Monitoring a change in persistence of a long range dependent time series*

Florian Heinen[1], Juliane Willert

Institute of Statistics, Faculty of Economics and Management,
Leibniz University of Hannover, D-30167 Hannover, Germany

Abstract
We consider the detection of a change in persistence of a long range dependent time series. The usual approach is to use one-shot tests to detect a change in persistence a posteriori in a historical data set. However, as breaks can occur at any given time and data arrives steadily it is desirable to detect a change in persistence as soon as possible. We propose the use of a MOSUM type test which allows sequential application whenever new data arrives. We derive the asymptotic distribution of the test statistic and prove consistency. We further study the finite sample behavior of the test and provide an empirical application.

JEL-Numbers: C12, C22
Keywords: Change in persistence · Long range dependency · MOSUM test

1 Introduction

The assumption of structural stability of an econometric model is a major issue in time series econometrics. If the parameter estimates stem from an unstable relationship they are not meaningful and additionally inference can be biased and forecasts yield inaccurate results (see e.g. Hansen (2001), Andrews and Fair (1988), Ghysel et al. (1997), Garcia and Perron (1996) or Clements and Hendry (1998)). In reaction to these findings a large amount of literature emerged that incorporated structural change in the inference techniques or analyzes forecasting subject to structural change more closely (see

*Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.
[1]Corresponding author:
Phone: +49-511-762-5383
Fax: +49-511-762-3923
E-Mail: heinen@statistik.uni-hannover.de

e.g. Perron (1989), Zivot and Andrews (1992) or Pesaran and Timmermann (2005)). Recently the possibility of a change in persistence, i.e. a change in the memory structure of the time series as a special case of structural instability, has become object of study (see e.g. Kim (2000), Kim et al. (2002), Busetti and Taylor (2004), Banerjee et al. (1992), Leybourne et al. (2003) or Leybourne et al. (2007)). This work has been placed within the $I(0)$ vs. $I(1)$, or vice versa, framework where the focus lies on short memory time series with an exponentially decaying autocorrelation structure.

However, since the seminal papers of Granger and Joyeux (1980) and Hosking (1981), long memory time series have become widely used in economics to model highly persistent time series as diverse as inflation rates or realized volatility (see e.g. Hassler and Wolters (1995) and Corsi et al. (2008)). Baillie (1996) provides an overview about various applications of long memory time series in economics.

Despite these facts little work has been done to test for a change in persistence in long range dependent time series. Notable exceptions are Beran and Terrin (1996), Ray and Tsay (2002), Sibbertsen and Kruse (2009) or Yamaguchi (2011). These test belong to the class of so-called "one-shot" tests (see Chu et al. (1996, p. 1045)), i.e. tests that are applied a posteriori to detect a structural break within a historical data set.

Because breaks can occur at any given time and also new data arrives steadily it is desirable for the applied econometrician to detect a change in persistence as soon as possible. This leads to a sequential testing problem (see Siegmund (1985) for an overview). As the usual "one-shot" tests work with constant critical values they cannot be applied sequentially given that the true null of no change would eventually be rejected with probability one (see Robbins (1970)). Starting with Bauer and Hackl (1978) a strand of literature has emerged that studies monitoring procedures that allow to detect structural change whenever new data arrives. Important contributions on this field are Chu et al. (1995), Kuan and Hornik (1995), Chu et al. (1996), Leisch et al. (2000), Altissimo and Corradi (2003), Zeileis et al. (2005), Andreou and Ghysels (2006) and Hsu (2007). These papers contribute to the literature on monitoring structural stability on different levels ranging from theoretical contributions to detecting structural change in the conditional mean or the conditional variance or comparing different types of rejection regions for the null.

In this paper we use a monitoring approach based on moving sums of residuals and place it into a long memory framework. We develop a procedure to detect an increase in persistence for the case that the process becomes non-stationary. This is important because an increase in persistence implies a loss of controllability for important macroeconomic time series such as inflation rate or the European overnight rate (EONIA) (see Sibbertsen and Kruse (2009) and Hassler and Nautz (2008)). Further, a change in persistence also affects forecast accuracy in long memory time series (see Heinen et al. (2009)).

The rest of the paper is organized as follows: In section 2 we describe the test procedure we use and develop the asymptotic behavior. We further discuss and motivate different forms of boundary functions for the test. In section 3 we undertake a simulation study to asses the finite sample performance of the monitoring test. Section 4 contains an empirical application before section 5 concludes. All proofs are collected in the appendix A.

2 Monitoring a change in persistence

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

$$\Phi(L)(1-L)^d y_t = \Theta(L)\varepsilon_t, \text{ with } \varepsilon_t \overset{iid}{\sim} (0,\sigma^2) \text{ and } t = 1,2,\ldots,T \ . \tag{1}$$

The differencing parameter d can take fractional values but is assumed to be $|d| < \frac{1}{2}$. Thus the process y_t is in the stationary region (see e.g. Beran (1995)).
Bauer and Hackl (1978) propose the use of moving sums of cumulated residuals (MOSUM) to detect parameter changes in regression models. These tests are further investigated by Chu et al. (1995).
We are interested in detecting a change in persistence, i.e. a change in the fractional differencing parameter d, in the monitoring period $T+1$ up to $[T\tau]$, $\tau > 1$. Where $[\cdot]$ denotes the integer part of its argument.
In particular, we test the null of no change in persistence, i.e. $d = d_0$ within the monitoring period where $|d_0| < \frac{1}{2}$, against the alternative of an increase in persistence. More formally we test the null that

$$H_0 : d_\ell = d_0, \quad \ell = T+1,\ldots,[T\tau] \ , \tag{2}$$

against the alternative that at some point in the monitoring period the persistence increases and $\frac{1}{2} < d_\ell < \frac{3}{2}$. Thus we test whether the process stays in the stationary region throughout the whole monitoring period or changes into the non-stationary region with an infinite variance at some point in the monitoring period. For the period from $t = 1,\ldots,T$ we follow Chu et al. (1996) and make the "noncontamination" assumption that

$$d_t = d_0, \quad t = 1,\ldots,T \ ,$$

with $|d_0| < \frac{1}{2}$. Consider for simplicity the case of an ARFIMA(0,d,0) process.
Let $\hat{e}_t$ be an ARFIMA(0,d,0) process as in (1) and $\hat{\sigma}^2 = T^{-1}\sum_{i=1}^{T}\hat{e}_i^2$ a consistent estimator of σ^2. Based on a moving sum of residuals obtained from a fixed window size $[Th]$,

$0 < h \leq 1$, the prototypical MOSUM test reads

$$MS_{T,h,d} = \max_{T+1 \leq k \leq [T\tau]} \sigma^{-1} T^{-\frac{1}{2}-d} \left| \sum_{i=k-[Th]+1}^{k} \hat{e}_i - \frac{[Th]}{T} \sum_{i=1}^{T} \hat{e}_i \right| , \tag{3}$$

for each value k in the monitoring period $T+1$ through $[T\tau]$.
The next theorem gives the asymptotic behavior of the test statistic in (3) if y_t follows a long range dependent process as in (1) and (2).

