by $\Rightarrow$ weak convergence and by $\xrightarrow{P}$ convergence in probability. We denote by d_0 the long memory parameter under the null hypothesis regardless of its specific value while we distinguish under the alternative hypothesis between values characterizing stationary $(0 \leq d_1 < 1/2)$ and non-stationary processes $(1/2 < d_2 < 3/2)$, respectively.

Theorem 1. *Under the null hypothesis $H_0 : X_t \sim I(d_0)\ \forall t$ with $1/2 < d_0 < 3/2$ the limiting distribution for $T \to \infty$ of R is given by*

$$T^{-2d_0} R \Rightarrow \frac{\inf_{\tau\in\Lambda} L_{d_0}^f(\tau)}{\inf_{\tau\in\Lambda} L_{d_0}^r(\tau)}$$

with

$$
\begin{aligned}
L^f_{d_0}(\tau) &= \int_0^{\tau} W^*_{d_0}(r,\tau)^2 dr \\
L^r_{d_0}(\tau) &= \int_0^{1-\tau} V^*_{d_0}(r,\tau)^2 dr \\
W^*_{d_0}(r,\tau)^2 &= \left(W_{d_0}(r) - \tau^{-1} \int_0^{\tau} W_{d_0}(r) dr \right)^2 \\
V^*_{d_0}(r,\tau)^2 &= \left(W_{d_0}(1-r) - (1-\tau)^{-1} \int_{\tau}^{1} W_{d_0}(r) dr \right)^2
\end{aligned}
$$

for the de-meaned case $(z_t = 1)$ *and*

$$
\begin{aligned}
L^f_{d_0}(\tau) &= \int_0^{\tau} W^{**}_{d_0}(r,\tau)^2 dr \\
L^r_{d_0}(\tau) &= \int_0^{1-\tau} V^{**}_{d_0}(r,\tau)^2 dr \\
W^{**}_{d_0}(r,\tau) &= W_{d_0}(r) - B_0(\tau) - rB_1(\tau) \\
V^{**}_{d_0}(r,\tau) &= -(W_{d_0}(1) - W_{d_0}(1-r)) - B^r_0(\tau) - rB^r_1(\tau) \\
B_0(\tau) &= 4\tau^{-1} \int_0^{\tau} W_{d_0}(r) dr - 6\tau^{-2} \int_0^{\tau} rW_{d_0}(r) dr \\
B_1(\tau) &= 6\tau^{-2} \int_0^{\tau} W_{d_0}(r) dr + 12\tau^{-3} \int_0^{\tau} rW_{d_0}(r) dr \\
B^r_0(\tau) &= 4(1-\tau)^{-1} \left(\int_{\tau}^{1} W_{d_0}(r) dr - (1-\tau) W_{d_0}(1) \right) \\
&\quad -6(1-\tau)^{-2} \left(\frac{(1-\tau)^2}{2} W_{d_0}(1) - \int_{\tau}^{1} W_{d_0}(r) dr \int_{\tau}^{1} rW_{d_0}(r) dr \right)
\end{aligned}
$$

$$
\begin{aligned}
B^r_1(\tau) &= 6(1-\tau)^{-2} \left(\int_{\tau}^{1} W_{d_0}(r) dr - (1-\tau) W_{d_0}(1) \right) \\
&\quad +12(1-\tau)^{-3} \left(\frac{(1-\tau)^2}{2} W_{d_0}(1) - \int_{\tau}^{1} W_{d_0}(r) dr + \int_{\tau}^{1} rW_{d_0}(r) dr \right)
\end{aligned}
$$

for the de-trended case $(z_t = [1,\ t]')$.

Theorem 1 shows that the limiting distribution depends strongly on the memory parameter. This behaviour of the limiting distribution leads to heavy size distortions of the original Leybourne and Kim (2007) test when long - range dependencies are neglected. Therefore, we recommend to use new critical values that are provided in section 4. However, as the critical values vary quite substantially with d this is not very handy in practice as new critical values have to be simulated for each value of d. Therefore, we give also response curves for finite sample critical values depending on d which are easy to implement and allow a fast computation of the critical values. As the variation of the critical values with sample size is minor this is an easy possibility for practical applications. However, see section 4 for further details on this issue. After establishing the limiting distribution under the null we have to prove consistency of the test.

Theorem 2. *Let* $0 \leq d_1 < 1/2$ *and* $1/2 < d_2 < 3/2$.

1. *Under the alternative of a break from stationary to non-stationary long memory, this is from* $I(d_1)$ *to* $I(d_2)$, *we obtain*
$$R = O_P(T^{d_1-d_2}).$$
2. *Under the alternative of a break from non-stationary to stationary long memory, this is from* $I(d_2)$ *to* $I(d_1)$ *we obtain*
$$R = O_P(T^{d_2-d_1}).$$

These results imply that a consistent test against the alternative of a break from non-stationary to stationary long memory is obtained by using critical values from the upper tail of the distribution whereas using the lower tail of the distribution leads to a consistent test against the alternative of breaking from stationary to non-stationary long memory.

Remark: Although the break was in both cases assumed to be from stationary to non-stationary long memory or vice versa the test has also a high power when the break is from stationary to stationary or from non-stationary to non-stationary long memory.

So far the breakpoint was assumed to be unknown. We therefore show that the breakpoint estimators given in Leybourne and Kim (2007) are also consistent in the long memory setup.

