Detecting multiple breaks in long memory the case of U.S. inflation

Uwe Hassler · Barbara Meller

Abstract   Multiple structural change tests by Bai and Perron (Econometrica 66:47–78, 1998) are applied to the regression by Demetrescu et al. (Econ Theory 24:176–215, 2008) in order to detect breaks in the order of fractional integration. With this instrument we tackle time-varying inflation persistence as an important issue for monetary policy. We determine not only the location and significance of breaks in persistence, but also the number of breaks. Only one significant break in U.S. inflation persistence (measured by the long-memory parameter) is found to have taken place in 1973, while a second break in 1980 is not significant.

Keywords   Fractional integration · Break in persistence · Unknown break point · Inflation dynamics

JEL Classification   C22 · E31

1 Introduction

Inflation persistence is an important issue for economists and especially for central bankers. This is because the degree of inflation persistence influences the extent to

U. Hassler
Goethe-Universität Frankfurt, Frankfurt am Main, Germany
e-mail: hassler@wiwi.uni-frankfurt.de

B. Meller (✉)
Deutsche Bundesbank, Frankfurt am Main, Germany
e-mail: barbara.meller@bundesbank.de

which central banks can control inflation. If inflation persistence is high, a shock to the price level increases inflation for a long period. In the worst case scenario, inflation might even follow the path of a random walk, making it impossible for central banks to bring it under control. In the best case, inflation is integrated of order zero. This implies that it reverts back to its initial level soon after a shock has occurred.

Not only is the level of inflation important in economic analysis but also the question of whether and when it has changed. If the occurrence and/or timing of a break are not accounted for properly, then inflation forecasts and policy decisions might be misguided. Despite its importance, there is still no agreement on the significance and dating of past changes in inflation persistence in the U.S. and elsewhere.[1] The diverse findings could be due to the fact that many studies ignore the fractionally integrated nature of inflation. This may lead to misspecification and incorrect test results. The early results presented by Geweke and Porter-Hudak (1983) along with the international evidence of Hassler and Wolters (1995) and Baillie et al. (1996), have long since established that inflation exhibits long memory. In view of this evidence, Kumar and Okimoto (2007) argue that tests for a change in inflation persistence using unit root tests or autoregressive coefficients may lead to incorrect conclusions. Their study is the first to use long memory techniques to determine a change in inflation persistence. It applies a visual judgment of rolling window estimates and analyzes two exogenously split subsamples. We go beyond this approach and attempt not only to answer the question of whether there has been a change in U.S. inflation persistence but also to determine the *data-driven* timing and the number of breaks.

Our paper contributes to the existing literature by proposing and applying a new procedure for determining the timing and the significance of breaks in the degree of fractional integration. We are not aware of any other test allowing for multiple breaks in long memory at unknown points in time. The test builds on a modified version of the lag-augmented LM (Lagrange Multiplier) test proposed by Demetrescu et al. (2008) where dummy variables account for potential breaks. The F type test statistic is computed from a regression of differences under the null hypothesis. Therefore, no $I(d)$ series, $d \neq 0$, enter the test regression under the null, and the estimators converge at the conventional $\sqrt{T}$ rate, with T denoting the sample size. Consequently, we can compare the maximum of a sequence of F statistics to critical values by Bai and Perron (1998, 2003b), see also Andrews (1993) for the case of just one break. The test is able to detect a break in the long-memory parameter even relatively close to the boundaries of the sample because it does not rely on a separate estimation of the long-memory parameter before and after potential breaks. Further, a sequence of tests makes it possible to estimate the number of breaks.

Since Stock's (2001) comment on the innovative study of Cogley and Sargent (2001), his warning not to confuse a change in volatility with a change in persistence has been taken seriously. Fortunately, our test inherits the properties of the lag-augmented LM test developed by Demetrescu et al. (2008): Using Eicker–White standard errors renders the test robust to unconditional heteroskedasticity of a very general nature, see Kew and Harris (2009). In fact, the variance process is

---

[1] For a review on changing inflation persistence in EMU member states see Tillmann (2012) and Meller and Nautz (2012).

essentially unrestricted, thus allowing for time-varying volatility except for explosive and degenerate cases.

We apply our new tests to monthly U.S. inflation rates in the period 1966–2008. While there is strong evidence for a break in long memory in October 1973, a second potential break in March 1980 turns out to be insignificant. Prior to making the long-memory analysis, a significant mean shift found in 1981 has been subtracted from the data. In addition, we observe a considerable decline in volatility during the eighties.

The rest of the paper is organized as follows. In the section which follows, we will discuss the model of fractional integration with a break in the order of integration. Next, in section three, we obtain a new Chow-type test for multiple breaks assuming the break dates are known a priori. Experimental evidence is collected showing that the test works extremely well in finite samples even if the order of integration is misspecified under the null hypothesis of no break. The fourth section is devoted to the case where the break points are not known. We propose performing the test as a max-Chow test in line with Andrews (1993) when testing against just one break, and generalizing this approach for several breaks by adopting tests developed by Bai and Perron (1998). The finite-sample performance is studied through simulations. In Sect. 5, we turn to the analysis of monthly U.S. inflation rates, allowing for breaks in the mean as well as for breaks in the order of integration. Our concluding remarks are made in the final section, while mathematical proofs are contained in the Appendix.

2 Breaks in long memory

As a starting point, let us recall how long memory is defined and interpreted within a fractionally integrated framework. Under the null hypothesis of no break the observed time series $\{y_t\}$ $(t = 1, \ldots, T)$ is integrated of order d,

$$\Delta^d y_t = (1 - L)^d y_t = e_t \sim I(0), \tag{1}$$

where $\{e_t\}$ is a stationary and invertible short-memory process integrated of order zero, $I(0)$, and L denotes the conventional lag operator. We write the Wold representation of $\{e_t\}$ in terms of zero mean white noise innovations $\{\varepsilon_t\}$, say $e_t = \sum_{k=0}^{\infty} \gamma_k \varepsilon_{t-k}$. Then integration of order zero means that $\sum_{k=0}^{\infty} \gamma_k$ is finite and different from zero. Fractional differences are defined through the usual binomial expansion,

$$(1 - L)^d = \sum_{i=0}^{\infty} \pi_{i,d} L^i, \quad \pi_{0,d} = 1, \ \pi_{i,d} = \frac{i - 1 - d}{i} \pi_{i-1,d}, \quad i \geq 1.$$

Similarly, one may expand the inverse filter with coefficients $\{\psi_{i,d}\}$,

$$y_t = (1 - L)^{-d} e_t = \sum_{i=0}^{\infty} \psi_{i,d} e_{t-i},$$

which provides a well defined stationary process only for $d < 0.5$. The impulse response coefficients $\{c_i\}$ of $\{y_t\}$ can be obtained by convolution of $\psi_{i,d}$ and γ_k such that:

$$y_t = \sum_{i=0}^{\infty} c_i \varepsilon_{t-i}.$$

Hassler and Kokoszka (2010) provide a necessary and sufficient condition which $\{\gamma_k\}$ has to obey for the impulse response coefficients $\{c_i\}$ to decay hyperbolically in d. Under this rather weak condition it holds true for $d > 0$ that

$$c_i \sim c\, i^{d-1} \quad, \text{i.e.,} \quad \lim_{i \to \infty} \frac{c_i}{i^{d-1}} = c \neq 0, \qquad (2)$$

where the constant c is defined in Hassler and Kokoszka (2010, Proposition 2.1). For $d = 1$, past innovations ε_{t-i} have a permanent effect on y_t, while for $0.5 \leq d < 1$ we observe nonstationarity with transitory shocks,[2] $c_i \to 0$ as $i \to \infty$. Finally, for $0 < d < 0.5$ the impulse response coefficients $\{c_i\}$ die out fast enough to be square-summable resulting in a stationary process, though still dying out slowly enough that $\{c_i\}$ is not summable, which characterizes long memory. In view of (2), d is interpreted as the degree of persistence or the memory parameter measuring how slowly the effect of past shocks dies out.

As an alternative hypothesis to (1) we model m breaks constituting $m + 1$ regimes,

$$y_t = (1 - L)^{-d_j} e_t, \quad t = T_{j-1} + 1, \ldots, T_j \quad j = 1, \ldots, m + 1, \qquad (3)$$

with $T_0 = 0$ and $T_{m+1} = T$. The null hypothesis of no breaks becomes

$$H_0 : d_2 - d_1 = \cdots = d_{m+1} - d_m = 0.$$

In what follows we prefer the parameterization

$$d_j = d + \theta_{j-1}, \quad j = 1, \ldots, m + 1, \; \theta_0 = 0, \qquad (4)$$

such that θ_j denotes the shift relative to the first period occurring at the jth break. The null hypothesis of interest may now be recast as

$$H_0 : \theta_1 = \cdots = \theta_m = 0. \qquad (5)$$

If a sudden shift in d is considered as too extreme in practice, there still may be a "smooth transition",[3]

---

[2] Such a feature is sometimes called "mean-reversion" although Phillips and Xiao (1999) argue that this is a misnomer given the nonstationarity.

[3] The model in (6) introduces a nonlinearity in $\Delta^d y_t$ which is not present under the null in (1). Baillie and Kapetanios (2007) and Baillie and Kapetanios (2008) found evidence in favour of nonlinearity in addition to long memory in many economic and financial time series. Contrary to (6), however, they instead assume a smooth transition autoregression or a similar nonlinear $I(0)$ model for $\Delta^d y_t$. An investigation of their tests under breaks in memory is beyond the scope of the present paper.

$$
y_t = \begin{cases}
y_0 + \sum_{i=0}^{t-1} \psi_{i,d}\, e_{t-i}, & t = 1, \ldots, T_1 \\[2mm]
y_{T_1} + \sum_{i=0}^{t-1-T_1} \psi_{i,d+\theta_1}\, e_{t-i}, & t = T_1 + 1, \ldots, T_2 \\[1mm]
\vdots & \\[1mm]
y_{T_m} + \sum_{i=0}^{t-1-T_m} \psi_{i,d+\theta_m}\, e_{t-i}, & t = T_m + 1, \ldots, T
\end{cases}
, \tag{6}
$$

where $\psi_{i,d+\theta_j}$ are the coefficients from expanding $(1-L)^{-d-\theta_j}$. In (6), only realizations e_t after a break contribute to the slowly evolving long memory after T_j.