Theorem 1. *Assume the process y_t follows an ARFIMA$(0,d,0)$ process as in (1) with $|d| < \frac{1}{2}$. Then, as $T \to \infty$, we have for $MS_{T,h,d}$ in (3) that*

$$MS_{T,h,d} \Rightarrow \frac{1}{\sigma} \max_{t \in [1,\tau]} \left| BB^0(t,d) - BB^0(t-h,d) \right| ,$$

where $BB^0(t,d)$ denotes a fractional Brownian Bridge depending on fractional Brownian motion with parameter d. $\Rightarrow$ denotes weak convergence on a function space.
Under the alternative of a break in persistence the test is consistent.

The limiting distribution thus depends on the increments of a fractional Brownian bridge which in turn depends on the differencing parameter d of the data generating process. Therefore the asymptotic critical values of $MS_{T,h,d}$ are determined by the boundary crossing probabilities of the increments of a fractional Brownian bridge:

$$\mathcal{P}\{MS_{T,h,d} \leq b\} = \mathcal{P}\left\{ \left| BB^0(t,d) - BB^0(t-h,d) \right| \leq b \right\} . \tag{4}$$

The use of the test statistic in (3) is beneficial because the sequential application of usual CUSUM tests as in Sibbertsen and Kruse (2009) with constant critical values will eventually reject a correct null of no change in persistence with probability one (see Robbins (1970)).
Generally, every strictly increasing function $b(t) = zq(t)$ could serve as a boundary function where z is some suitable scaling factor and $q(t)$ is some monotonically increasing function in time. However if the boundary function grows too slowly the monitoring test will commit the type one error almost surely as it will detect a break in persistence with probability one. On the contrary if the boundary grows too quickly the test will loose power because a break in persistence cannot be detected anymore. For the short memory case a variety of different boundary functions have been proposed (see Andreou and Ghysels (2006, p. 92) for an overview). In particular Altissimo and Corradi (2003) derive a boundary function based on the almost sure asymptotically uniform equicontinuity of the Brownian bridge obtaining an almost sure boundary function. This is convenient because it gives the rate of convergence with which the sequence of functions

converges to a relatively compact set in the sense of an Arzelà-Ascoli theorem (see e.g. Davidson (1994, p. 335)). This provides useful information as we are interested in the behavior of the limiting distribution independently of the test statistic. We also derive almost sure results similar to the ones obtained by Altissimo and Corradi (2003) which are collected in the next theorem.

Theorem 2. *Let* $BB^0(t,d) = B(t,d) - tB(1,d)$ *be a fractional Brownian bridge. Then,* $d_T^{-1}|BB^0(t,d)|$ *is almost surely asymptotically uniform equicontinuous in* $t \in [0,1]$*. With* $d_T := \sqrt{2T^{2d+1}\log\log(T)}$*.*

The use of this theorem is that it provides the rate with which the increment of the fractional brownian bridge becomes asymptotically uniform equicontinuous. In the proof this derived to be $\sqrt{2\log\log(T)}$. Hence, if we use this growth rate for the boundary function we will obtain a slowly growing function and therefore detect a change in persistence but at the same time this function is independent of the long memory parameter under the null d_0.
Different forms of the boundary function are possible. For example one could use the boundary function

$$b_1(t) = z\sqrt{2t\log_2(t)}\,, \tag{5}$$

where $\log_2(t) := \log(\log(t))$. This boundary function is based on the law of iterated logarithm and is motivated by the fastest detection of change because it grows as slowly as possible. From theorem 2 we deduce the boundary function

$$b_2(t) = z\sqrt{2\log_2(t)}\,. \tag{6}$$

Because both boundary functions rely on the square root of a logarithm one needs to find a way to deal with values $\leq \log(1)$ to ensure real valued boundaries. One way of doing so is to define

$$\log_2'(t) := \begin{cases} 1, & \text{if } t \leq \exp(1) \\ \log\log(t), & \text{if } t > \exp(1)\,, \end{cases}$$

similar to Leisch et al. (2000). Another way which avoids the constant behavior of the boundary function at the beginning of the monitoring period is to define

$$\log_2''(t) := \begin{cases} t, & \text{if } t \leq \exp(1) \\ \log\log(t), & \text{if } t > \exp(1)\,. \end{cases}$$

Formally this leads to four possible boundary functions

$$
\begin{aligned}
b_3(t) &= z\sqrt{2t\log_2'(t)} && (7)\\
b_4(t) &= z\sqrt{2t\log_2''(t)} && (8)\\
b_5(t) &= z\sqrt{2\log_2'(t)} && (9)\\
b_6(t) &= z\sqrt{2\log_2''(t)}\,. && (10)
\end{aligned}
$$

One could think of different boundary functions such as functions that are dependent on the long memory parameter under the null to account for the gradually increasing variance of the process. However, unreported simulations showed that such a boundary function does not perform satisfactorily and we therefore restrict ourselves to the above boundary functions.

3 Monte Carlo evidence

We start by providing some Monte Carlo evidence on the small sample behavior of the usual MOSUM test as considered in Leisch et al. (2000) under long range dependence. Table 1 shows some of the simulation results.