Theorem 3. *Denote by τ_0 the true breakpoint. Then, for*

$$\hat{\tau} = \inf_{\tau \in \Lambda} K^f(\tau)$$

we have

$$\hat{\tau} \xrightarrow{P} \tau_0,$$

if the alternative is a break from stationary to non-stationary long memory. For a break from non-stationary to stationary long memory a consistent estimator for the breakpoint τ_0 is given by

$$\hat{\tau} = \inf_{\tau \in \Lambda} K^r(\tau).$$

As usual for long memory processes the convergence is slower than in the Leybourne and Kim (2007) situation. However, this does not change the consistency result in general. Finally, we have to evaluate the behaviour of the test when $X_t \sim I(d_0)\ \forall t$ with $0 \leq d_0 < 1/2$. This is the situation where no break in persistence occurs but on the other hand our null hypothesis from Theorem 1 is wrong as the data generating process exhibits stationary long memory. We have

Theorem 4. *Let $X_t \sim I(d_0)\ \forall t$ with $0 \leq d_0 < 1/2$. Then,*

$$R \xrightarrow{P} 1.$$

In this situation the test has a degenerated limit distribution. As the limit theorems for non-stationary long memory processes hold for $1/2 < d_0 < 3/2$, it is reasonable to integrate the time series in this case before applying the adjusted test. For $0 \leq d_0 < 1/2$ the memory parameter of the integrated series is between 1 and $3/2$. Thus, the results in Theorem 1 to Theorem 3 still hold in this situation allowing us to construct a consistent and correctly sized test. By this approach we can overcome the problem of the Leybourne and Kim (2007) test to have a degenerated limiting distribution when the original series is stationary and therefore obtaining a conservative test in this situation.

So far, the value of d_0 has been assumed to be known. Of course, the true value of d_0 is unknown in practice and has to be estimated. However, our Monte Carlo results in the following section show that the adjusted test performs well if d_0 is estimated by a consistent estimator.

4 Monte Carlo study

In this section we evaluate the finite sample behaviour of the adjusted CUSUM of squared type test. All simulations are computed in the open-source statistical programming language R (2004). We consider the sample size $T = 500$ and use $M = 2000$ Monte Carlo repetitions for each experiment, while simulated critical values are obtained by setting $T = 10000$ and $M = 20000$. The memory parameter d is treated as unknown in order to achieve realistic conditions and is therefore estimated. We use the log-periodogram regression introduced by Geweke and Porter-Hudak (1983) with a rate of frequencies of $o(T^{0.8})$ which is MSE-optimal. First, we investigate the behaviour of the Leybourne and Kim (2007) test when the DGP exhibits long-range dependencies without a break in persistence, i.e. $(1-B)^d X_t = \varepsilon_t$ with $\varepsilon_t \overset{iid}{\sim} N(0,1)$. The long memory parameter d takes the values 0.00, 0.10, 0.25, 0.40, 0.60, 0.75, 0.90 and 1.00. This means that both, the null and the alternative hypothesis of the Leybourne and Kim (2007) test are wrong except the case of $d = 1$. As the Leybourne and Kim (2007) test is known to be conservative for a process being constantly $I(0)$ we would expect a similar behaviour in our setup. As we can see from Table 1 the originally proposed test exhibits serious size distortions in the presence of long-range dependencies resulting in a conservative test. Unsurprisingly, the empirical size is closer to the nominal significance level

Table 1: Empirical size using unadjusted critical values

	de-meaning						de-trending					
d	1.0L	5.0L	10.0L	10.0U	5.0U	1.0U	1.0L	5.0L	10.0L	10.0U	5.0U	1.0U
0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.40	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.60	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.75	0.0	0.0	1.7	1.5	0.0	0.0	0.0	1.0	2.1	2.4	0.1	0.0
0.90	0.0	2.5	6.5	6.9	2.3	0.0	0.0	3.3	6.2	6.3	2.6	0.0
1.00	1.1	5.2	10.0	10.3	5.1	0.9	1.1	5.2	9.8	10.1	4.9	1.2

Notes: Sample size is $T = 500$, xL and xU denote the x-th lower and upper quantile of R under long-range dependencies, respectively.

Table 2: Empirical size using estimated response curves

	de-meaning						de-trending					
d	1.0L	5.0L	10.0L	10.0U	5.0U	1.0U	1.0L	5.0L	10.0L	10.0U	5.0U	1.0U
0.00	1.3	5.6	10.7	9.5	4.8	1.1	0.8	5.1	10.0	9.6	5.1	1.2
0.10	1.1	5.0	9.8	10.0	4.9	1.2	1.4	4.9	9.4	10.1	4.9	1.1
0.25	1.0	5.7	10.9	9.8	4.9	1.2	1.2	4.8	9.8	9.1	4.4	0.7
0.40	0.8	4.9	9.8	10.9	5.7	1.8	1.0	3.8	10.5	9.0	4.9	0.9
0.60	0.9	5.4	11.0	11.1	6.1	1.6	1.5	5.8	11.0	11.2	5.5	1.1
0.75	0.6	4.0	8.4	9.8	4.8	0.7	1.5	5.3	11.2	9.4	4.4	1.1
0.90	0.7	5.0	10.6	8.4	4.1	0.5	0.8	5.2	10.3	9.4	4.5	0.9
1.00	1.0	5.3	11.2	9.7	4.9	1.0	0.6	5.2	10.0	9.5	5.0	1.4

Notes: Sample size is $T = 500$, xL and xU denote the x-th lower and upper quantile of R under long-range dependencies, respectively.

as d approaches one. The test is correctly sized if $d = 1$. However, even for non-stationary DGPs the size distortions are not negligible. This results underline the need for adjusted critical values taking the long-range dependencies into account.