As in Bai and Perron (1998), we assume that the potential break points are determined by break fractions λ_j, i.e. $T_j = [\lambda_j T]$, where $[\cdot]$ denotes the integer part. In fact, treating the break points as unknown parameters, it makes sense to distinguish true break fractions λ_j^0 from those estimated from the data, $\widehat{\lambda}_j$. To reduce the notational burden we have ignored such a distinction in the exposition so far. As usual in this literature, we maintain the standard assumption that the true break points all grow with the sample size, such that each subsample contains an increasing number of observations.

Assumption 1 For the true break fractions it holds true that

$$
0 = \lambda_0^0 < \lambda_1^0 < \cdots < \lambda_m^0 < \lambda_{m+1}^0 = 1.
$$

There exists a considerable body of literature that deals with a break from an I(0) to an I(1) process (and vice versa), starting with tests pioneered by Kim (2000) and Busetti and Taylor (2004). Hassler and Scheithauer (2011) showed that those tests have power against fractional alternatives, too. If we wish to allow for $d \neq 0$ under the null, however, then d would have to be estimated first in order to apply such tests to differenced data. It is not clear how the preliminary estimation step would affect the subsequent test.

Some recent papers have proposed alternative procedures to detect breaks in long memory at an unknown time. Referencing the least-squares principle, Gil-Alana (2008) discusses a procedure allowing for breaks in the memory parameter and/or the mean and a linear time trend, but this technique does not allow to establish significance. Sibbertsen and Kruse (2009) discuss a CUSUM of squares-based test and find that the critical values depend on the unknown parameter d. This requires a preliminary consistent estimation $\widehat{d}$ under H_0; such an estimate can be very volatile in smaller samples resulting in unreliable subsequent inference. Martins and Rodrigues (2012) propose a related procedure relying on a recursive forward and backward estimation where critical values depend again on the unknown parameter d. Further, Ray and Tsay (2002) adopt a Bayesian perspective and apply Markov Chain Monte Carlo methods to estimate the posterior probability and size of a change in the order of integration. Finally, Beran and Terrin (1996) suggest using non-overlapping subsamples to compute (approximate) maximum likelihood [ML] estimates of d, $\widehat{d}_1$, $t = 1, \ldots, T_1$, and $\widehat{d}_2$, $t = T_1 + 1, \ldots, T$, where T_1 is varied systematically. The test statistic builds

on the maximum difference $|\widehat{d}_2 - \widehat{d}_1|$. The limiting distribution established by Beran and Terrin (1999) coincides upon squaring with the one given by Andrews (1993) as supremum of so-called tied-down Bessel processes. It was derived under the sufficient assumption of $\sqrt{T}$-consistent estimators, see Andrews (1993, Theorem 1). Consequently, Beran and Terrin (1996) work with a parametric approximation to ML requiring a fully specified model for the $I(0)$ component $\{e_t\}$, see also Yamaguchi (2011), likewise working with an approximation to ML. The asymptotic theory does not seem to hold for semiparametric estimators converging with a slower rate than $\sqrt{T}$. This is one further motivation for our proposal, since in the regression framework by Demetrescu et al. (2008) $\sqrt{T}$-consistency is maintained, see also Proposition 2 below. The major advantage, however, of the regression approach is that it extends naturally to multiple breaks along the lines of Bai and Perron (1998).

3 Tests with known break points

In some cases, economists have an idea of the timing of a potential break point in persistence or wish to know the impact of a certain event on persistence. In the context of this paper, the inauguration of a new central bank governor might be an event that induces a break in inflation persistence. Alternatively, economists might be interested in the impact of a new inflation target or a new monetary policy regime on inflation persistence. Therefore, the case of known break fractions is an interesting starting point for which we will first derive a test statistic from the Lagrange Multiplier [LM] principle under simplifying assumptions before then turning to extensions that are relevant in practice.

3.1 Under iid assumptions

Working with finite samples of size T the theoretical difference operator from (1) has to be adjusted. Given a finite past starting value with the first observation y_1, the infinite expansion is truncated in practice. We call the truncated differences $\Delta_{t,y}^d$ instead of $\Delta^d y_t$, and denote them by x_t for brevity,

$$x_t = \Delta_{t,y}^d = \sum_{j=0}^{t-1} \pi_{j,d} y_{t-j}, \ t = 1, \ldots, T. \tag{7}$$

This amounts to assuming that past values of $\{y_t\}$ are zero for $t \leq 0$. Such processes are also classified as "type II" contrasting the more conventional "type I" processes; see e.g., Robinson (2005) for a discussion. To derive an LM test we will further assume absence of short memory.

Assumption 2 Let $\{e_t\} = \{\varepsilon_t\}$, $t \in \mathbb{Z}$, from (3) be an iid series with mean 0 and variance σ^2. The starting values are set equal to zero, $y_t = 0$ for $t \leq 0$.

To set up the score function in the Appendix we have to assume a Gaussian pseudo-log-likelihood function, although Gaussianity is not required for the limiting distribution below.

Proposition 1 *Under (3) with (4), (5), and Assumption 2, the LM statistic becomes—with true break points—$T_j = [\lambda_j^0 T]$*

$$LM = \frac{6}{\sigma^4 \pi^2} \sum_{j=1}^{m} \frac{\left(\sum_{t=T_j+1}^{T_{j+1}} x_t x_{t-1}^* \right)^2}{T_{j+1} - T_j}, \tag{8}$$

where

$$x_{t-1}^* = \sum_{i=1}^{t-1} i^{-1} x_{t-i} \tag{9}$$

with $\{x_t\}$ from (7).

Proof See Appendix. □

Note that LM with $\{x_t\}$ from (7) is computed under the null of no break, i.e. (5) and (4), which means that the differencing parameters are equal to d in all subsamples. The summation in LM starts with the second sample after T_1, but the information of the first sample is contained in $\{x_{t-1}^*\}$. Along the lines of Breitung and Hassler (2002, Theorem 1), LM can be approximated by an $F-$ statistic testing for $\psi_1 = \cdots = \psi_m = 0$ in the following regression estimated by ordinary least squares [OLS],

$$x_t = \sum_{j=1}^{m} \widehat{\psi}_j \, x_{t-1}^* \, D_t(\lambda_j^0) + \widehat{\varepsilon}_t, \quad t = 2, \dots, T, \tag{10}$$

with the step dummy variables ($j = 1, \dots, m$)

$$D_t(\lambda_j^0) = \begin{cases} 1, & t = [\lambda_j^0 T] + 1, \dots, [\lambda_{j+1}^0 T] \\ 0, & \text{else} \end{cases}. \tag{11}$$

For the usual F statistic it is straightforward to obtain (with $SSR = \sum_{t=2}^{T} \widehat{\varepsilon}_t^2$)

$$\frac{T - m}{m \, SSR} \sum_{j=1}^{m} \frac{\left(\sum_{t=T_j+1}^{T_{j+1}} x_t x_{t-1}^* \right)^2}{\sum_{t=T_j+1}^{T_{j+1}} \left(x_{t-1}^* \right)^2} = \frac{LM}{m} + o_p(1),$$

since $(T_{j+1} - T_j)^{-1} \sum_{t=T_j+1}^{T_{j+1}} \left(x_{t-1}^* \right)^2$ converges to $\sigma^2 \pi^2 / 6$ under Assumption 1.

Breitung and Hassler (2002) consider testing for the parameter value d (assuming a priori that there is no break) within

$$x_t = \widehat{\phi} x_{t-1}^* + \widehat{\varepsilon}_t, \quad t = 2, \ldots, T. \tag{12}$$

We now merge regressions (10) and (12) (and will argue in Sect. 3.3 that this robustifies against a misspecification of d):

$$x_t = \widehat{\phi} x_{t-1}^* + \sum_{j=1}^{m} \widehat{\psi}_j \, x_{t-1}^* \, D_t(\lambda_j^0) + \widehat{\varepsilon}_t, \quad t = 2, \ldots, T. \tag{13}$$

A break in fractional integration is indicated by means of the usual F statistic $F(\lambda_1^0, \ldots, \lambda_m^0)$ from (13) testing for the null

$$H_0 : \psi_1 = \cdots = \psi_m = 0. \tag{14}$$

The following result can be established, where "$\xrightarrow{d}$" stands for the convergence in distribution.

Proposition 2 *Under the assumptions of Proposition 1 and Assumption 1 it follows for the estimators from (13) that*

$$\sqrt{T} \, (\widehat{\phi}, \widehat{\psi}_1, \ldots \, \widehat{\psi}_m)' \; \xrightarrow{d} \; \mathcal{N}_{m+1} (0, \Sigma)$$

as $T \to \infty$, where Σ has full rank and is given in the Appendix. Hence,

$$m \, F(\lambda_1^0, \ldots, \lambda_m^0) \xrightarrow{d} \chi^2(m),$$

where $\chi^2(m)$ denotes a chi-squared distribution with m degrees of freedom.

Proof See Appendix.   □

Remark 1 In practice, the variables entering (13) will have a mean different from zero that has to be accounted for. Deterministic components have to be extracted prior to the regression, see e.g. Robinson (1994), such that $\{x_t\}$ can be considered as a zero mean variable.

3.2 Extensions

Assumption 2 is too restrictive for practical purposes and can be relaxed considerably. We indicate generalizations without going into technical details and omit formal proof, as our test statistic is related to statistics handled in the papers referenced below. A valid set of conditions replacing Assumption 2 is now adopted from Hassler et al. (2009).