	$\tau=4$			$\tau=6$			$\tau=8$		
d	$h=0.25$	$h=0.5$	$h=1$	$h=0.25$	$h=0.5$	$h=1$	$h=0.25$	$h=0.5$	$h=1$
0.1	66.22	57.26	51.08	70.04	62.80	53.72	72.52	67.12	57.02
0.2	98.06	96.38	91.90	99.28	97.94	94.90	99.76	98.98	96.86
0.3	99.96	99.90	99.64	100.00	100.00	99.96	100.00	100.00	99.96
0.4	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 1: Empirical size of the fluctuation test by Leisch et al. (2000) [in %] for $T=250$ and $\alpha=5\%$.

As expected the generalized fluctuation test does not keep its size. Even if the long memory is only moderately present the test does not allow a secure conclusion whether a change in persistence is present or not because the boundary functions are too narrow. In order to assess the finite sample performance of the monitoring procedure described in section 2 we consider different values for the long memory parameter $d=0.1,0.2,0.3,0.4$, the monitoring window $h=0.25,0.5,0.75,1$ and the out-of-sample monitoring period $\tau=2,4,6,8,10$. We also consider different sample sizes of $T=200,250,300$ and the different boundary functions $b_i(t)$, for $i=3,\dots,6$, from (7) to (10) for the simulations. The number of Monte Carlo repetitions is set to $M=10000$ and the levels of significance are set to $\alpha = 1\%, 5\%, 10\%$.[1]

Boundary function $b_3(t)$									
	$\tau = 4$			$\tau = 6$			$\tau = 8$		
d	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$
0.1	6.86	7.10	8.59	6.26	6.95	9.03	7.07	6.97	8.60
0.2	6.77	6.80	8.73	6.47	6.56	8.52	5.75	6.82	8.97
0.3	5.68	7.12	9.76	6.16	7.08	9.54	6.21	7.08	10.24
0.4	10.77	12.29	16.43	9.96	12.82	16.18	10.07	12.24	16.09

Table 2: Empirical size of the monitoring procedure [in %] for $T = 250$ and $\alpha = 5\%$.

Table 2 shows the size results for the boundary function motivated by the law of iterated logarithm. Using this boundary we obtain a procedure that is generally oversized. This overrejection of the correct null becomes more severe as the degree of persistence increases and/or the monitoring window h increases.
Table 3 displays the respective results based on the almost sure results from theorem 2. These results are more promising compared to the ones of boundary $b_3(t)$ as the nominal size level is better adhered to. Looking at the dependencies between the size, the long memory parameter d, the monitoring window h and the monitoring period τ we see that a moderate window size of $h = 0.5$ or $h = 0.75$ is generally preferable regardless of the monitoring period τ. If the persistence increases a reduced window size of $h = 0.5$ yields the most accurate size results. Reducing the window size even further to $h = 0.25$, however, leads to overrejection again as unreported results show.
As the boundary function $b_6(t)$ is only a slight modification of boundary function $b_5(t)$ the same argument as above applies to the results in table 4. The only difference is that the test overrejects somewhat when using boundary function $b_6(t)$.

Boundary function $b_5(t)$									
	$\tau = 4$			$\tau = 6$			$\tau = 8$		
d	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$
0.1	7.55	7.02	7.09	7.12	6.16	6.69	7.30	6.56	6.29
0.2	6.76	5.85	6.56	6.64	5.85	5.60	5.90	5.66	5.23
0.3	5.15	4.87	4.87	5.08	4.28	4.29	5.16	3.89	3.92
0.4	5.61	5.10	5.78	4.86	4.12	4.28	4.32	3.63	3.89

Table 3: Empirical size of the monitoring procedure [in %] for $T = 250$ and $\alpha = 5\%$.

[1] Some of the results here and in the sequel are unreported to save space but can be obtained from the authors on request.

Boundary function $b_6(t)$									
	$\tau = 4$			$\tau = 6$			$\tau = 8$		
d	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$
0.1	7.00	7.52	8.65	6.75	6.71	8.42	7.34	7.05	7.59
0.2	7.13	6.92	8.60	6.69	6.31	7.33	6.02	6.20	6.70
0.3	6.12	6.96	8.94	5.96	5.86	7.26	5.84	5.58	6.49
0.4	10.26	11.22	13.96	8.39	9.61	10.83	7.67	8.06	9.53

Table 4: Empirical size of the monitoring procedure [in %] for $T = 250$ and $\alpha = 5\%$.

The size results for the $\alpha = 10\%$ level are unreported but show the same general behavior of the previously discussed results. However, in this setting it becomes even more obvious that the boundary function $b_5(t)$ yields the best performance over all considered settings.

Generally the size distortions are minor and acceptable and also comparable to the short memory case as reported in Leisch et al. (2000).

In an empirical setting the long memory parameter d_0 is unknown and has to be estimated. We therefore conduct the size experiment again but this time using an estimated d_0. Generally every consistent estimation method is applicable but estimators that converge faster than the asymptotic distribution to the true value of d_0 are preferable. One such estimator is the approximate maximum likelihood estimator proposed by Beran (1995) which is $\sqrt{T}$ consistent. Another popular method to estimate d_0 is the log-periodogram regression (see Geweke and Porter-Hudak (1983)). The rate of convergence of this estimator is $\sqrt{m}$ where m is the number of frequencies used. The estimator is consistent as long as $(m\log(m))/n \to 0$ as $m, n \to \infty$, with n being the sample size (see Hurvich et al. (1998)). In our simulations we use this estimator with $T^{4/5}$ frequencies. The results are reported for the $\alpha = 5\%$ level in table 5.

Boundary function $b_5(t)$									
	$\tau = 4$			$\tau = 6$			$\tau = 8$		
d	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$	$h = 0.5$	$h = 0.75$	$h = 1$
0.1	8.68	7.62	8.50	8.62	7.14	7.48	7.72	7.20	6.42
0.2	7.12	6.62	6.02	7.44	6.08	5.82	6.36	5.86	5.42
0.3	5.74	4.98	5.26	5.14	4.28	4.40	4.26	4.24	4.04
0.4	5.36	4.86	5.38	4.32	4.36	4.34	3.74	3.42	4.04

Table 5: Empirical size of the monitoring procedure with estimated d_0.