Next, we simulate the asymptotic distribution of the R statistic dependening on d in the following 99 cases: $d = 0.51, \ldots, 1.49$. Due to the fact that adjusted critical values depend on d they have to be tabulated for a wide range of possible values of d. As this is rather burdensome we fit polynomial functions in d to the sequence of critical values depending on d. This response curve is given by

$$q_\alpha(d) = \sum_{i=0}^{s} \beta_i d^i \,, \tag{2}$$

where q_α denotes the α-quantile of the asymptotic distribution of R and takes values on the grid 0.51,...,1.49 consisting of 99 equally spaced points. The polynomial order s is set equal to nine. Additionally, we tried other settings, but $s = 9$ appeared to be a satisfying choice. Parameters β_i are estimated via OLS. Following the general-to-specific approach, we eliminate in each step the most insignificant power of d and re-estimate the function by OLS until no further non-rejection of the hypotheses $H_0 : \beta_i = 0$ occur at the five percent level of significance. The final estimates are reported in Appendix B. In Figure 1 we display the simulated quantiles (y-axis) depending on $d = 0.51, \ldots, 1.49$ (x-axis) and the fit of the response curves (solid line) for the 5% and 95% level of significance for de-meaned and de-trended data.

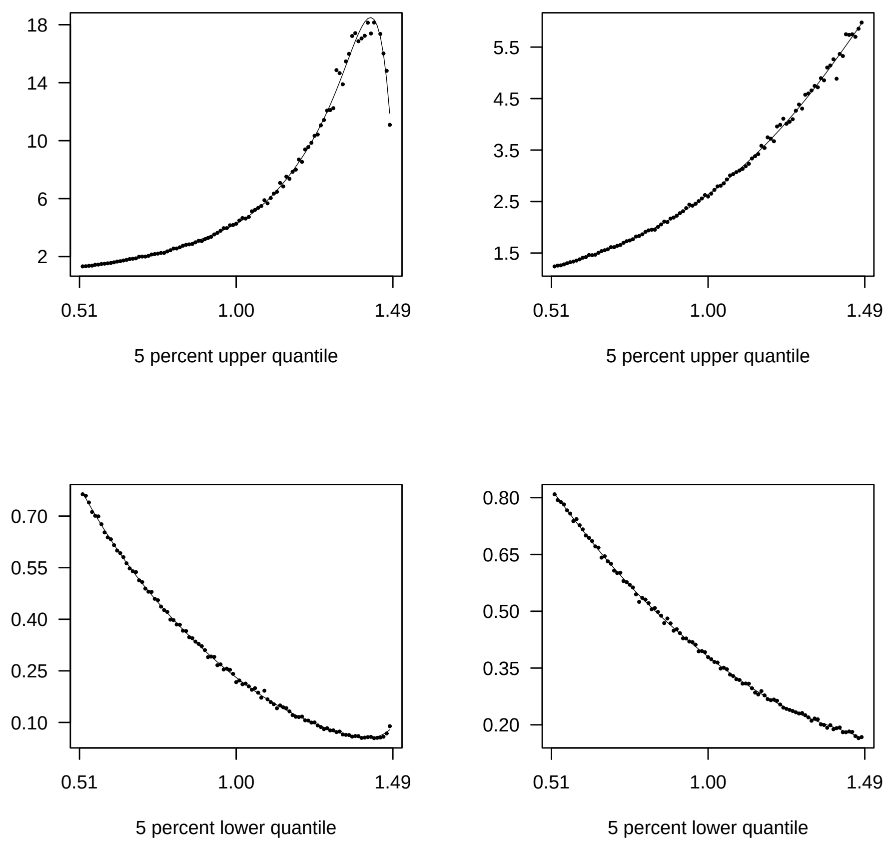

Figure 1: Simulated quantiles of the R statistic under long memory with fitted response curves for de-meaned (left) and de-trended data (right).

Using the fitted response curves we can approximate critical values easily. Beside the simplicity of this approach it is reasonable in our opinion as the variation of critical values is with the memory parameter and not with the sample size. The results in Table 2 that we discuss in a moment underline this argument.

Using adjusted critical values obtained from the fitted response curves we now revisit the empirical size of the test. The results are reported in Table 2. Note that time series with $\hat{d}_0 < 1/2$ are integrated in order to avoid a degenerated limit distribution under the null hypothesis. Henceforth, we actually consider a range of values from $d = 0.60$ to 1.40. The test has satisfying properties even though there are some minor distortions for values of d in the neighborhood of 1/2. It might happen that the estimated value of d is less (greater) then 1/2 under H_0 when the true value is greater (less) than 1/2, which means that we wrongly integrate (not integrate) the time series and obtain therefore a biased test result. However, the results in Table 2 suggest that this is not really a serious problem.

Next we consider the power of the test based on adjusted critical values. As the test can be seen as correctly sized, there is no need for size-adjusted critical values. For all power experiments we consider three different locations of the breakpoint, at the beginning $(\tau = 0.3)$, the middle $(\tau = 0.5)$ and the end $(\tau = 0.7)$ of the sample period. The long memory parameter takes the same values as before. The simulation results are given in Table 3 and 4 for de-meaned and de-trended data, respectively. We consider breaks from stationary to non-stationary long memory (upper left part of Tables 3 and 4) and vice versa (upper right part of Tables 3 and 4). The consistency results in Theorem 2 suggest that the test is consistent in both cases.

Furthermore, we consider in the lower part of Tables 3 and 4 breaks inside of the stationary $(0 \leq d < 1/2)$ (lower left part of Tables 3 and 4) and non-stationary region $(1/2 < d < 3/2)$ (lower right part of Tables 3 and 4). Note that the latter experiments are not covered by any of our theorems but that they might be relevant in empirical applications.

Overall, the power results of the adjusted test using de-meaned data (Table 3) is good and confirms the consistency result. For quite extreme breaks, e.g. 0.00 to 0.90, the power is almost hundred percent. Unsurprisingly, the power decreases for less extreme breaks, e.g. 0.00 to 0.60. For a given value of d_2, the power decreases with increasing d_1 (left part). For a given value of d_2, the power decreases with decreasing d_1 (right part). At first sight, it might be not intuitive that the power for breaks of equal distance, e.g. 0.25 to 0.75 and 0.40 to 0.90, is not the same. This is an artefact of the adjusted test introduced by integrating the time series if the estimated value of d is located in the stationary region. In addition, the fact that the breakpoint influences the estimate of d under H_0 further complicates the interpretation. However, the test is able to detect switches of the long memory parameter within the stationary and non-stationary region. The main conclusions are not changing when looking at the results for de-trended data.