Assumption 3 Let $\{e_t\}$ from (3) be a stable autoregressive process of order p,

$$e_t = \sum_{i=0}^{p} a_i e_{t-i} + \varepsilon_t,$$

driven by a strictly stationary and ergodic martingale difference series $\{\varepsilon_t\}$ with variance σ^2 satisfying an eight-order cumulant condition.

Let us briefly comment on generalizations going beyond the previous section (Assumption 2).

First, Assumption 3 relaxes the assumption of independence and instead assumes lack of correlation, maintaining that the innovations form a martingale difference series. In case of conditional homoskedasticity, $E(\varepsilon_t^2|\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = \sigma^2$, the asymptotic results of the previous section will not change, see, for example, Robinson (1991). In case of conditional heteroskedasticty, however, it is necessary to employ Eicker–White standard errors as advocated by Demetrescu et al. (2008). With such robustified standard errors the limiting distribution remains unchanged. More generally, this even holds true for unconditional heteroskedasticty of very general form, $E(\varepsilon_t^2) = \sigma_t^2$, where the variance process allows for smooth shifts as well as sudden breaks, see Kew and Harris (2009).

Second, upon fractional differencing one often observes additional short memory correlation in $\{e_t\}$. To account for autocorrelation, we follow Demetrescu et al. (2008) and augment the test regression with lagged endogenous variables,

$$x_t = \widehat{\phi} x_{t-1}^* + \sum_{j=1}^{m} \widehat{\psi}_j \, x_{t-1}^* \, D_t(\lambda_j^0) + \sum_{i=1}^{p} \widehat{a}_i \, x_{t-i} + \widehat{\varepsilon}_t. \tag{15}$$

In fact, Demetrescu et al. (2008) allow for more general processes $\{e_t\}$ than in Assumption 3. Their assumptions accommodate many short-memory AR (∞) processes that can be approximated with growing p. Since the regressors in (15) are not orthogonal, Demetrescu et al. (2008) advise against data-driven lag-length selection, as the model selection step affects subsequent inference about ψ_j even asymptotically, see, for example, Leeb and Pötscher (2005). Instead, they advocate choosing the lag length p in (15) by deterministically following the rule of thumb

$$p = \left[4(T/100)^{1/4} \right], \tag{16}$$

which was originally proposed by Schwert (1989). Although it lacks optimality properties it is widely used in applied econometrics. Demetrescu et al. (2011) collected further experimental support for its usefulness in practice in that it balances the trade-off between power and control of size under H_0.

Third, the starting value condition in Assumption 2 is not crucial. Note that, the sequence of regressors $\{x_{t-1}^*\}$ is only asymptotically stationary. Without zero starting values, the stationary, non-observable counterpart is $x_{t-1}^{**} = \sum_{j=1}^{\infty} j^{-1} \Delta^d y_{t-j}$ with y_t being from (3) under H_0. The difference between x_{t-1}^{**} and x_{t-1}^* becomes negligible

with growing sample size, as already stressed by Demetrescu et al. (2008) and more recently by Hassler et al. (2009).

3.3 Misspecification of d

When testing for (5), Propositions 1 and 2 assume the true d to be known a priori, which will rarely be the case in practice. Often, practitioners will estimate the unknown differencing parameter before testing for a break, which will result in fractional misspecification when computing the differences. Therefore, we now consider the model

$$\Delta^{d+\delta} y_t = e_t, \quad t = 1, \ldots, T, \tag{17}$$

where $\delta \neq 0$ is the degree of misspecification. Consequently, the differences $\{x_t\}$ from (7) building on Δ^d are not $I(0)$ but rather $I(\delta)$, and hence serially correlated and therefore correlated with $\{x_{t-1}^*\}$. Hence, it is easy to show for $\widehat{\psi}_j$ from (10) that $\widehat{\psi}_j \nrightarrow 0$, and that the LM statistic diverges as T increases. To compensate for this effect we proposed combining regression (10) with the original proposal of Breitung and Hassler (2002). The latter test building on (12) is consistent, and a violation of the specified order of integration ($\delta \neq 0$) will be captured by $\widehat{\phi} \rightarrow E(x_t x_{t-1}^*)$. This motivates the regression (13) instead of (10). Admittedly, under the more realistic null model (17) instead of (1) the asymptotic distribution of the estimators from (13) is not obvious. Although **local** power results (for $\delta = c/\sqrt{T}$) are available from Tanaka (1999, Theorem 3.1) or Demetrescu et al. (2008, Proposition 3), it is not clear how they generalize for a fixed δ. Still, we have the conjecture in Remark 2 that the approximation in Proposition 2 is valid a guideline under $\delta \neq 0$, too.

Remark 2 Let us assume the null model (17) with $\delta \neq 0$, while the test regression is computed with $\{x_t\}$ from (7) relying on Δ^d. We then expect that estimates $\widehat{\phi}$ significantly different from 0 in (13) or (15) will account for (at least moderate) misspecification $\delta \neq 0$, such that $\chi^2(m)$ provides a valuable approximation for the multiple of the F statistic under the null of no break in fractional integration.

To back Remark 2, we report results from a computer exercise for the case $m = 1$, which corresponds to a classical Chow test applied to the regression (13) or the lag-augmented version (15). We simulate time series with $T = 500$ observations and test for a break assuming as break fraction $\lambda_1 = 0.5$. The data is simulated with standard normal iid innovations $e_t = \varepsilon_t$ entering (17) with $d = 0$ and no break, such that the observables are integrated of order δ. The parameter δ measures the degree of misspecification. All experiments rely on 1,000 replications. We computed the size (at nominal 1, 5, and 10 % level) of the F test $F(0.5)$ from Proposition 2, i.e., from regression (13) without lags. Figure 1 shows experimental sizes for the range of $-0.4 \leq \delta \leq 0.4$ and $\delta = \pm 1$. They tend to be below the nominal ones except for $\delta = 1$ where we observe size distortions which would be unacceptable in practice. The size properties distinctly improve when working with the lag-augmented regression. This is not surprising since the lags capture some of the serial correlation of $\{x_t\}$
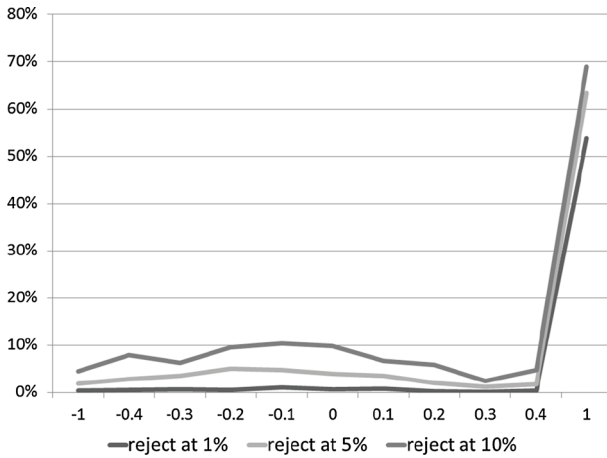
Fig. 1 Rejection rates from (13) plotted against δ from (17)
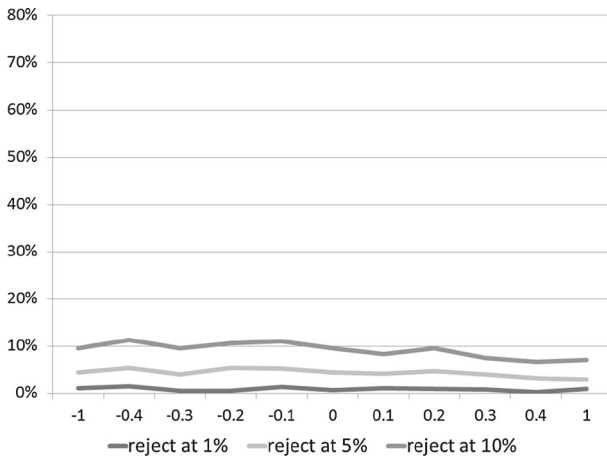

Fig. 2 Rejection rates from (15) plotted against δ from (17)

stemming from misspecification. Figure 2 displays the rejection rates of $F(0.5)$ from (15) when compared with quantiles from $\chi^2(1)$. Even for a misspecification as strong as $\delta = 1$ the size distortion is negligible. Hence, Fig. 2 soundly supports our conjecture in Remark 2 that a (moderate) misspecification due to the estimation of d in practice will leave the test valid.

4 Tests with unknown break points

Let us now turn to the interesting situation of where the timing of potential breaks in long memory is not known a priori. First, we adopt the tools by Bai and Perron (1998) to determine the number of breaks and to test for their significance. We then investigate their finite sample behavior in our context through Monte Carlo experimentation.

4.1 Implementation

We stick to the regression Eq. (15), only that the break fractions are now not known but varied over the sample. To underline this difference, we write

$$x_t = \widehat{\phi} x_{t-1}^* + \sum_{j=1}^{m} \widehat{\psi}_j \, x_{t-1}^* \, D_t(\lambda_j) + \sum_{i=1}^{p} \widehat{a}_i \, x_{t-i} + \widehat{\varepsilon}_t, \tag{18}$$

where the step dummies $D_t(\lambda_j)$ are defined as in (11) but with λ_j and hence $T_j = [\lambda_j T]$ varying. Under Assumption 3 all variables are (asymptotically) stationary, and the stage is set to perform a multiple change analysis along the lines of Bai and Perron (1998).