We observe small size distortions for smaller values of d_0 and larger monitoring periods but generally the size is well kept even if we estimate the long memory parameter. When the persistence changes from stationary to non-stationary the MOSUM test will eventually detect this with probability one due to consistency (see theorem 1).[2] Therefore it is more interesting how fast a change in persistence can be detected.

To study the detection delay we consider breaks from the stationary region, namely $d_0 = 0.1, 0.2, 0.3, 0.4$, to the non-stationary region, $d_1 = 0.6, 0.7, 0.8, 0.9, 1$. The break occurs within the monitoring period at $t^* = [\rho\tau T]$, where $\rho = 0.3, 0.5, 0.7$ and $\tau = 2, 4, 6, 8, 10$ as above and $[\cdot]$ denotes the integer part of its argument. We use a sample size of $T = 250$ and the boundary functions $b_i(t)$, for $i = 3, \ldots, 6$, from (7) to (10). As an example the average detection delay for the $\alpha = 5\%$ level for the boundary function $b_5(t)$ for different breaks is displayed in tables 6, 7 and 8.

Boundary function $b_5(t)$									
$\tau = 2$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	53.33	43.40	37.23	77.61	63.15	49.58	91.34	72.67	56.20
0.2	79.23	63.41	54.45	112.58	95.82	81.91	126.86	112.28	95.50
0.3	112.44	92.12	76.31	144.79	137.08	122.53	158.95	156.00	148.96
0.4	133.96	118.80	99.73	150.59	158.25	151.10	157.33	172.92	172.75
$\tau = 4$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	57.89	46.73	38.50	91.25	71.51	57.13	116.31	90.09	69.96
0.2	94.31	72.12	59.49	149.92	118.26	98.78	186.03	152.03	127.12
0.3	160.65	109.40	86.93	222.47	183.72	152.57	265.43	240.25	210.58
0.4	225.64	167.07	121.51	272.32	258.26	214.56	293.90	309.68	282.91
$\tau = 6$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	53.52	41.29	32.36	96.13	71.78	56.47	124.22	94.77	72.26
0.2	98.55	69.55	55.80	167.16	124.39	104.44	217.05	169.95	138.28
0.3	190.16	111.91	85.96	282.65	204.74	165.30	339.06	276.86	231.03
0.4	315.36	192.13	126.20	380.04	315.26	239.72	419.28	396.05	320.28

Table 6: Average detection delay of the monitoring procedure for $T = 250$, $\alpha = 5\%$ and $\rho = 0.3$.

Table 6 shows the results for the case of an early break within the monitoring period. As one expects the detection is easier and therefore faster if the difference between d_0 and d_1 is large. Consequently the detection delay is rather small if the persistence changes from stationary, say $d_0 = 0.2$, long memory to non-stationary, say $d_1 = 0.8$, and even

[2] This has also been confirmed in unreported simulations.

faster if the process becomes a unit root process after the break. In fact, the detection delay for larger breaks is comparable with the short memory case (see table 3 in Leisch et al. (2000)). This is encouraging given the well known slow rate of convergence in long memory time series. Another result is that it is easier and faster to detect a change in persistence if the width of the monitoring window $[Th]$ is rather small. Detection delays for values of $h = 0.25$ and $h = 0.5$ are generally smaller compared to larger values of h. This is also in line with the findings of Leisch et al. (2000) for the short memory case. It is well known also in related areas of the structural change literature (see e.g Pesaran and Timmermann (2005) for results regarding forecasts under structural breaks) that smaller windows of data are usually better to detect and deal with structural change. The results for later breaks within the monitoring period are shown in tables 7 and 8. The general conclusions from above remain valid but the detection delay becomes even smaller if the breaks occurs later. This is also a similar behavior to the short memory case reported in Leisch et al. (2000).

Boundary function $b_5(t)$									
$\tau = 2$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	44.55	35.39	28.67	65.85	54.45	43.62	74.23	60.25	48.20
0.2	65.49	55.62	47.31	89.16	82.66	72.46	96.55	91.89	81.20
0.3	85.56	78.49	68.26	110.59	112.02	106.93	118.37	123.46	118.71
0.4	89.62	92.35	84.78	99.62	115.70	117.50	95.06	110.72	125.06
$\tau = 4$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	51.01	35.76	29.86	88.67	68.46	52.90	113.94	87.37	67.87
0.2	88.44	64.41	51.54	141.47	116.89	98.79	172.76	150.65	129.69
0.3	143.60	108.29	82.25	190.48	174.08	152.71	223.04	217.80	204.48
0.4	169.54	153.31	114.90	204.31	218.35	201.99	225.85	250.99	241.94
$\tau = 6$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	36.59	24.66	16.15	87.80	65.26	47.83	122.35	90.52	67.80
0.2	86.31	58.32	41.59	158.73	120.80	99.18	204.27	169.67	140.36
0.3	174.06	106.13	74.87	242.96	201.27	163.20	289.19	264.62	229.60
0.4	235.73	181.23	121.38	287.28	279.41	231.81	303.02	332.77	302.43

Table 7: Average detection delay of the monitoring procedure for $T = 250$, $\alpha = 5\%$ and $\rho = 0.5$.

Boundary function $b_5(t)$									
$\tau = 2$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	28.02	24.45	18.84	39.55	35.59	30.44	41.95	38.81	31.69
0.2	39.24	38.55	33.06	53.02	52.25	50.44	54.65	53.90	52.27
0.3	45.85	53.02	50.46	61.29	69.01	69.73	50.52	60.77	69.68
0.4	27.52	44.96	50.19	21.71	38.06	52.93	-77.72	-86.60	-61.95
$\tau = 4$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	38.81	24.41	16.72	69.65	58.49	44.95	86.56	79.32	63.26
0.2	66.51	51.73	40.02	101.49	101.71	92.10	121.38	117.65	112.30
0.3	90.24	82.78	69.57	124.50	133.20	131.38	146.83	156.16	159.54
0.4	77.71	100.90	92.53	96.81	138.45	144.58	93.48	140.91	166.83
$\tau = 6$	$h = 0.25$			$h = 0.5$			$h = 0.75$		
d_0	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$	$d_1 = 0.6$	$d_1 = 0.8$	$d_1 = 1$
0.1	12.08	4.92	-5.71	68.34	50.98	33.94	99.05	76.97	55.14
0.2	63.08	37.87	20.33	120.52	103.78	87.11	155.37	142.54	128.07
0.3	104.40	85.53	57.51	163.12	161.03	145.76	188.54	197.29	196.56
0.4	110.33	132.11	96.32	139.56	185.50	185.40	148.29	215.15	222.53

Table 8: Average detection delay of the monitoring procedure for $T = 250$, $\alpha = 5\%$ and $\rho = 0.7$.