After evaluating the power we consider the small sample performance of the simple breakpoint

Table 3: Power Experiment with de-meaned data

d	τ			d	τ		
	0.3	0.5	0.7		0.3	0.5	0.7
$0.00 \rightarrow 0.60$	81.9	84.1	78.3	$1.00 \rightarrow 0.00$	98.7	100.0	100.0
$0.00 \rightarrow 0.75$	99.2	97.8	87.9	$1.00 \rightarrow 0.10$	99.1	100.0	100.0
$0.00 \rightarrow 0.90$	100.0	100.0	97.0	$1.00 \rightarrow 0.25$	99.5	100.0	99.1
$0.00 \rightarrow 1.00$	100.0	100.0	99.6	$1.00 \rightarrow 0.40$	99.3	99.7	95.8
$0.10 \rightarrow 0.60$	69.0	78.4	64.3	$0.90 \rightarrow 0.00$	93.4	99.9	100.0
$0.10 \rightarrow 0.75$	96.7	97.2	82.7	$0.90 \rightarrow 0.10$	95.4	99.8	99.7
$0.10 \rightarrow 0.90$	99.9	100.0	97.2	$0.90 \rightarrow 0.25$	96.1	100.0	98.1
$0.10 \rightarrow 1.00$	100.0	100.0	99.8	$0.90 \rightarrow 0.40$	96.6	98.5	89.6
$0.25 \rightarrow 0.60$	42.1	70.6	50.7	$0.75 \rightarrow 0.00$	65.5	94.5	99.0
$0.25 \rightarrow 0.75$	86.5	97.5	83.2	$0.75 \rightarrow 0.10$	69.9	95.7	97.0
$0.25 \rightarrow 0.90$	98.2	100.0	98.0	$0.75 \rightarrow 0.25$	80.9	96.7	85.8
$0.25 \rightarrow 1.00$	99.5	100.0	99.6	$0.75 \rightarrow 0.40$	84.8	88.0	59.6
$0.40 \rightarrow 0.60$	20.8	50.0	51.2	$0.60 \rightarrow 0.00$	20.2	48.6	67.1
$0.40 \rightarrow 0.75$	54.7	88.7	86.4	$0.60 \rightarrow 0.10$	25.4	55.1	61.3
$0.40 \rightarrow 0.90$	87.4	98.9	98.0	$0.60 \rightarrow 0.25$	37.5	64.3	40.9
$0.40 \rightarrow 1.00$	95.8	99.9	99.6	$0.60 \rightarrow 0.40$	49.4	47.8	19.2
$0.00 \rightarrow 0.10$	82.7	84.8	82.5	$1.00 \rightarrow 0.60$	81.9	86.4	68.5
$0.00 \rightarrow 0.25$	77.4	82.0	81.0	$1.00 \rightarrow 0.75$	45.1	45.1	36.8
$0.00 \rightarrow 0.40$	64.6	79.6	77.0	$1.00 \rightarrow 0.90$	13.9	14.5	12.4
$0.10 \rightarrow 0.25$	59.5	68.1	67.3	$0.90 \rightarrow 0.60$	68.5	65.9	46.3
$0.10 \rightarrow 0.40$	48.5	62.9	61.9	$0.90 \rightarrow 0.75$	24.7	25.0	15.8
$0.25 \rightarrow 0.40$	21.7	33.1	32.3	$0.75 \rightarrow 0.60$	36.4	31.4	16.2

Table 4: Power Experiment with de-trended data

d		τ		d		τ	
	0.3	0.5	0.7		0.3	0.5	0.7
$0.00 \rightarrow 0.60$	45.1	84.7	94.2	$1.00 \rightarrow 0.00$	100.0	99.9	97.2
$0.00 \rightarrow 0.75$	64.1	95.8	99.5	$1.00 \rightarrow 0.10$	100.0	99.9	96.6
$0.00 \rightarrow 0.90$	88.6	99.7	99.9	$1.00 \rightarrow 0.25$	99.9	99.6	95.2
$0.00 \rightarrow 1.00$	96.5	99.9	100.0	$1.00 \rightarrow 0.40$	99.0	98.8	89.3
$0.10 \rightarrow 0.60$	40.8	76.6	87.1	$0.90 \rightarrow 0.00$	99.9	99.7	90.8
$0.10 \rightarrow 0.75$	61.1	93.7	97.6	$0.90 \rightarrow 0.10$	99.6	99.8	90.0
$0.10 \rightarrow 0.90$	89.3	99.6	99.7	$0.90 \rightarrow 0.25$	98.5	98.6	85.4
$0.10 \rightarrow 1.00$	97.1	99.9	99.9	$0.90 \rightarrow 0.40$	97.1	95.6	73.7
$0.25 \rightarrow 0.60$	33.6	57.4	64.9	$0.75 \rightarrow 0.00$	99.1	96.1	64.8
$0.25 \rightarrow 0.75$	53.7	87.4	92.0	$0.75 \rightarrow 0.10$	96.5	94.6	63.1
$0.25 \rightarrow 0.90$	85.5	98.4	99.0	$0.75 \rightarrow 0.25$	91.0	88.6	57.2
$0.25 \rightarrow 1.00$	95.0	99.7	99.8	$0.75 \rightarrow 0.40$	80.3	71.3	42.8
$0.40 \rightarrow 0.60$	19.6	33.1	41.7	$0.60 \rightarrow 0.00$	94.1	84.3	46.3
$0.40 \rightarrow 0.75$	39.1	71.7	77.9	$0.60 \rightarrow 0.10$	85.9	75.9	42.6
$0.40 \rightarrow 0.90$	72.5	94.4	96.4	$0.60 \rightarrow 0.25$	65.4	58.7	32.7
$0.40 \rightarrow 1.00$	89.3	98.5	99.5	$0.60 \rightarrow 0.40$	42.1	34.5	19.0
$0.00 \rightarrow 0.10$	7.3	13.5	21.8	$1.00 \rightarrow 0.60$	85.0	85.3	63.6
$0.00 \rightarrow 0.25$	27.5	39.4	49.5	$1.00 \rightarrow 0.75$	48.0	47.1	34.4
$0.00 \rightarrow 0.40$	52.4	69.3	75.8	$1.00 \rightarrow 0.90$	14.5	15.7	12.3
$0.10 \rightarrow 0.25$	12.8	20.5	26.0	$0.90 \rightarrow 0.60$	65.5	64.7	39.8
$0.10 \rightarrow 0.40$	33.6	46.6	53.5	$0.90 \rightarrow 0.75$	23.4	24.0	17.6
$0.25 \rightarrow 0.40$	13.4	16.3	25.7	$0.75 \rightarrow 0.60$	29.3	26.6	16.6