On top of the model and Assumption 1 concerning the true break fractions, we now assume that in the empirical aplication each sample segment has a minimal length determined by a trimming parameter $\epsilon > 0$:

$$\frac{T_j - T_{j-1}}{T} \geq \epsilon, \quad j = 1, \ldots, m+1.$$

The limiting distributions depend on the trimming, and Bai and Perron (1998) provide critical values for $\epsilon = 0.05$. Bai and Perron (2003a), however, recommend the usage of $\epsilon = 0.15$ in order to have better size properties in finite samples. For the rest of the paper we will work with $\epsilon = 0.15$ relying on corresponding critical values from response surface regressions by Bai and Perron (2003b).[4] F statistics $F(\lambda_1, \ldots, \lambda_m)$ testing for $\psi_1 = \cdots = \psi_m = 0$ from (18) are computed for all possible break points subject to

$$\Lambda_\epsilon = \{(\lambda_1, \ldots, \lambda_m) : |\lambda_j - \lambda_{j-1}| \geq \epsilon, \; j = 1, \ldots, m+1\}, \quad \lambda_0 = 0, \; \lambda_{m+1} = 1.$$

The maximum across all F statistics is called $sup F(m)$,

$$sup F(m) = \max_{\Lambda_\epsilon} \left(F(\lambda_1, \ldots, \lambda_m) \right).$$

It can easily be determined by a grid search for moderate sample sizes and small m. For large values of m, Bai and Perron (2003a) recommend the principle of dynamic programming. Critical values are available up to $m = 9$. For $m = 1$, this corresponds to a max-Chow test in line with Andrews (1993). The candidates for breaks are the arguments maximizing $sup F(m)$ (or, as an equivalent, minimizing the sum of squared residuals from (18)):

$$(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_m) = \arg\max_{\Lambda_\epsilon} \left(F(\lambda_1, \ldots, \lambda_m) \right).$$

---

[4] The critical values are available from an unpublished appendix to Bai and Perron (2003b) posted on the homepage of Pierre Perron.

In many cases, we do not want to specify a specific number (m) of potential breaks a priori. We would prefer to determine m from the data. To this end, Bai and Perron (1998) suggest a so-called double maximum test which we do not investigate here. Instead, we adopt their third proposal to test for the null hypothesis of ℓ breaks versus the alternative of $\ell+1$ changes building on a test statistic $sup F(\ell+1|\ell)$. To determine the number of breaks, Bai and Perron (2003a) advocate a sequence of tests, 0 vs. 1, 1 vs. 2, and so on; if $sup F(\ell+1|\ell)$ is not significant for $\ell \geq m$, then the number of breaks is determined as $m = \ell$. Obviously, $sup F(1|0) = sup F(1)$. In general, the statistic $sup F(\ell+1|\ell)$ is computed in the following way: determine the break points assuming ℓ breaks, $(\widehat{\lambda}_1, \dots, \widehat{\lambda}_\ell)$. For each of the $\ell + 1$ segments, determine the F statistic testing for $m = 1$ break at unknown time in segment j, say $sup F_j(1)$. If the overall maximal value, $\max_{j=1,\dots,\ell+1} sup F_j(1)$ is sufficiently large, then the null of ℓ breaks is rejected in favor of $\ell + 1$ breaks. The critical values are again available from Bai and Perron (1998, 2003b).

4.2 Monte Carlo evidence

For this section, we simulated time series with $T = 500$ observations, based on standard normal iid innovations $\{\varepsilon_t\}$. The true data generating process is from (6). We computed the size (at nominal 1, 5, and 10 % level) and power. All the rejection frequencies rely on 1,000 replications. In (6) we choose $d = 0$ without loss of generality and vary $|\theta|$ within the experiments between 0 and 1.

In a preliminary analysis we investigate the effect of a misspecified order of integration when testing for an unknown break point, analogously to the experiment leading to Fig. 2 in Sect. 3.3. For this purpose we evaluated with $m = 1$ the maximum $sup F(1)$ from (18) with with $p = 5$ lags under (17). The resulting empirical sizes resemble the ones in Figure 2 and are presented in Fig. 3. Hence, the issue of a misspecified order of integration does not seem to be of major concern, which is of course good news for applied purposes.

For all further simulations the true order of integration is not assumed to be known but estimated, which parallels the situation in real life. We use the so-called exact local Whittle [ELW] estimator proposed by Shimotsu and Phillips (2005). With the estimated $\widehat{d}$, the differences $\{x_t\}$ are constructed. F statistics are from regression (18) with $p = 5$ (according to (16)).

4.2.1 Case of one break

First, we focus on the situation where the data-generating process [DGP] has $m = 1$ change. Table 1 shows the empirical size and power of the $sup F(1)$ test, the mean of the estimated break fractions, their standard deviation and their root mean squared error for different values of $\theta_1 = \theta$ in (3). Figure 4 visualizes the power and the size of the test for $|\theta_1| = |\theta| \leq 0.4$ and $|\theta| = 1$. The unknown break data is in the middle of the sample, $\lambda_1^0 = 0.5$.

The simulation results in Table 1 correspond to expectations. The larger the difference in the order of integration before and after the break, the easier the break is
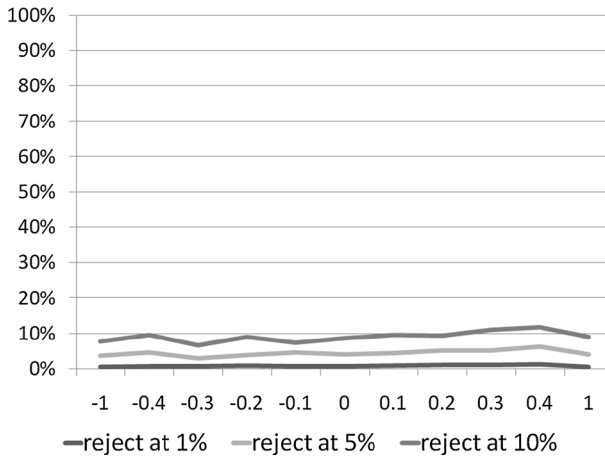
Fig. 3 Rejection rates from (18) plotted against δ from (17)

Table 1 Rejection of the null hypothesis of $supF(1)$ for different values of θ

θ	1 % CV (%)	5 % CV (%)	10 % CV (%)	$\overline{\hat{\lambda}_1}$	$\sigma(\hat{\lambda}_1)$	RMSE($\hat{\lambda}_1$)
-1	100.0	100.0	100.0	0.50	0.02	0.02
-0.4	97.0	99.4	99.7	0.49	0.07	0.07
-0.3	73.9	91.0	95.4	0.48	0.10	0.10
-0.2	29.0	53.5	67.2	0.48	0.15	0.15
-0.1	5.1	16.0	25.9	0.49	0.21	0.21
-0	0.6	4.5	8.7	0.49	0.24	0.24
-0.1	4.9	14.5	22.8	0.51	0.21	0.21
-0.2	25.5	48.4	60.9	0.51	0.15	0.16
-0.3	66.8	87.3	93.8	0.51	0.10	0.10
-0.4	94.6	98.6	99.3	0.51	0.06	0.06
-1	100.0	100.0	100.0	0.50	0.02	0.02

Notes The table shows, for different values of θ, how often the null hypothesis was rejected in the Monte Carlo study using the 1, 5, and 10 % critical values. The table also reports the mean of the estimated break fraction, its standard deviation and its root mean squared error. Simulation is conducted under the basic set-up: $T = 500$, $\lambda_1^0 = 0.5$, $\{\varepsilon_t\}$ white noise

detected and correctly allocated. In other words, the larger θ is in absolute terms, the higher the rejection rate and the smaller the RMSE($\hat{\lambda}_1$). Overall, the performance of our test in a finite sample is satisfactory. The size of the test is good: close to 1, 5, 10 % at the corresponding significance levels. The power is extremely high if the difference in the long-memory parameter before and after the break is greater than 0.3 in absolute terms. Even if the difference is only ±0.2, the power is still high. Figure 4 depicts the symmetry of the rejection rates with respect to θ around zero.

Next, we investigate the performance of the $supF(1)$ test in the light of a number of variations in the simulation set-up. In the left-hand graph of Fig. 5, the 5 %-rejection
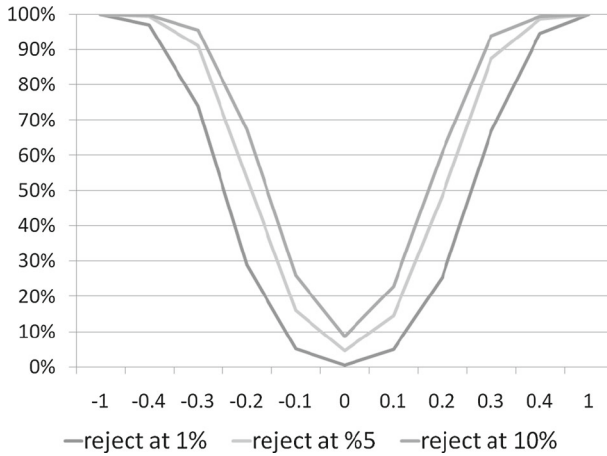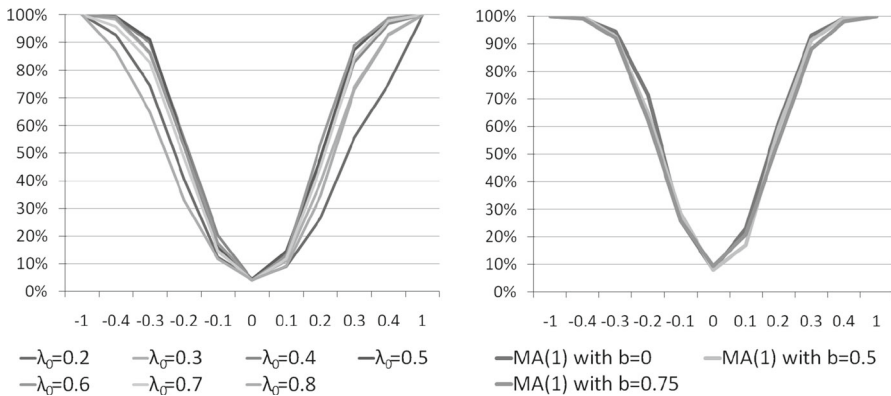
Fig. 4 $\sup F(1)$ plotted against θ, $m = 1$