4 Empirical Application

To illustrate the use of the monitoring approach we analyze monthly US price inflation series from Stock and Watson (2005).[3] In particular we consider the first difference of the logarithmic implied price deflator for durable goods. This series has also been under investigation from Cavaliere and Taylor (2008) who report a change in persistence from $I(0)$ to $I(1)$. However, they did not consider the possibility of fractional integration in the series although inflation related time series are likely to show long memory behavior (see e.g. Hassler and Wolters (1995)). The sample spans from 01/1959 to 12/2003. The series is depicted in figure 1.

[3]The data is available at Mark Watson's website at: http://www.princeton.edu/~mwatson/wp.html.

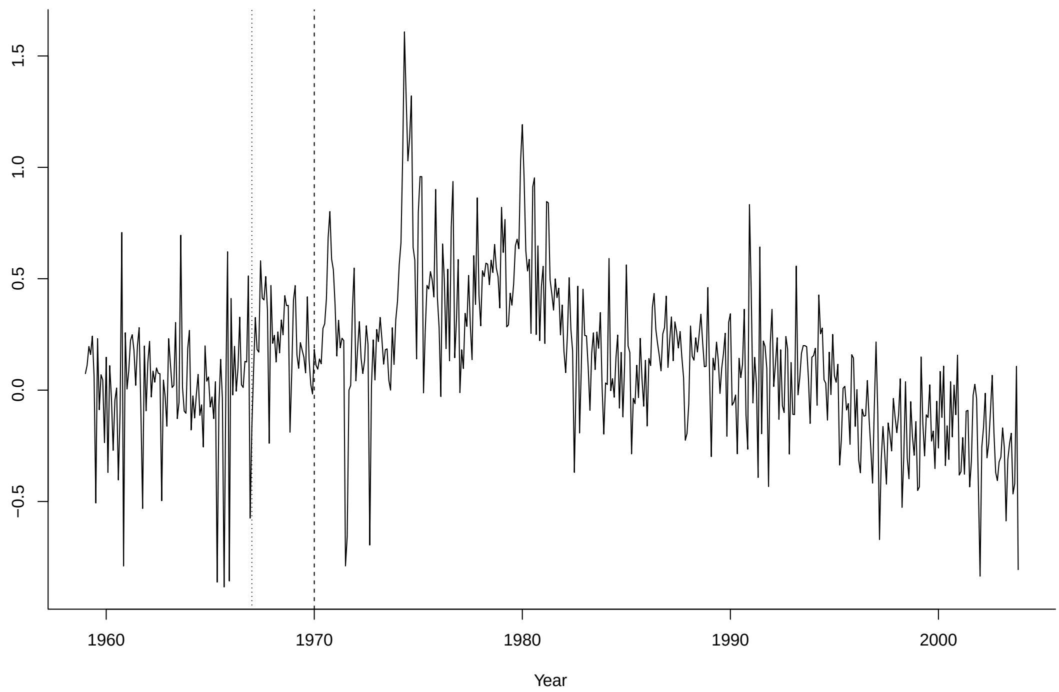

Figure 1: First difference of logarithmic price deflator for durable goods.

To determine the value of the long memory parameter we use log-periodogram regression as proposed by Geweke and Porter-Hudak (1983). The decision of how many frequencies should be used in the regression is a trade-off between reducing the bias and reducing the asymptotic variance. We use $T^{1/2}$ frequencies to deal with potential short memory components in the data (see e.g. Agiakloglou et al. (1993)). For the whole sample this yields an estimate of $\hat{d} = 0.61$. This value is highly significant as judged by its p-value which is $< 1e-03$.

To test whether a change in persistence can be detected in the data we apply the CUSUM of squares test for a change in persistence proposed by Sibbertsen and Kruse (2009) to the whole sample. This leads to a test statistic of $R = 0.0373$ which is significant at the $\alpha = 5\%$ level in favor of an increasing persistence. The estimated breakpoint is at $t^* = 107$ which is 11/1967 (the dotted line in figure 1).

To use the monitoring approach we split the sample in an in-sample part ranging from 01/1959 to 12/1965 and leave the rest as monitoring period. This yields a $\tau \approx 5$. The estimated d_0 within the in-sample period is $\hat{d}_0 = 0.23$.

For the application of the MOSUM test we use the boundary function $b_5(t)$ and set $h = 0.5$. The first time the sequence of test statistics exceeds the $\alpha = 1\%$ boundary function is at $t = 55$ in the monitoring period. This is equivalent to an estimated breakpoint at $t^* = 139$ which is 06/1970 (the dashed line in figure 1). The first time the sequence of test statistics exceeds the $\alpha = 5\%$ and $\alpha = 10\%$ boundary functions is only one period earlier.

The estimation of d_1 in the monitoring period yields $\hat{d}_1 = 0.68$. Thus we can confirm a change in persistence with high probability from stationary long memory to non-stationary long memory.
Notably the detection delay is rather short and we obtain a fast indication of the change in persistence from using the monitoring procedure.

5 Conclusion

Detecting a change in persistence as soon as possible is of paramount interest because structural change affects the subsequent analysis of the data heavily. The usual approach is to use one-shot tests to detect a change in persistence a posteriori. However, these tests cannot be applied sequentially because a correct null of no change would eventually be rejected with probability one. We propose a monitoring procedure based on moving sums that allows to detect a change in the long memory parameter of a long range dependent time series whenever new data arrives. By means of a Monte Carlo experiment we show good size properties and also study the detection delay when a change in persistence occurs. Depending on the width of the monitoring window and the difference between the pre- and post-break long memory parameter the detection is rather fast. Smaller monitoring windows generally prove more useful to detect a change in persistence early and also larger differences between the long memory parameters are detected faster.
In an empirical illustration of the method we are able to confirm a change in persistence from stationary to non-stationary long memory in an inflation time series.