Table 5: Small sample performance of breakpoint estimator

d	τ	de-meaning 0.3	de-meaning 0.5	de-meaning 0.7	de-trending 0.3	de-trending 0.5	de-trending 0.7
$U[0,0.4] \rightarrow U[0.6,1]$	$\hat{\tau}^f$	0.378	0.546	0.725	0.442	0.557	0.730
	$\text{se}(\hat{\tau}^f)$	0.139	0.076	0.035	0.183	0.087	0.038
$U[0.6,1] \rightarrow U[0,0.4]$	$\hat{\tau}^r$	0.275	0.453	0.610	0.271	0.440	0.568
	$\text{se}(\hat{\tau}^r)$	0.034	0.075	0.148	0.036	0.087	0.178
$U[0,0.2] \rightarrow U[0.8,1]$	$\hat{\tau}^f$	0.327	0.521	0.717	0.340	0.524	0.717
	$\text{se}(\hat{\tau}^f)$	0.058	0.037	0.026	0.078	0.043	0.026
$U[0.8,1] \rightarrow U[0,0.2]$	$\hat{\tau}^r$	0.281	0.477	0.672	0.279	0.473	0.660
	$\text{se}(\hat{\tau}^r)$	0.025	0.039	0.055	0.027	0.044	0.076

Notes: Sample size is $T = 500$, $U[i,j]$ denotes the uniform distribution with lower and upper bound i and j, respectively; $\hat{\tau}^f$ and $\hat{\tau}^r$ denote the break point estimator based on the forward (f) and reversed series (r), respectively; $\text{se}(\cdot)$ is the standard error.

estimator, see Theorem 3. We use the same three different break points as before, i.e. $\tau = 0.3, 0.5, 0.7$. The memory parameter switches from the stationary to the non-stationary region and vice versa. For both regions we draw the memory parameter from a uniform distribution in order to cover a wide range of possible values in a small number of experiments. The upper and lower bounds are set equal to $[0, 0.4]$ and $[0.6, 1]$ for the stationary and the non-stationary region, respectively. In a more restrictive setting we set them equal to $[0, 0.2]$ and $[0.8, 1]$. Results for a sample size of five hundred observations are reported in Table 5. The overall impression of the breakpoint estimator's performance is satisfying. Noteworthy, we observe that the breakpoint estimator performs worse in the case of de-trended data which is due to an additional nuisance parameter that has to be estimated, c.f. Leybourne and Kim (2007). We further note that $\hat{\tau}^f$ and $\hat{\tau}^r$ perform better in the more restrictive setting because the break point becomes easier to detect.

5 Empirical Illustration

In this section we consider an empirical application of the adjusted test to study whether there is any change in persistence in US inflation. Hassler (1995) reported that such time series

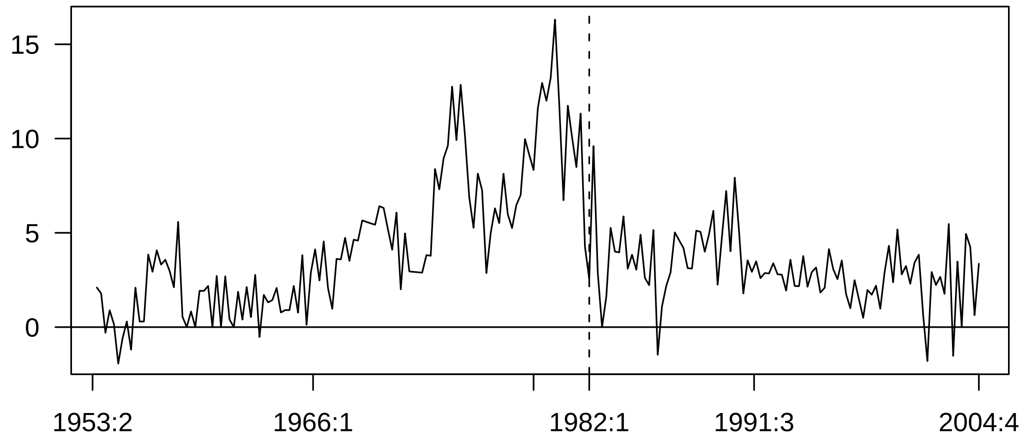

Figure 2: Time series plot of US inflation.

exhibits long-range dependencies. In addition, there is a small but growing literature dealing with the persistence of US inflation. Main questions in this literature are the measurement of persistence and potential structural breaks in it. In a recent article, Pivetta (2007) come to the conclusion that persistence of US inflation is approximately constant over time. The authors argue that their conclusion is in line with Stock and Watson (2003) as well as O'Reilly (2005). However, Kim et al. (2002) provide evidence for a decline in persistence at the very end of the seventies. Nonetheless, to the best of our knowledge, previous studies are not concerned with long-range dependencies and there is no previous study dealing with a structural break in the long memory parameter regarding inflation time series. Therefore, we try to add some new evidence by applying our long - memory adjusted test.