Fig. 5 $\sup F(1)$: different break fractions and MA(1), nominal level of 5 %

rates are plotted against θ for different values of the true unknown break fraction: $\lambda_1^0 \in \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$. It is remarkable how well breaks are detected where there are only 150 observations before or after the break if $\theta > 0.2$, that is for $\lambda_1^0 = 0.2$ and $\lambda_1^0 = 1 - 0.2$. Where $\theta \leq 0.2$, and $\lambda_1^0 = 0.2$ or $\lambda_1^0 = 1 - 0.2$, the power is low. For all other cases, the power is high and the RMSE($\hat{\lambda}_1$) (not reported here) are comparable to those reported in Table 1.[5]

The right-hand graph in Fig. 5 contains the 5 %-rejection rates plotted against θ for three different moving average parameters. To allow for the short memory of the time series $\{e_t\}$ in (3), we consider an MA(1) process,

$$e_t = \varepsilon_t + b\,\varepsilon_{t-1}.$$

---

[5] Tables containing corresponding information as reported in Table 1 are available for all variations to the simulation set-up reported in this subsection.
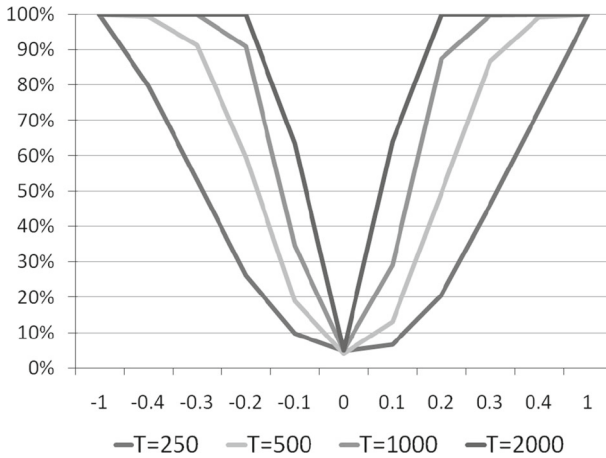
Fig. 6 $supF(1)$: different sample sizes, nominal level of 5 %

The MA(1) coefficient b takes on the values 0.00 (white noise), 0.50 and 0.75. Due to the lagged variables included in regression (18), the size and power of the $supF(1)$ are hardly affected, see Fig. 5.

Figure 6 shows the power of the test for different sample sizes, $T \in \{250, 500, 1000, 2000\}$. Unsurprisingly, the power decreases as the sample size decreases. For $T = 250$, the test is only of limited use but if T is greater, the test has good and even excellent power properties.

Next, we present a number of rejection frequencies of $supF(2)$ testing against two breaks where the true DGP only has one change. The results from Fig. 7 can be compared with $supF(1)$ from Fig. 4. In particular, at the 10 % level we observe that $supF(2)$ is mildly conservative under the null hypothesis. Consequently, it displays less power than $supF(1)$, which is not surprising since $supF(1)$ specifies the number of (potential) breaks correctly. Further, we present results for sequential testing under one break in Fig. 8. The left-hand graph contains rejection frequencies for $F(1|0)$, which coincide of course with $F(1)$ from Fig. 4. The right-hand graph in Fig. 8 shows the empirical sizes of $supF(2|1)$ at conventional levels. Given one break, $supF(2|1)$ tends to be mildly conservative; only in case of $\theta = 1$, are the experimental sizes above the nominal ones.

We briefly summarize the findings for $m = 1$. The power of the $supF(1)$ test depends especially on the difference in the order of integration before and after the break. If the difference is larger than 0.3, the power is very good. Furthermore, the power is almost unaffected by variations in the true break fraction or the value of the moving average coefficient in the case of MA(1) short memory. The power of the test is good for samples with at least 500 observations. Its size properties are quite satisfactory throughout all the simulation set-ups in that the power does not come at the price of a too liberal test.

Testing against two breaks we observe that $supF(2)$ and $supF(2|1)$ are both mildly conservative in that the experimental size tends to be smaller than the nominal one.
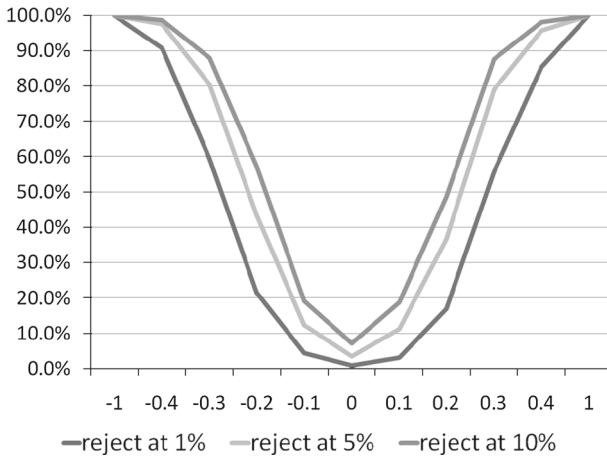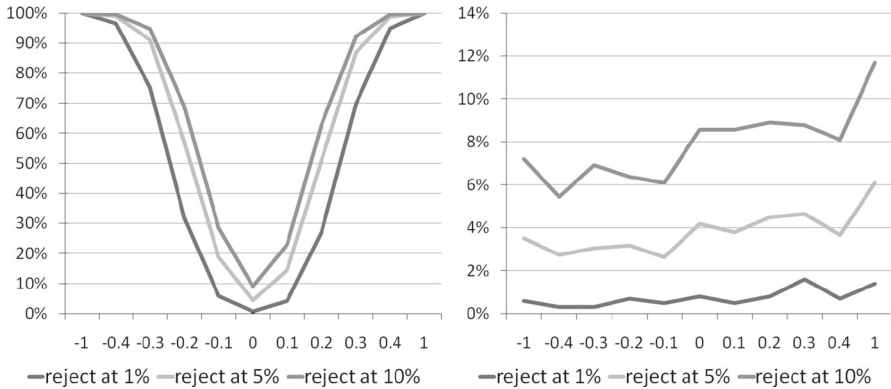
Fig. 7 $supF(2)$ plotted against θ, $m = 1$

Fig. 8 $supF(1|0)$ and $supF(2|1)$ plotted against θ, $m = 1$

4.2.2 Case of two breaks

The only difference to the previous experiments is that the DGP in this section has two breaks ($m = 2$). The true break fractions are $\lambda_1^0 = 1/3$ and $\lambda_2^0 = 2/3$. We consider the following scheme of breaks

$$d_1 = 0, \ d_2 = \theta, \ d_3 = 0,$$

which means in the notation of (4): $\theta_1 = \theta$ and $\theta_2 = 0$.

The size and power results of $supF(1)$ are given in the left-hand graph of Fig. 9. Clearly, the power curve is not as steep as in Fig. 4 because in the present DGP the second change returns to the original level.
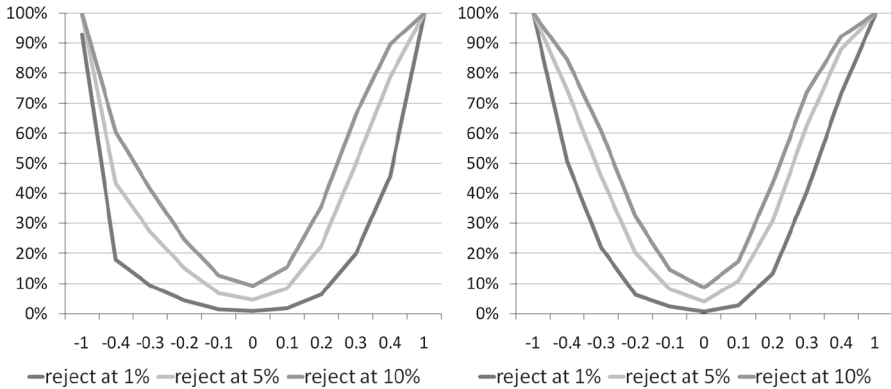
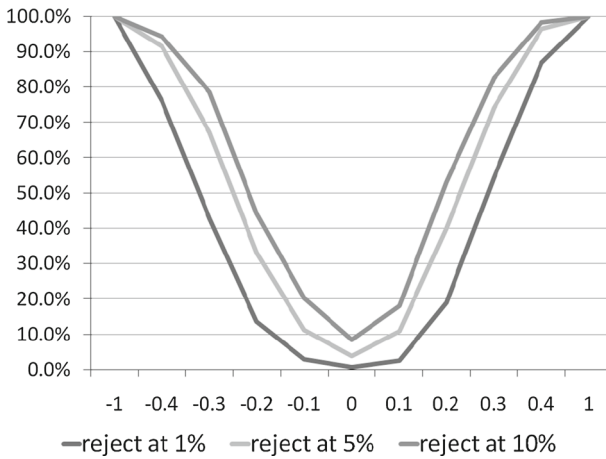**Fig. 9** $sup F(1|0)$ and $sup F(2|1)$ plotted against θ, $m = 2$

Fig. 10 $sup F(2)$ plotted against θ, $m = 2$

The size and power of $sup F(2|1)$ and $sup F(2)$ are depicted in Figures 9 (right-hand graph) and 10, respectively. While the size is very similar, we observe that $sup F(2)$ outperforms $sup F(2|1)$ in terms of power.

5 U.S. inflation

We use the monthly U.S. consumer price index (CPI) collected by the Organization for Economic Cooperation and Development. The sample runs from January 1966 until June 2008, yielding 509 observations. Inflation is computed as the annualized monthly change in CPI: $p_t = 1,200(\log(CPI_t) - \log(CPI_{t-1}))$.