A Appendix

A.1 Proof of theorem 1

First, let $k=[Tt]$ for each value in the monitoring period then write the test statistic as

$$\begin{aligned}
MS_{T,h,d} &= \max_{T+1\le k\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d}\left|\sum_{i=k-[Th]+1}^{k}\hat{e}_i-\frac{[Th]}{T}\sum_{i=1}^{T}\hat{e}_i\right| \\
&= \max_{T+1\le[Tt]\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d}\left|\sum_{i=1}^{[Tt]}\hat{e}_i-\frac{[Tt]}{T}\sum_{i=1}^{T}\hat{e}_i-\sum_{i=1}^{[Tt]-[Th]}\hat{e}_i+\frac{[Tt]-[Th]}{T}\sum_{i=1}^{T}\hat{e}_i\right| .
\end{aligned}$$

Then using the FCLT for fractionally integrated processes (see Sowell (1990) and Davidson and de Jong (2000)) and the continuous mapping theorem (CMT) we have

$$\begin{aligned}
MS_{T,h,d} &\Rightarrow \max_{T+1\le[Tt]\le[T\tau]} \sigma^{-1}\left|B(t,d)-tB(1,d)-B(t-h,d)+(t-h)B(1,d)\right| \\
&= \max_{T+1\le[Tt]\le[T\tau]} \sigma^{-1}\left|BB^0(t,d)-[B(t-h,d)-(t-h)B(1,d)]\right| \\
&= \max_{T+1\le[Tt]\le[T\tau]} \sigma^{-1}\left|BB^0(t,d)-BB^0(t-h,d)\right| ,
\end{aligned}$$

where $BB^0(t,d)$ denotes a fractional Brownian bridge.

To prove consistency we consider that at some point in the monitoring period, say k^*, the persistence changes from stationary long memory with $0<d_0<\frac{1}{2}$ to non-stationary long memory with $\frac{1}{2}<d_1<\frac{3}{2}$ and then split the test statistic into its stationary and non-stationary parts. We write the test statistic as

$$\begin{aligned}
MS_{T,h,d_0} &= \max_{T+1\le k\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=k-[Th]+1}^{k}\hat{e}_i-\frac{[Th]}{T}\sum_{i=1}^{T}\hat{e}_i\right| \\
&= \max_{T+1\le[rT]\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\underbrace{\sum_{i=1}^{[rT]}\hat{e}_i}_{I}-\underbrace{\sum_{i=1}^{[rT]-[hT]}\hat{e}_i}_{II}-\frac{[Th]}{T}\underbrace{\sum_{i=1}^{T}\hat{e}_i}_{III}\right| ,
\end{aligned}$$

where $k=[rT]$ for some $r>1$. Part III only contains $I(d_0)$ variables due to the noncontamination assumption.

We have to distinguish two cases:

(i) $k^*\le[rT]-[Th]$ $\Rightarrow$ in this case both I and II contain $I(d_1)$ variables

(ii) $[rT]-[Th]\le k^*\le[rT]$ $\Rightarrow$ in this case only I contains $I(d_1)$ variables.

Ad (i):

The case (i) is depicted in figure 2 where $[rT]$ is denoted by k_1 and $[rT]-[Th]$ is denoted by k_0. The gray shaded area is the monitoring window.

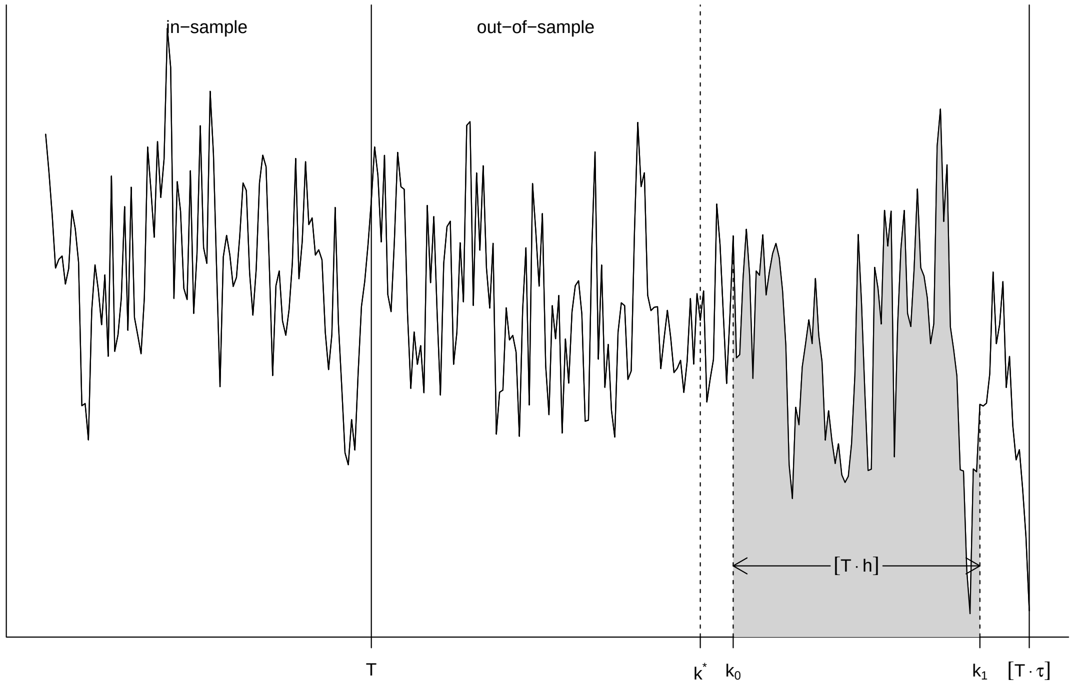

Figure 2: MOSUM case (i).

