Distinguishing between long-range dependence and deterministic trends

by

Philipp Sibbertsen[1] **and Ioannis Venetis**

Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund, Germany

Cardiff Business School, Cardiff University, Cardiff CF10 3EU, UK

Version May 2003

Abstract

We provide a method for distinguishing long-range dependence from deterministic trends such as structural breaks. The method is based on the comparison of standard log-periodogram regression estimation of the memory parameter with its tapered counterpart. The difference of these estimators provides the desired test. Its asymptotic distribution depends on the true memory parameter under the null, and is therefore estimated by bootstrapping. The test is applied to inflation rates of three industrialized countries.

KEY WORDS: Long memory, trends, log-periodogram regression, inflation rates

JEL - classification: C 12; C 22; C 14

[1]corresponding author; e-mail: sibberts@statistik.uni-dortmund.de

1 Introduction

It is a well discussed problem that structural breaks as well as slowly decaying trends are often misspecified as long-range dependence. On the other hand, long memory can easily be mistaken as a break in trend. Beginning with Bhattacharya et al.(1983) several authors constructed trends which artificially produce a Hurst effect and thus look as if having long memory (Diebold/Inoue, 2001). It is stated in many papers that standard methodology fails in this context. Krämer/Sibbertsen(2002) proved among others that tests on structural breaks are not able to distinguish structural breaks and long-range dependence and Giraitis et al.(2001) showed that R/S-based tests on long memory fail also for a quite general class of trends.

But the question whether a data set has real or spurious long memory has deep impact to many economic applications. For example, there is evidence of long-range dependence in the volatilities of many stock returns (Krämer et al., 2002), which will influence the price of options based on this stock (Bollerslev/Mikkelsen, 1996). Also long-memory time series allow for optimal long-term forecasts which would not be possible in a model with a deterministic trend disturbed by some independent or short-memory noise process. Bos et al.(1999) consider the problem of long-range dependence and level shifts in inflation rates. For an overview about the problem of distinguishing long memory and major deterministic trends see Sibbertsen(2002a).

But nevertheless there is still no method at hand for distinguishing both of these phenomena. So far, approaches dealing with this problem focused mainly on R/S-methodology. But Künsch(1986) already showed that the periodogram is able to distinguish monotonic trends and long memory. Even though Künschs results are not valid for non-monotonic trends it indicates that periodogram based methods seem more appropriate than rescaled-range based approaches. In this context Sibbertsen(2003) found by Monte Carlo that log-periodogram based estimates for the memory parameter allow to distinguish quite general deterministic trends and long memory. The test constructed in this paper is based on this idea. To the series under test standard log-periodogram regression is applied as well as tapered log-periodogram regression. Both estimates are consistent for the memory parameter. But whenever major trends are present in the data the standard log-periodogram estimator behaves completely different than its tapered counterpart. Hence our test statistic is based on the difference of these estimators. This idea is similar to Hausman tests (Hausman, 1978). However, the limiting distribution under the null of no trend depends on the true memory parameter. Therefore, it will be estimated by bootstrapping.

The rest of the paper is organized as follows. In the next section long memory is defined and log-periodogram regression estimators are explained. Section 3 introduces the test statistic and its main properties and in section 4 bootstrap methods for estimating the limiting distribution of the statistic are described. Section 5 contains some Monte Carlo results concerning the power of the test. Application to inflation rates of three industrialized countries, namely the US, UK and Germany, is given in section 6. Section 7 concludes.

2 Log-Periodogram Regression

In this section long-memory time series as well as log-periodogram regression estimators for the memory parameter are introduced. A time series X_t is said to exhibit long memory or long-range dependence if the correlation function $\rho(k)$ behaves for $k \to \infty$ as

$$\lim_{k\to\infty} \frac{\rho(k)}{c_\rho k^{2d-1}} = 1. \tag{1}$$

Here c_ρ is a positive constant and $d \in (0, 0.5)$ denotes the memory parameter. This means that observations far away from each other are still strongly correlated. Thus the correlations of a long-memory process decay slowly that is with a hyperbolic rate and consequently they are no longer summable. This is the most important difference to short-memory processes.