5.1 Preliminary analysis

It has been argued that long memory may be spurious and caused by breaks in the mean or by regime shifts. In particular, Lobato and Savin (1998) raised the question of whether the long memory in inflation is due to deterministic shifts. See also Sibbertsen (2004) for a corresponding survey paper. In order to avoid any confusion between mean shifts and long memory, we allow for a shift in the overall mean while seasonally demeaning at the same time. The demeaned inflation rate becomes

$$
y_t = \begin{cases} p_t - \hat{\mu}_1(\tau_0) - seas_t, \ t = 1, \ldots, [\tau_0 T] \\ p_t - \hat{\mu}_2(\tau_0) - seas_t, \ t = [\tau_0 T] + 1, \ldots, T \end{cases}
$$

where τ_0 is the unknown, potential break fraction, $\hat{\mu}_1(\tau_0)$ $(\hat{\mu}_2(\tau_0))$ is the estimated mean before (after) the break point and $seas_t$ is the effect of seasonality.[6] In order to find τ_0, we adopt an approach developed by Hsu (2005) who modified the local Whittle [LW] estimator for d, discussed by Robinson (1995). In the same way we modify the more refined exact local Whittle [ELW] estimator by Shimotsu and Phillips (2005). In a grid search over $\tau \in [0.15, 0.85]$, d is estimated while accounting for a mean shift and seasonality at the same time. The modified criterion function is

$$
R(d; \tau) = \log(G(d; \tau)) - \frac{2d}{B} \sum_{i=1}^{B} \log(\lambda_i)
$$

with

$$
G(d; \tau) = \frac{1}{B} \sum_{i=1}^{B} I_{\Delta^d y}(\lambda_i; \tau), \tag{19}
$$

where λ_i are harmonic frequencies $\lambda_i = 2\pi i / T, i = 1, \ldots, B$, and the bandwidth B is usually chosen according to

$$
B = T^{\alpha}, \quad 0.5 < \alpha < 0.8.
$$

Further, $I_{\Delta^d y}(\lambda_i; \tau)$ denotes the periodogram evaluated from $\Delta^d y_t$ for a given mean shift fraction τ. Denote the conditional ELW estimator obtained for given τ in a first minimization as $\widehat{d}(\tau)$, while a second optimization step is necessary to find the change-point estimator $\widehat{\tau}$:

$$
\widehat{\tau} = \arg \min_{\tau \in [0.15, 0.85]} R(\widehat{d}(\tau); \tau).
$$

---

[6] Seasonality is accounted for by twelve monthly dummies (dum_{seas}), the break in mean is accounted for by a mean dummy (dum_{μ}) taking on the value one before and the value 0 after $[\tau_0 T]$. The variable y_t is the residual of the regression of p_t on dum_{μ} and dum_{seas}.

Table 2 Estimation of the order of integration

B	$T^{0.60} = 42$	$T^{0.65} = 57$	$T^{0.70} = 78$	$T^{0.75} = 107$
$\widehat{d}(\widehat{\tau})$	0.44	0.34	**0.35**	0.29
CI	[0.31, 0.57]	[0.23, 0.45]	**[0.25, 0.44]**	[0.21, 0.37]
$\widehat{\tau} T$	1981/8	1981/10	**1981/10**	1982/7
HR	2.07	2.65	**2.61**	3.07

Notes $\widehat{d}(\widehat{\tau})$ is the modified ELW estimate of d of the seasonally adjusted inflation which is demeaned accounting for one mean shift. The bandwidth is denoted by B. The 90 % confidence interval of the estimation of d is given in parentheses. Furthermore, the date of the mean shift and the HR statistic is given. The statistics correspond to the most appropriate choice of B are highlighted

The modified ELW estimator for the memory parameter d is $\widehat{d}(\widehat{\tau})$. Since the estimator $\widehat{\tau}$ converges to the true normalized change point τ_0 (see Lavielle and Ludeña 2000), Hsu (2005) argues that the limiting distribution is not affected. From Shimotsu and Phillips (2005) we conclude

$$2\sqrt{B} \, (\widehat{d}(\widehat{\tau}) - d) \xrightarrow{d} \mathcal{N}(0, 1), \qquad (20)$$

which allows to compute approximate confidence intervals.

Next, we wish to test whether the mean shift is significant, $H_0 : \mu_1 = \mu_2$, using a test statistic proposed by Hidalgo and Robinson (1996):

$$HR = T^{d-0.5} \frac{\hat{\mu}_1(\tau_0) - \hat{\mu}_2(\tau_0)}{\sqrt{\Omega}} \sim \mathcal{N}(0, 1),$$

where Ω depends on $G(d; \tau)$. To obtain a feasible version of the test statistic, the unknown parameters are replaced by the estimators $\widehat{\tau}$ and $\widehat{d}(\widehat{\tau})$.

We repeat the empirical analysis for different values of the bandwidth: $B \in \{T^{0.60}, T^{0.65}, T^{0.70}, T^{0.75}\}$. The candidate for the break fraction τ_0 lies in the interval 1981/8 to 1982/7, depending on the bandwidth B, see Table 2. For all choices of bandwidth, the mean shift is clearly significant according to the HR statistic. The timing of the mean shifts roughly coincide with the break date found by Hsu (2005) and is consistent with previous literature. Among others, Meltzer (2006) and Stock (2001) describe the level of inflation as high in the 1970s and early 1980s and low afterwards.

Table 2 reports the estimates of the order of integration. Needless to say, the appropriate choice of B is of crucial importance. If B is chosen too small, the estimate has a great standard deviation and might be imprecise. By contrast, choosing B too large results in a bias due to short memory components. Our estimate of d seems to stabilize for $B = T^{0.65}$ and $B = T^{0.70}$ while the estimate for $B = T^{0.75}$ seems to exhibit a small downward bias. Therefore, the choice $B = T^{0.70}$ maximizes the number of observations that do not lead to a bias: the results corresponding to this choice of B are highlighted below. The order of integration of inflation in the whole sample period is 0.35, with $B = T^{0.70}$, implying that inflation is stationary.
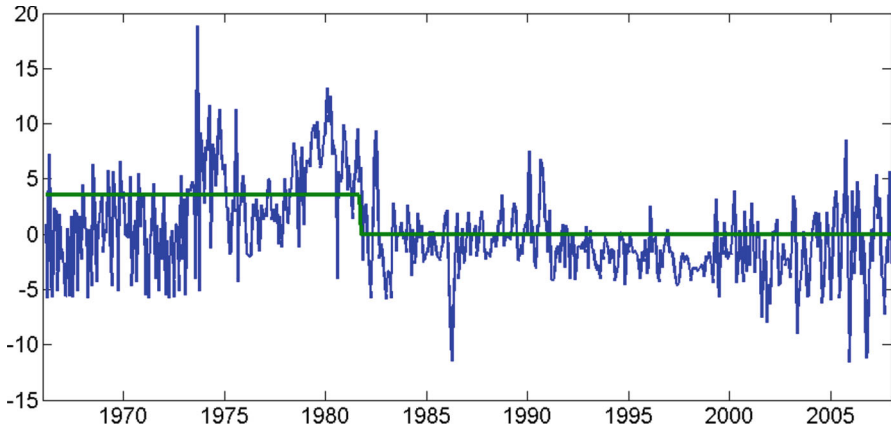
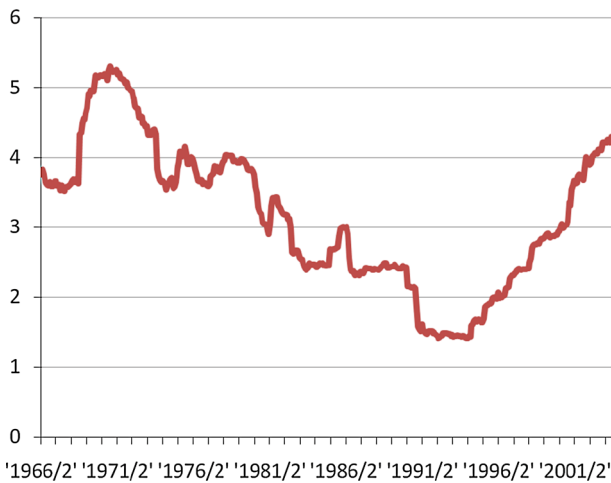**Fig. 11** Monthly U.S. inflation—seasonally demeaned, illustrating mean shift

Fig. 12 Rolling standard deviations for π_t (window of 5 years)

We investigate whether there is a second break in the mean. To this end, we proceeded sequentially, subtracting the first mean-shift from the series and searching for a second mean-shift.[7] The second break is insignificant, even at the 10 % significance level. For this reason, we only account for one shift in the mean. In Fig. 11, we plot inflation adjusted for seasonal means and illustrate the mean shift. In order to obtain our variable of interest, π_t, we then additionally adjusted for the mean shift.

Next, we visually investigate whether there has been a change in variance by inspecting the rolling standard deviations of inflation $s_t(\pi)$ for π_t (5 years window), depicted in Fig. 12. We observe that the eighties were characterized by a reduction in volatility.

---

[7] As an alternative to the sequential procedure we also allow for two mean shifts simultaneously and obtain similar break points and p values.
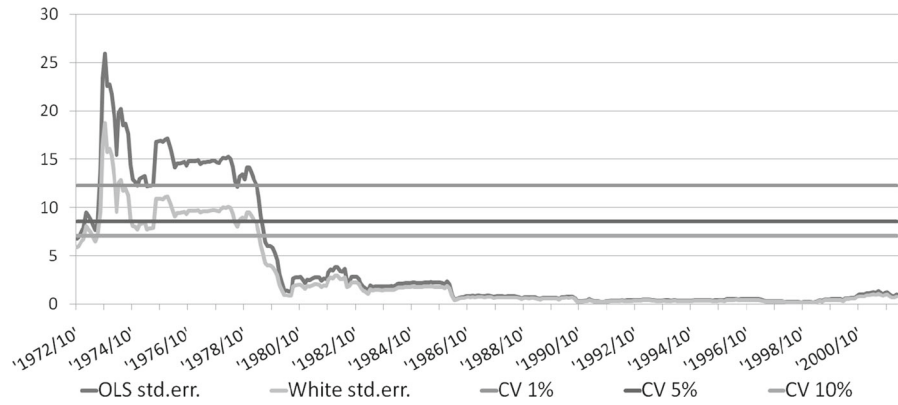
Fig. 13 $F(\lambda_1)$ (usual and Eicker–White standard errors) with critical values for a search in the interval 15–85 % of the observations

To account for this variance heterogeneity, we report Eicker–White standard errors in the next section as advocated by Demetrescu et al. (2008) and Kew and Harris (2009).