Write the test statistic as

$$
\begin{aligned}
MS_{T,h,d_0} &= \max_{T+1\le[rT]\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=1}^{k^*}\hat{e}_i + \sum_{i=k^*+1}^{[rT]}\hat{e}_i - \sum_{i=1}^{k^*}\hat{e}_i - \sum_{i=k*+1}^{[rT]-[hT]}\hat{e}_i - \frac{[hT]}{T}\sum_{i=1}^{T}\hat{e}_i\right| \\
&= \max_{T+1\le[rT]\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|-\frac{[hT]}{T}\sum_{i=1}^{T}\hat{e}_i\right| + \max_{T+1\le[rT]\le[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=[rT]-[hT]}^{[rT]}\hat{e}_i\right| .
\end{aligned}
$$

Now, the first part is $I(d_0)$ and is correctly standardized. Therefore, using the arguments from above it converges to $-hB(1,d_0)$ which is the standard deviation of the fractional Brownian motion. For the second part the standardization is obtained from d_0 but the variables are $I(d_1)$ and so the expression diverges and we obtain

$$MS_{T,h,d_0} = o_p(1) + O_p\left(T^{d_1-d_0}\right) . \tag{11}$$

Ad (ii):

The situation (ii) is depicted in figure 3.

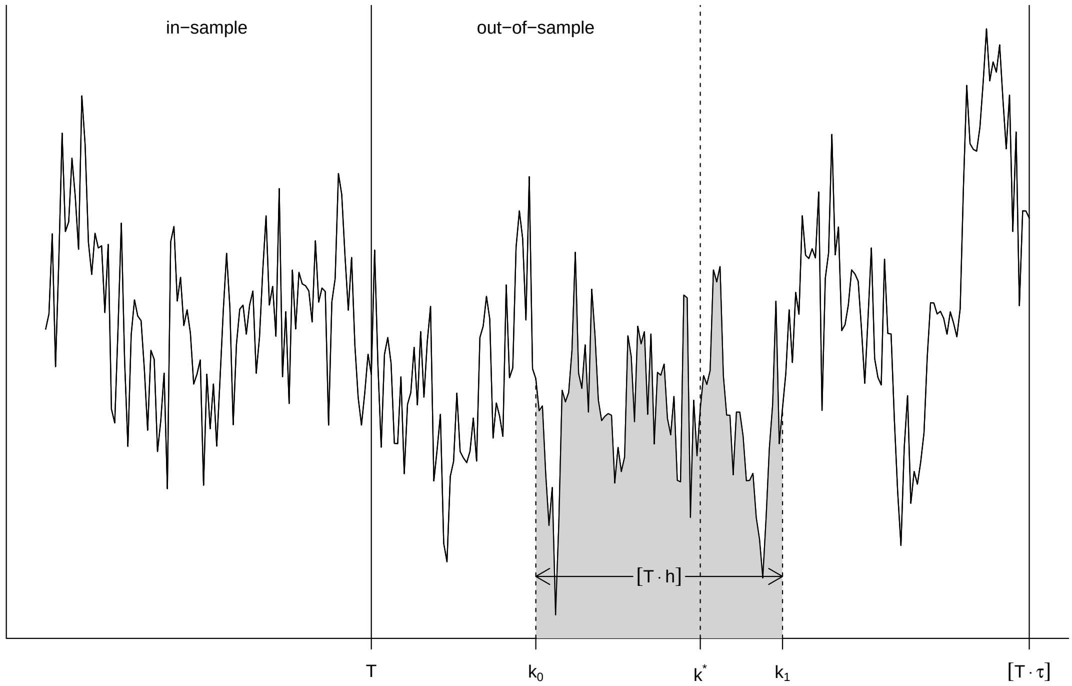

Figure 3: MOSUM case (ii).

Now only I contains $I(d_1)$ variables. Write the test statistic as

$$\begin{aligned}
MS_{T,h,d_0} &= \max_{T+1\leq[rT]\leq[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=1}^{[rT]}\hat{e}_i - \sum_{i=1}^{[rT]-[hT]}\hat{e}_i - \frac{[hT]}{T}\sum_{i=1}^{T}\hat{e}_i\right| \\
&= \max_{T+1\leq[rT]\leq[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=1}^{k^*}\hat{e}_i + \sum_{i=k^*+1}^{[rT]}\hat{e}_i - \sum_{i=1}^{[rT]-[hT]}\hat{e}_i - \frac{[hT]}{T}\sum_{i=1}^{T}\hat{e}_i\right| \\
&= \max_{T+1\leq[rT]\leq[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=k^*+1}^{[rT]}\hat{e}_i\right| + \max_{T+1\leq[rT]\leq[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\sum_{i=[rT]-[hT]}^{k^*}\hat{e}_i\right| \\
&\quad - \max_{T+1\leq[rT]\leq[T\tau]} \sigma^{-1}T^{-\frac{1}{2}-d_0}\left|\frac{[hT]}{T}\sum_{i=1}^{T}\hat{e}_i\right| .
\end{aligned}$$

With the arguments from case (i) we obtain

$$MS_{T,h,d_0} = O_p\left(T^{d_1-d_0}\right) + o_p(1) + o_p(1)\,, \tag{12}$$

where the second part of the above expression does not expand with T anymore and therefore vanishes as $T \to \infty$. □

A.2 Proof of theorem 2

Denote by $d_T := \sqrt{2T^{2d+1}\log\log(T)}$. By the reverse triangle inequality we have for some $r \in [0,1]$

$$d_T^{-1}\left|B(Tr,d) - rB(T,d) - (B(Tr',d) - r'B(T,d))\right| \leq d_T^{-1}\left|B(Tr,d) - B(Tr',d)\right| + d_T^{-1}\left|(r-r')B(T,d)\right| ,$$

for distinct values r and r'. Using the notation from Altissimo and Corradi (2003, p. 232) we write $S(r,\delta) = (r' : |r-r'| \leq \delta)$. Now, by the fact that (see Davidson (1994, p. 335 ff.))