We use the quarterly CPI data from Lanne (2006) and transform it to annualized inflation by computing $y_t = 400 \ln(\text{CPI}_t/\text{CPI}_{t-1})$. The CPI data spans from 1953:1 to 2004:4 implying 207 observations. The time series plot is depicted in Figure 2. The graph suggests a decline of persistence in the second half of the sample which might be a result of Volcker's policy to pull inflation down from its high level in the seventies. The vertical line at 1982:1 shows the estimated break point of our test which will be discussed below.

Under the null hypothesis we obtain an estimated value of d_0 via the GPH approach with MSE-optimal rate of frequencies that equals 0.617 indicating non-stationary long-memory. We apply the adjusted test without integrating the time series under consideration since the asymptotic distribution of the test statistic is not degenerated as long as $d_0 > 1/2$ holds. Furthermore, we de-mean the data in a first step, since a clear linear trend is not obvious. Testing the null hypothesis of constant memory against decreasing memory gives a test statistic of 1.801 which is significant at the ten and five percent level of significance. Note that we make use of our estimated response curves to approximate the relevant critical values. The test result suggests that there is a decline in the persistence of US inflation. Interestingly, the estimated break point is 1982:1 which is nine quarters after the begin of Volcker's chairmanship at the Federal Reserve. When estimating the long memory parameter d before and after this breakpoint we get 0.862 and 0.246 which can be viewed as a sharp decline in persistence. Finally, we tested the null hypothesis of constant memory for the time period after 1982:1. Note that we have to integrate the time series once, since the asymptotic distribution is degenerated for $0 \leq d_0 < 1/2$. The test statistic is now 1.667 and insignificant at conventional levels. Although this result is based only on 91 observations, it suggests that there is no additional break after 1982:1.

6 Conclusion

In this paper we present a modification of a test proposed by Leybourne and Kim (2007) that allows for long memory dynamics. In particular, the test is constructed for the null hypothesis that there is no change in the long-memory parameter d against the alternative that it breaks from a stationary value $(0 \leq d < 1/2)$ to a non-stationary one $(1/2 < d < 3/2)$ or vice versa. We derive several asymptotic properties of the test statistic under long-range dependent DGPs and show that the asymptotic distribution depends on d. Therefore, we propose response curves based on estimates for d to obtain the relevant critical value easily and show by means of a Monte Carlo study that this approach works well. Furthermore, the power of the test is good and a simple breakpoint estimator has satisfying properties. Finally, we apply the test to US inflation data and find a break from non-stationary to stationary long-memory in the early eighties.

Appendix A

Proof of Theorem 1: For the proof of the theorem let us consider the de-meaned case first. The test statistic was defined by

$$R = \frac{\inf_{\tau\in\Lambda} K^f(\tau)}{\inf_{\tau\in\Lambda} K^r(\tau)},$$

with

$$K^f(\tau) = [\tau T]^{-2} \sum_{t=1}^{[\tau T]} \hat{v}_{t,\tau}^2$$

and

$$K^r(\tau) = (T-[\tau T])^{-2} \sum_{t=1}^{T-[\tau T]} \tilde{v}_{t,\tau}^2.$$

For the nominator we have

$$T^{-d_0-\frac{1}{2}}\hat{v}_{t,\tau} = T^{-d_0-\frac{1}{2}}x_t - T^{-d_0-\frac{1}{2}}\overline{x}(\tau).$$

We have

$$T^{-d_0-\frac{1}{2}}x_{[rt]} \Rightarrow W_{d_0}(r)$$

with W_{d_0} denoting fractional Brownian motion with parameter d_0. Furthermore we have

$$\begin{aligned} T^{-d_0-\frac{1}{2}}\overline{x}(\tau) &= T^{-d_0-\frac{1}{2}}[\tau T]^{-1}\sum_{t=1}^{[\tau T]} x_t \\ &= T^{-d_0-\frac{3}{2}}\tau^{-1}\sum_{t=1}^{[\tau T]} x_t \\ &\Rightarrow \tau^{-1}\int_0^\tau W_{d_0}(r)dr. \end{aligned}$$

Application of the continuous mapping theorem gives

$$\begin{aligned} T^{-d_0-\frac{1}{2}}\hat{v}_{[rt]} &\Rightarrow W_{d_0}(r) - \tau^{-1}\int_0^\tau W_{d_0}(r)dr \\ &=: W^*_{d_0}(r,\tau) \end{aligned}$$

and thus for the nominator

$$
\begin{aligned}
T^{-2d_0}K^f(\tau) &= \tau^{-2}\int_0^{\tau}(T^{-d_0-\frac{1}{2}}\hat{v}_{[rT]})^2 dr \\
&\Rightarrow \tau^{-2}\int_0^{\tau} W^*_{d_0}(r,\tau)^2 dr.
\end{aligned}
$$

Similarly we obtain for the denominator

$$
\begin{aligned}
T^{-d_0-\frac{1}{2}}\tilde{v}_{[rT]} &\Rightarrow W_{d_0}(1)-W_{d_0}(1-r)+(1-\tau)^{-1}\int_0^{1-\tau}(W_{d_0}(1)-W_{d_0}(1-r)dr) \\
&= W_{d_0}(1-r)-(1-\tau)^{-1}\int_{\tau}^{1} W_{d_0}(r)dr \\
&=: V^*_{d_0}(r,\tau).
\end{aligned}
$$

Again using the continuous mapping theorem we obtain for the denominator

$$
\begin{aligned}
T^{-2d_0}K^r(\tau) &= (1-\tau)^{-2}\int_0^{1-\tau}(T^{-d_0-\frac{1}{2}}\tilde{v}_{[rT]})^2 dr \\
&\Rightarrow (1-\tau)^{-2}\int_0^{1-\tau} V^*_{d_0}(r,\tau)^2 dr.
\end{aligned}
$$

Combining the result for the nominator and the denominator gives the result.