5.2 Testing against changes in inflation persistence

We now turn to the estimation of a change in persistence in U.S. inflation rates. As a first step, we apply the difference filter to the adjusted inflation rates (π_t):

$$x_t = (1 - L)^{\hat{d}} \pi_t,$$

where $\hat{d} = 0.35$ is the estimated order of integration of the whole sample as reported in Table 2. Note that the precise value of d used for differencing is not of major importance since we observed a considerable robustness with respect to misspec-ification, see Remark 2. Next, we estimate regression (18) with $m = 1$ using $p = \left[4(509/100)^{1/4}\right] = 6$ lags, and compute a sequence of F statistics, $F(\lambda_1)$, see Fig. 13. Their maximum values, $supF(1)$, are clearly significant, irrespective of whether $F(\lambda_1)$ is computed using usual or Eicker–White standard errors. Both versions of the test detect the break in October 1973.

Similarly, we observe that $supF(2)$ is significant at the 1 % level: the critical value is 9.36 while $supF(2)$ takes on the values 14.86 and 10.14 with usual standard errors and with Eicker–White robustified standard errors, respectively. Again, the first break is found in October 1973, while the second one is located in March 1980. Note, that $supF(1)$ is larger than $supF(2)$, suggesting that there is only one break.[8]

To verify whether there is a second change in persistence or not, we apply the $supF(2|1)$ test. In Fig. 14 we present $F_1(1)$ and $F_2(1)$ computed for the segments

---

[8] Bai and Perron (1998) also investigate a double maximum test, not considered in this paper. The number of break points is found by taking the maximum over all $supF(m)$ test statistics, where $m = 1, 2, ..., 5$. This maximum value is then compared to critical values in order to determine the significance. This suggests that in our analysis there is only one break point.
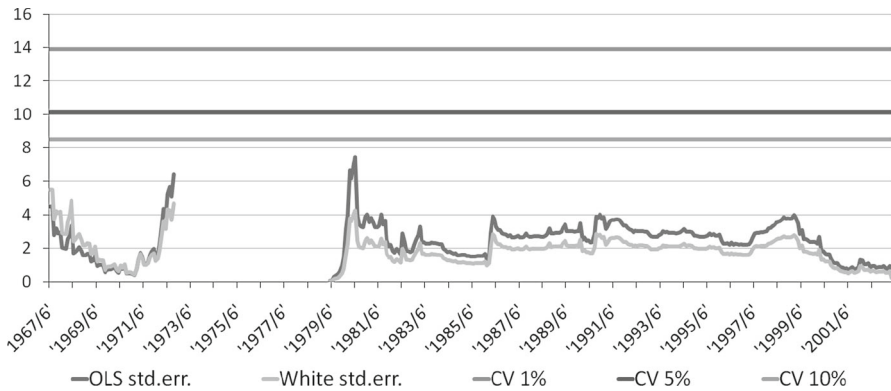
Fig. 14 $F_1(1)$ and $F_2(1)$ (OLS and Eicker–White standard errors) with critical values in search for a second break

before and after October 1973, respectively. The maximum thereof, $sup F(2|1)$, found in June 1980, is below 8.51 and hence not significant at the 10 % level, irrespective of whether robust Eicker–White standard errors are used or not.

As a robustness check, we test for a change in persistence of inflation without accounting for mean shifts. The results of these tests are similar and we come to the same conclusion. There is a break taking place in October 1973 with $sup F(2|1) = 46.56$ and 40.31, for OLS and Eicker–White standard errors, respectively. There is no evidence for a significant second break: $sup F_1(1) = 0.78$ and 1.11 (with the potential break point being June 1967) and $sup F_2(1) = 6.58$ and 4.58 (with the potential break point being April 1986), for OLS and Eicker–White standard errors, respectively.

One virtue of our approach is that it can find a change in d even with few—at least 150—observations before or after the break.[9] On this account, we were able to detect an early break in persistence taking place in 1973. Moreover, we can deduce the direction of the change in persistence from the sign of $\widehat{\psi}$. A positive coefficient indicates an increase in persistence after a break, while a negative coefficient indicates a decrease, where the dummy variable $D_t(\lambda)$ is defined as in (11).[10] In our estimation, $\widehat{\psi}$ is positive, leading us to the conclusion that inflation persistence has increased since 1973. Naturally, we would like to know the order of integration before and after the break. However, the short time period does not allow us to reliably estimate the order of integration before the break.[11]

---

[9] As becomes evident in Fig. 5, the power of the test increases with the difference in the order of integration before and after the break. Another factor is the total number of observations in the whole sample, see Fig. 6. If the difference in the order of integration is at least 0.3 and the sample size is 500, the test has a rejection rate of more than 80 % if there are at least 150 observations left before and after the break.

[10] This interpretation is inferred directly from Proposition 3 in Demetrescu et al. (2008) and the derivations of this paper.

[11] The order of integration was estimated to be 0.22 for $B = T^{0.70}$ with a 90 % confidence interval [0.01, 0.43]. However, the estimation depends heavily on single observations.

The order of integration after 1973/10 can be estimated more reliably. The point estimate is 0.27 with a 90 % confidence interval of [0.15, 0.39], for $B = T^{0.70}$. It is worth noting that this confidence interval overlaps with the confidence interval of d estimated over the whole sample. This is not surprising as the rather long second subsample starting in 1973/10 dominates the estimation results obtained for the whole sample.

To sum up, we conclude that inflation persistence increased after 1973 and stayed constant thereafter. The estimate of d is about 0.27 after the break. By looking at the confidence intervals, we come to the conclusion that inflation neither has short memory ($d \leq 0$) nor is nonstationary ($d \geq 0.5$). In addition to the break in persistence, we have evidence for a break in the mean and a trending behavior of the variance.

6 Concluding remarks

We proposed new tests against breaks in the order of fractional integration, which are built on the change-test methodology applied to the lag-augmented LM regression by Demetrescu et al. (2008). The procedures are $sup - F$ tests, specifically following Andrews (1993) in the case of one potential break and more generally following Bai and Perron (1998). In particular, the latter authors allow for a sequence of tests to determine the unknown number of changes. Monte Carlo simulations indicate that the power of the tests essentially depends on the size of the changes. Breaks relatively close to the end or beginning of the sample can be detected with remarkable reliability. Not knowing the true order of integration and working with estimated values does not affect the performance of the tests.

Using the new tools, we investigate whether inflation persistence, i.e. the order of integration of inflation, in the U.S. has changed. In order to forestall spuriously high orders of integration, we adjust inflation rates by accounting for a shift in the mean where the break point is determined endogenously. Testing adjusted inflation, we find an increase in its persistence in October 1973. A second potential break in March 1980 is not significant at the 10 % level. This result does not change if we do not account for a mean shift.

Many studies measure inflation persistence as the largest autoregressive root [LARR] or as the sum of autoregressive coefficients [SARC]. Those measures cannot discriminate between different degrees of long-run persistence, see Kumar and Okimoto (2007) and Gadea and Mayoral (2006). Therefore, it is not surprising that most of these studies do not find evidence for a break in persistence. In contrast, most studies using the order of integration as a measure of persistence, which in a wider sense also includes the studies of Cogley and Sargent (2001) and Cogley and Sargent (2005), find time-varying persistence. The studies find breaks taking place in the early 1970s, the early 1980s and/or the early 1990s. Employing Eicker–White standard errors, our tests are robust to the apparent time-varying inflation volatility (see, for example, Stock and Watson (2007) or Pivetta and Reis (2007) for evidence). We are led to the conclusion that there is only one change in persistence and this took place in 1973. This break date coincides with the end of the Bretton Woods system, a sharp increase in oil prices

and the start of an episode of high inflation.[12] Breaks in the eighties, documented in the literature, might be attributed to mean shifts or the decrease in inflation volatility.

Appendix

Proof of Proposition 1

Under Gaussianity and Assumption 2 the pseudo-log-likelihood function becomes

$$\mathcal{L}(\theta_1, \ldots, \theta_m; d, \sigma^2) = -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \varepsilon_t^2,$$

with $\varepsilon_t = \Delta^{d+\theta_{j-1}} y_t$ for $t = T_{j-1} + 1, \ldots, T_j$ using (3) and (4), or $\varepsilon_t = \Delta^{d+\theta_j} y_t$ for $t = T_j + 1, \ldots, T_{j+1}$, such that $(j = 1, \ldots, m)$

$$\frac{\partial \varepsilon_t}{\partial \theta_j} = \begin{cases} (\log(1-L)) (1-L)^{d+\theta_j} y_t, & t = T_j + 1, \ldots, T_{j+1} \\ 0 & \text{else} \end{cases}.$$

With $\log(1-L) = -\sum_{j=1}^{\infty} j^{-1} L^j$ we obtain for the score vector evaluated under the null (where $\Delta^d y_t = \varepsilon_t$)

$$S = \left(\frac{\partial \mathcal{L}}{\partial \theta_j} \bigg|_{\theta_j=0} \right)_{j=1,\ldots,m} = \left(\frac{1}{\sigma^2} \sum_{t=T_j+1}^{T_{j+1}} \varepsilon_t \varepsilon_{t-1}^{**} \right)_{j=1,\ldots,m} \qquad \text{with } \varepsilon_{t-1}^{**} = \sum_{j=1}^{\infty} j^{-1} \varepsilon_{t-j},$$

where $\{\varepsilon_{t-1}^{**}\}$ is a stationary process with variance

$$\sigma_{**}^2 = \text{Var}\left(\varepsilon_{t-1}^{**}\right) = \sigma^2 \sum_{j=1}^{\infty} j^{-2} = \sigma^2 \frac{\pi^2}{6}.$$

To construct the LM statistic we compute the Fisher information as the outer product of gradients,

$$\mathcal{I} = \text{E}\left(S\,S'\right) = \frac{\sigma_{**}^2}{\sigma^2} \text{diag}\left(T_2 - T_1, \ldots, T_{m+1} - T_m\right).$$

---

[12] The events are not described in order to indicate causality but rather in order to integrate the break date into its historical background.