$$\sup_{\theta\in\Theta}\sup_{\theta'\in S(\theta,\delta)}\left|f_n(\theta') - f_n(\theta)\right| \leq 2\sup_{\theta\in\Theta}|f_n(\theta)|$$

and the LIL for the fractional Brownian motion (see e.g. Taqqu (1977)) we have for the second part of the right side

$$\limsup_{T\to\infty}\sup_{r\in[0,1]}\sup_{r'\in S(r,\delta)} d_T^{-1}\left|(r-r')B(T,d)\right| \leq 2\delta\sigma ,$$

with σ the variance of the fractional Brownian Motion. As $\delta \to 0$ the whole part approaches zero which ensures the asymptotic uniform equicontinuity almost surely.
For the first part of the right hand side we have by self-similarity

$$\begin{aligned}
\limsup_{T\to\infty}\sup_{r\in[0,1]}\sup_{r'\in S(r,\delta)} d_T^{-1}\left|B(Tr) - B(Tr')\right| &= \limsup_{T\to\infty}\sup_{r\in[0,1]}\sup_{r'\in S(r,\delta)} d_T^{-1}\left|T^{d+1/2}B(r) - T^{d+1/2}B(r')\right| \\
&= \limsup_{T\to\infty}\sup_{r\in[0,1]}\sup_{r'\in S(r,\delta)} T^{d+1/2}d_T^{-1}\left|B(r) - B(r')\right| .
\end{aligned}$$

Now note that

$$d_T = \sqrt{2T^{2d+1}\log\log(T)} = \sqrt{T^{2d+1}}\sqrt{2\log\log(T)} = T^{d+1/2}\sqrt{2\log\log(T)} .$$

Therefore we obtain

$$\limsup_{T\to\infty}(2\log\log(T))^{-\frac{1}{2}}\sup_{r\in[0,1]}\sup_{r'\in S(r,\delta)}\left|B(r) - B(r')\right| .$$

Because $|B(r) - B(r')|$ is almost surely Hölder continuous of order strictly less than H (see Biagini et al. (2008, p. 11)) and $\limsup_{T\to\infty}(2\log\log(T))^{-\frac{1}{2}}$ tends to zero as $T \to \infty$ it follows that the above expression is almost surely asymptotically uniform equicontinuous. □

References

Agiakloglou, C., Newbold, P., and Wohar, M. (1993). Bias of an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, 14:235–246.

Altissimo, F. and Corradi, V. (2003). Strong rules for detecting the number of breaks in a time series. *Journal of Econometrics*, 117:207–244.

Andreou, E. and Ghysels, E. (2006). Monitoring disruptions in financial markets. *Journal of Econometrics*, 135:77–124.

Andrews, D. W. K. and Fair, R. C. (1988). Inference in nonlinear econometric models wiht structural change. *The Review of Economic Studies*, 55:615–639.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73:5–59.

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

Bauer, P. and Hackl, P. (1978). The use of MOSUMs for quality control. *Technometrics*, 20:431–436.

Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society. Series B*, 57:659–672.

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

Biagini, F., Hu, Y., Øksendahl, B., and Zhang, T. (2008). *Stochastic calculus for fractional brownian motion and applications*. Springer, New York.

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

Cavaliere, G. and Taylor, A. M. R. (2008). Testing for a change in persistence in the presence of non-stationary volatility. *Journal of Econometrics*, 147:84–98.

Chu, C., Hornik, K., and Kuan, C. (1995). MOSUM tests for parameter constancy. *Biometrika*, 82:603–617.

Chu, C., Stinchcombe, M., and White, H. (1996). Monitoring structural change. *Econometrica*, 64:1045–1065.

Clements, M. P. and Hendry, D. F. (1998). *Forecasting economic time series*. Cambridge University Press.

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, 27:46–78.

Davidson, J. (1994). *Stochastic limit theory*. Oxford University Press.

Davidson, J. and de Jong, R. (2000). The functional central limit theorem and weak convergence to stochastic integrals II: Fractionally integrated processes. *Econometric Theory*, 16:643–666.

Garcia, R. and Perron, P. (1996). An analysis of the real interest rate under regime shifts. *The Review of Economics and Statistics*, 78:111–125.

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

Ghysel, E., Guay, A., and Hall, A. (1997). Predictive tests for structural change with unknown breakpoint. *Journal of Econometrics*, 82:209–233.

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

Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in U.S. labor productivity. *The Journal of Economic Perspectives*, 15:117–128.

Hassler, U. and Nautz, D. (2008). On the persistence of the Eonia spread. *Economics Letters*, 10:184–187.

Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, 13:37–45.

Heinen, F., Sibbertsen, P., and Kruse, R. (2009). Forecasting long memory time series under a break in persistence. CREATES Research Paper 2009-53.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68:165–176.

Hsu, C. (2007). The MOSUM of squares test for monitoring variance changes. *Finance Research Letters*, 4:254–260.

Hurvich, C. M., Deo, R., and Brodsky, J. (1998). The mean-squared error of Geweke and Porter-Hudaks estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19:19–46.

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

Kim, J., Belaire Franch, J., and Badilli Amador, R. (2002). Corrigendum to "Detection of a change in persistence of a liner times series". *Journal of Econometrics*, 109:389–392.

Kuan, C. and Hornik, K. (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews*, 14:135–161.

Leisch, F., Hornik, K., and Kuan, C. (2000). Monitoring structural changes with the generalized fluctuation test. *Econometric Theory*, 16:835–854.

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

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

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

Pesaran, M. H. and Timmermann, A. (2005). Small sample properties of forecasts from autoregressive models under structural breaks. *Journal of Econometrics*, 129:183–217.

Ray, B. K. and Tsay, R. S. (2002). Bayesian methods for change-point detection in long-range dependent processes. *Journal of Time Series Analysis*, 23:687–705.

Robbins, H. (1970). Statistical methods related to the law of iterated logarithms. *The Annals of Mathematical Statistics*, 41:1397–1409.

Sibbertsen, P. and Kruse, R. (2009). Testing for a change in persistence under long-range dependencies. *Journal of Time Series Analysis*, 30:263–285.

Siegmund, D. (1985). *Sequential Analysis*. Springer, New York.

Sowell, F. (1990). The fractional unit root distribution. *Econometrica*, 58:495–505.

Stock, J. H. and Watson, M. W. (2005). Implications of dynamic factor analysis for VAR models. NBER working paper No. 11467.

Taqqu, M. (1977). Law of the iterated logarithm for sums of non-linear functions of gaussian variables that exhibit a long range dependence. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40:203–238.

Yamaguchi, K. (2011). Estimating a change point in the long memory parameter. *Journal of Time Series Analysis*, 32:304–314.

Zeileis, A., Leisch, F., Kleiber, C., and Hornik, K. (2005). Monitoring structural change in dynamic econometric models. *Journal of Applied Econometrics*, 20:99–121.

Zivot, E. and Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 10:251–270.