The result for the de-meaned and de-trended case is obtained by applying standard results for linear regression with long-memory errors. We consider the forward statistic

$$
K^f(\tau) = [\tau T]^{-2}\sum_{t=1}^{[\tau T]}\hat{v}_t^2,
$$

where $\hat{v}_t = x_t - \hat{\alpha} - \hat{\beta}t$ are the residuals from the OLS regression of x_t on the vector $z_t = [1,t]'$, $t = 1,\ldots,[\tau T]$. It is well known that $T^{-d_0-1/2}(\hat{\alpha}-\alpha) \Rightarrow B_0(\tau)$ and $T^{-d_0-1/2}(\hat{\beta}-\beta) \Rightarrow B_1(\tau)$ and $B_0(\tau)$ and $B_1(\tau)$ given as in the Theorem. Therefore, we obtain

$$
\begin{aligned}
T^{-d_0-1/2}\hat{v}_{[\tau T]} &= T^{-d_0-1/2}v_{[\tau T]} - T^{-d_0-1/2}(\hat{\alpha}-\alpha) - rT^{-d_0-1/2}(\hat{\beta}-\beta) \\
&\Rightarrow W_{d_0}(r) - B_0(\tau) - rB_1(\tau) \\
&\equiv W^{**}_{d_0}(r,\tau).
\end{aligned}
$$

As the forward statistic is a continuous functional of $T^{-d_0-1/2}\hat{v}_{[\tau T]}$ we obtain using the CMT

$$
\begin{aligned}
T^{-2d_0}K^f(\tau) &= \tau^{-2}\int_0^{\tau}(T^{-d_0-1/2}\hat{v}_{[\tau T]})^2 dr \\
&\Rightarrow \tau^{-2}\int_0^{\tau} W^{**}_{d_0}(r,\tau)^2 dr \\
&\equiv L^f(\tau).
\end{aligned}
$$

The proof for the reverse statistic is analogous and therefore omitted. For the remainder of the Appendix we omit proofs for the de-meaned and the de-trended case for the brevity of notation as they are straightforward. $\diamondsuit$

Proof of Theorem 2: First we prove the first part of the theorem. This is we assume a breakpoint that the DGP breaks from a stationary to a non-stationary long-memory process. Let us first consider the situation of $\tau \leq \tau_0$, where τ_0 denotes the true breakpoint. This means that $X_t \sim I(d_1)$ with $0 \leq d_1 < 1/2$. Have in mind that the standardization of the test statistic is obtained from $H_0 : X_t \sim I(d_0)$ with $d_1 \neq d_0$. In the stationary part we have $d_0 \geq d_1$. In this situation we obtain:

$$\begin{aligned} T^{-2d_0+1} K_{d_0}^f(\tau) &= \tau^{-1} T^{-2d_0} [\tau T]^{-1} \sum_{t=1}^{[\tau T]} v_t^2 \\ &\xrightarrow{P} \tau^{-1} O(T^{d_1 - d_0}). \end{aligned}$$

In the case of $d_1 = d_0$ the upper expression converges to $\tau^{-1}\gamma_0$ with γ_0 denoting the variance of X_t. For $d_1 < d_0$, wich is the relevant case in practise, this expression tends to zero with a rate depending on the difference of the true d_0 before the break and the hypothetic memory parameter.

We next consider the situation of $\tau > \tau_0$ where we split $K_d^f(\tau)$ up in its stationary and its non-stationary part. Have in mind that the true DGP is of order $1/2 < d_2 < 3/2$ after the break with $d_2 > d_0$ in the non-stationary part.

$$\begin{aligned} T^{-2d_0} K_{d_0}^f(\tau) &= \tau^{-2} T^{2-2d_0} \sum_{t=1}^{[\tau T]} x_t^2 - \tau^{-3} \left(T^{3/2-d_0} \sum_{t=1}^{[\tau T]} y_t \right)^2 \\ &= \tau^{-2} \left(T^{2-2d_0} \sum_{t=1}^{[\tau_0 T]} x_t^2 + T^{2-2d_0} \sum_{t=[\tau_0 T]+1}^{[\tau T]} x_t^2 \right) \\ &\quad -\tau^{-3} \left(T^{3/2-d_0} \sum_{t=1}^{[\tau_0 T]} y_t + T^{3/2-d_0} \sum_{t=[\tau_0 T]+1}^{[\tau T]} y_t \right)^2 \\ &= \tau^{-2} T^{2-2d_0} \sum_{t=[\tau_0 T]+1}^{[\tau T]} x_t^2 - \tau^{-3} \left(T^{3/2-d_0} \sum_{t=[\tau_0 T]+1}^{[\tau T]} y_t \right)^2 + o_P(1) \\ &\xrightarrow{P} O_P(T^{d_2 - d_0}). \end{aligned}$$

From these considerations we see that the limit of $T^{-2d_0+1}K^f_{d_0}(\tau)$ is given by $\tau^{-1}O_P(T^{d_1-d_0})+\infty 1_{(\tau>\tau_0)}$ which is obviously minimized by τ_0. Thus, we have

$$T^{-2d_0+1}\inf_{\tau\in\Lambda} K^f_{d_0}(\tau) \xrightarrow{P} O_P(T^{d_1-d_0}).$$