Hence, we obtain

$$S' \mathcal{I}^{-1} S = \frac{1}{\sigma_{**}^2 \sigma^2} \sum_{j=1}^{m} \frac{\left(\sum_{t=T_j+1}^{T_{j+1}} \varepsilon_t \varepsilon_{t-1}^{**} \right)^2}{T_{j+1} - T_j}.$$

Since the LM statistic is evaluated under H_0, we replace ε_t with $\Delta^d y_t$. Given the starting value assumption in Assumption 2, this coincides with x_t defined in (7). Consequently, ε_{t-1}^{**} equals x_{t-1}^* from (9), and the LM statistic becomes LM from (8) as required.   □

Proof of Proposition 2

Write the regression equation (13) in obvious matrix notation, $y = X \widehat{\beta} + \widehat{\varepsilon}$, with

$$y' = (x_2, \ldots, x_T),$$
$$\widehat{\beta}' = (\widehat{\phi}, \widehat{\psi}_1, \ldots, \widehat{\psi}_m),$$

and X containing $(x_1^*, \ldots, x_{T-1}^*)'$ as the first column, while the other columns contain zeros and segments of $(x_{T_j}^*, \ldots, x_{T_{j+1}-1}^*)'$. Under Assumption 2, we have $x_t = \varepsilon_t \sim iid(0, \sigma^2)$. The required limiting distributions can be obtained as set out by Robinson (1991) or Tanaka (1999), see also Hassler and Breitung (2006, Lemma A):

$$\frac{X'X}{T} \xrightarrow{p} \sigma^2 \frac{\pi^2}{6} \Lambda^0$$

where $\xrightarrow{p}$ stands for convergence in probability, and

$$\Lambda^0 = \begin{pmatrix} 1 & (\lambda^0)' \\ \lambda^0 & \mathrm{diag}(\lambda_2^0 - \lambda_1^0, \ldots, \lambda_{m+1}^0 - \lambda_m^0) \end{pmatrix}, \quad (\lambda^0)' = (\lambda_2^0 - \lambda_1^0, \ldots, \lambda_{m+1}^0 - \lambda_m^0),$$

and

$$\frac{X'y}{\sqrt{T}} \xrightarrow{d} \mathcal{N}_{m+1} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \sigma^4 \frac{\pi^2}{6} \Lambda^0 \right).$$

Consequently, $\sqrt{T} \widehat{\beta}$ follows a limiting normal distribution with $\Sigma = \frac{6}{\pi^2} (\Lambda^0)^{-1}$. Define the $m \times (m+1)$ matrix R with $R \widehat{\beta} = (\widehat{\psi}_1, \ldots, \widehat{\psi}_m)'$. The F statistic becomes

$$F(\lambda_1^0, \ldots, \lambda_m^0) = \frac{T - m - 1}{m} \frac{\widehat{\beta}' R' \left(R (X'X)^{-1} R' \right)^{-1} R \widehat{\beta}}{\widehat{\varepsilon}' \varepsilon},$$

and its limiting distribution follows the usual way.   □

References

Andrews DWK (1993) Tests for parameter instability and structural change with unknown change point. Econometrica 61:821–56

Bai J, Perron P (1998) Estimating and testing linear models with multiple structural changes. Econometrica 66:47–78

Bai J, Perron P (2003a) Computation and analysis of multiple structural change models. J Appl Econ 18:1–22

Bai J, Perron P (2003b) Critical values for multiple structural change tests. Econ J 6:72–78

Baillie RT, Chung C-F, Tieslau MA (1996) Analysing inflation by the fractionally integrated ARFIMA–GARCH model. J Appl Econ 11:23–40

Baillie RT, Kapetanios G (2007) Testing for neglected nonlinearity in long-memory. J Bus Econ Stat 25:447–461

Baillie RT, Kapetanios G (2008) Nonlinear models for strongly dependent processes with financial applications. J Econ 147:60–71

Beran J, Terrin N (1996) Testing for a change of the long-memory parameter. Biometrika 83:233–233

Beran J, Terrin N (1999) Correction: testing for a change of the long-memory parameter. Biometrika 86:627–638

Breitung J, Hassler U (2002) Inference on the cointegration rank in fractionally integrated processes. J Econ 110:167–185

Busetti F, Taylor AMR (2004) Tests of stationarity against a change in persistence. J Econ 123:33–66

Cogley T, Sargent TJ (2001) Evolving post-world war II U.S. inflation dynamics. NBER Macroecon Ann 16:331–373

Cogley T, Sargent TJ (2005) Drift and volatilities: monetary policies and outcomes in the post WWII U.S. Rev Econ Dyn 8:262–302

Demetrescu M, Hassler U, Kuzin V (2011) Pitfalls of post-model-selection testing: experimental quantification. Empir Econ 40:359–372

Demetrescu M, Kuzin V, Hassler U (2008) Long memory testing in the time domain. Econ Theory 24:176–215

Gadea MD, Mayoral L (2006) The persistence of inflation in OECD countries: a fractionally integrated approach. Int J Central Banking 2:51–104

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

Gil-Alana LA (2008) Fractional integration and structural breaks at unknown periods of time. J Time Ser Anal 29:163–185

Hassler U, Breitung J (2006) A residual-based lm-type test against fractional cointegration. Econometric Theory 22:1091–1111

Hassler U, Kokoszka P (2010) Impulse responses of fractionally integrated processes with long memory. Econ Theory 26:1855–1861

Hassler U, Rodrigues PMM, Rubia A (2009) Testing for general fractional integration in the time domain. Econ Theory 25:1793–1828

Hassler U, Scheithauer J (2011) Detecting changes from short to long memory. Stat Pap 52:847–870

Hassler U, Wolters J (1995) Long memory in inflation rates: international evidence. J Bus Econ Stat 13:37–45

Hidalgo J, Robinson PM (1996) Testing for structural change in a long-memory environment. J Econ 70:159–174

Hsu C-C (2005) Long memory or structural changes: an empirical examination on inflation rates. Econ Lett 88:289–294

Kew H, Harris D (2009) Heteroskedasticity-robust testing for a fractional unit root. Econ Theory 25:1734–1753

Kim J-Y (2000) Detection of change in persistence of a linear time series. J Econ 95:97–116

Kumar MS, Okimoto T (2007) Dynamics of persistence in international inflation rates. J Money, Credit Banking 39:1457–1479

Lavielle M, Ludeña C (2000) The multiple change-points problem for the spectral distribution. Bernoulli 6:845–869

Leeb H, Pötscher BM (2005) Model selection and inference: facts and fiction. Econ Theory 21:21–59

Lobato IN, Savin NE (1998) Real and spurious long-memory properties of stock-market data. J Bus Econ Stat 16:261–268

Martins LF, Rodrigues PMM (2012) Testing for persistence change in fractionally integrated models: an application to world inflation rates. Comput Stat Data Anal doi:10.1016/j.csda.2012.07.021

Meller B, Nautz D (2012) Inflation persistence in the euro area before and after the European Monetary Union. Econ Model 29:1170–1176

Meltzer AH (2006) From inflation to more inflation, disinflation, and low inflation. Am Econ Rev 96:185–188

Phillips PCB, Xiao Z (1999) A primer on unit root testing. J Econ Surv 12:423–470

Pivetta F, Reis R (2007) The persistence of inflation in the United States. J Econ Dyn Control 31:1326–1358

Ray BK, Tsay RS (2002) Bayesian methods for change-point detection in long-range dependent processes. J Time Ser Anal 23:687–705

Robinson PM (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. J Econ 47:67–84

Robinson PM (1994) Efficient tests of nonstationary hypotheses. J Am Stat Assoc 89:1420–1437

Robinson PM (1995) Gaussian semiparametric estimation of long range dependence. Ann Stat 23:1630–1661

Robinson PM (2005) The distance between rival nonstationary fractional processes. J Econ 128:283–399

Schwert GW (1989) Tests for unit roots: a Monte Carlo investigation. J Bus Econ Stat 7:147–159

Shimotsu K, Phillips PCB (2005) Exact local whittle estimation of fractional integration. Ann Stat 33:1890–1933

Sibbertsen P (2004) Long memory versus structural breaks: an overview. Stat Pap 45:465–515

Sibbertsen P, Kruse R (2009) Testing for a break in persistence under long-range dependencies. J Time Ser Anal 30:263–285

Stock JH (2001) Evolving post-world war II U.S. inflation dynamics: comment. NBER Macroecon Ann 16:379–387

Stock JH, Watson MW (2007) Why has U.S. inflation become harder to forecast? J Money, Credit Banking 39:3–33

Tanaka K (1999) The nonstationary fractional unit root. Econ Theory 15:549–583

Tillmann P (2012) Has inflation persistence changed under EMU? Ger Econ Rev 13:86–102

Yamaguchi K (2011) Estimating a change point in the long memory parameter. J Time Ser Anal 32:304–314