For the reversed series we obtain by similar arguments for $\tau \leq \tau_0$:

$$T^{-2d_0}K^r_{d_0}(\tau) \xrightarrow{P} O_P(T^{d_2-d_0}).$$

This gives us

$$\inf_{\tau\in\Lambda} K^r_{d_0}(\tau) = O_P(T^{d_2+d_0-1})$$

which gives us the first result of the theorem. The result in (2) is obtained by similar arguments as above. $\diamondsuit$

Proof of Theorem 3: Let us assume a break from stationary long memory to non-stationary long memory, that is $0 \leq d_1 < 1/2$ and $1/2 < d_2 < 3/2$. The hypothetical memory parameter is denoted by d_0 with $d_1 \leq d_0 \leq d_2$. From Theorem 2 we know that the limit of $T^{-2d_0+1}K^f_{d_0}(\tau)$ is given by $O_P(T^{d_1-d_0})1_{(\tau\leq\tau_0)} + \infty 1_{(\tau>\tau_0)}$ which is obviously minimized by τ_0. The result follows now by similar arguments as in Leybourne and Kim (2007).

The proof for the second part of the theorem, that is the break from non-stationary to stationary long memory, is analogous and therefore omitted here. $\diamondsuit$

Proof of Theorem 4: Because of the symmetry of Λ around 0.5 we have

$$\begin{aligned}
\inf_{\tau\in\Lambda} T^{-2d_0}K^f_{d_0}(\tau) &\xrightarrow{P} \inf_{\tau\in\Lambda} \tau^{-1}\gamma_0 \\
&= \lambda_u^{-1}\gamma_0 \\
\inf_{\tau\in\Lambda} T^{-2d_0}K^r_{d_0}(\tau) &\xrightarrow{P} \inf_{\tau\in\Lambda}(1-\tau^{-1})\gamma_0 \\
&= (1-\lambda_l^{-1})\gamma_0 \\
&= \lambda_u^{-1}\gamma_0,
\end{aligned}$$

where λ_u and λ_l denote the upper and the lower bound of the interval Λ respectively. This proves the theorem. $\diamondsuit$

Appendix B

Table 6: Estimated response curves for de-meaned data

Quantile	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
1.0L	1.063	0	0	0	-41.002	133.627	-183.98	131.206	-47.89	7.102
5.0L	1.601	0	-7.486	9.449	0	0	-17.596	25.299	-13.724	2.688
10.0L	-221.524	2316.11	-10522.512	27414.943	-45191.318	48907.541	-34769.527	15666.998	-4062.561	462.173
10.0U	5145.518	-54469.126	252323.451	-671384.183	1131196.84	-1252080.53	910897.739	-420239.255	111628.329	-13015.697
5.0U	10493.76	-110784.01	511682.48	-1357262	2279365.93	-2514370.73	1822761.11	-837851.29	221721.78	-25752.77
1.0U	-1174.527	0	58540.259	-312952.617	792898.52	-1170633.31	1062803.45	-586254.848	180679.152	-23898.266

Notes: xL and xU denote the x-th lower and upper quantile of R. OLS estimates for β_i $(i = 0, 1, \ldots, 9)$ in (2) are reported in columns; $\beta_i = 0$ means that the parameter is set equal to zero.

Table 7: Estimated response curves for de-trended data

Quantile	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
1.0L	1.051	0	0	-4.815	0	18.496	-25.406	13.556	-2.63	0
5.0L	1.151	0	0	-9.281	21.702	-21.366	9.999	-1.824	0	0
10.0L	-0.455	0	53.424	-234.177	459.766	-499.311	310.551	-103.809	14.485	0
10.0U	1.054	0	0	0	3.328	-3.117	0.868	0	0	0
5.0U	1.008	0	0	0	8.274	-13.18	8.509	-1.971	0	0
1.0U	1.187	0	0	0	6.272	-5.03	1.557	0	0	0

Notes: See Table 6.

References

Banerjee, A., R. Lumsdaine, and J. Stock (1992), “Recursive and sequential tests of the unit root and trend break hypothesis: theory and international evidence.” *Journal of Business and Economics Statistics*, 10, 271–288.

Beran, J., and N. Terrin (1996), “Testing for a change of the long-memory parameter.” *Biometrika*, 83, 627–638.

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

Geweke, J., and S. Porter-Hudak (1983), “The estimation and application of long - memory time series models.” *Journal of Time Series Analysis*, 4, 221–238.

Granger, Joyeux R., C. (1980), “An introduction to long-range time series models and fractional differencing.” *Journal of Time Series Analysis*, 1, 15–30.

Hassler, Wolters J., U. (1995), “Long memory in inflation rates: International evidence.” *Journal of Business and Economics Statistics*, 13, 1326–1358.

Kim, J. (2000), “Detection of change in persistence of a linear time series.” *Journal of Econometrics*, 95, 97–116.

Kim, J., J. Belaire Franch, and R. Badilli Amador (2002), “Corrigendum to ”Detection of change in persistence of a linear time series.”,” *Journal of Econometrics*, 109, 389–392.

Lanne, M. (2006), “Nonlinear Dynamics of Interest Rate and Inflation.” *Journal of Applied Econometrics*, 21, 1157–1168.

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

Leybourne, Taylor R., S., and T. Kim (2007), “CUSUM of squares - based tests for a change in persistence.” *Journal of Time Series Analysis*, 28, 408–433.

O'Reilly, Whelan K., G. (2005), "Has Euro Area inflation persistence changed over time?" *The Review of Economics and Statistics*, 87, 709–720.

Pivetta, Reis R., F. (2007), "The persistence of inflation in the United States." *Journal of Economic Dynamics and Control*, 31, 1326–1358.

R, Development Core Team (2004), "R: A language and environment for statistical computing." *Available at www.r-project.org.*

Stock, J., and M. Watson (2003), "Has the business cycle changed and why?" *In: M. Gertler and K. Rogoff, Editors, NBER Macroeconomics Annual 2002, MIT Press, Cambridge*.