Using the spectral density of the process we can obtain an equivalent definition of long memory which is the base for log-periodogram regression estimates. In this context a time series X_t is said to exhibit long memory if the spectral density $f(\lambda)$ behaves for $\lambda \to 0$ as

$$\lim_{\lambda \to 0} \frac{f(\lambda)}{c_f |\lambda|^{2d}} = 1. \tag{2}$$

Here c_f is a positive constant and again $d \in (0, 0.5)$ denotes the memory parameter. Thus the spectral density has a pole at the origin.

Long-memory processes can be represented as $I(d)$-processes with fractional $d \in (0, 0.5)$. This can be done by generalising Box/Jenkins(1976) ARMA-models to ARFIMA-models. ARFIMA-models were introduced by Granger/Joyeux(1980) and independently by Hosking(1981). Allowing also for short-memory terms we have the representation

$$\Phi(B)(1-B)^d X_t = \Psi(B)\varepsilon_t, \tag{3}$$

where B denotes the Backshift operator, ε_t is a mean zero finite variance white noise process and $\Phi(z), \Psi(z)$ denote the autoregressive and moving average polynomials respectively. For an exact definition of fractional integration and further details about long-memory processes see Beran(1994) or Sibbertsen(1999). For an overview about long-range dependence in economics see Bailie(1996).

One possibility of estimating the memory parameter d is log-periodogram regression introduced by Geweke/Porter-Hudak(1983) (further referred as GPH-estimation). This approach is based on the representation (2) of the spectral density of a long-memory process near the origin. The idea is to estimate the spectral density by using the periodogram. Taking the logarithm gives a linear regression model. For defining the estimator exact denote with

$$I_X(\lambda_j) := \frac{1}{2\pi N} |\sum_{t=1}^{N} X_t \exp(\frac{-it2\pi j}{N})|^2$$

the periodogram of the process X_t. The GPH-estimator is now defined as the least-squares estimator of d based on the regression equation

$$\log I_X(\lambda_j) = \log c_f - 2d \log \lambda_j + \log \xi_j, \tag{4}$$

where λ_j denotes the $j-th$ Fourier frequency, that is $\lambda_j = 2\pi j/n$ and the ξ_j are identically distributed errors with $E[\log \xi_j] = -0.577$, known as Euler constant.

Besides simplicity the main advantage of the GPH-estimator is that it does not require any further knowledge about short-term components. Consistency of the estimator can also be obtained without knowledge of the distribution of the data generating process (Robinson, 1995 or Hurvich et al., 1998). Only for proving asymptotic normality it is required that the data generating process is normally distributed.

Disadvantages of this approach result from the fact that the errors in the regression equation (4) are not independent. Another problem is that the representation (2) of the spectral density holds only near the origin. Thus, a trade off between bias and variance has to be made by taking the optimal number of frequencies used for the estimation. Whereas Geweke/Porter-Hudak(1983) proposed a number of $N^{1/2}$, which is still used in many applications, Hurvich et al.(1998) showed that a rate of $N^{4/5}$ is MSE-optimal. This rate will be used in this paper. Here and in the following N denotes the sample size.

The standard GPH-estimator can be modified by using the tapered periodogram instead of the standard periodogram. Hurvich/Ray(1995) and Velasco(1999) showed that the tapered version gives better results in the case of non-stationary long-memory processes that is $d > 0.5$. As we see in the next section this holds also true for non-stationarities produced by deterministic trends. The idea of data tapers is to apply a smoothing function to the data which gives smaller weights to the low frequencies in the periodogram. Low frequencies are important in the case of non-stationarities. Thus the influence of the trend is reduced by the taper.

The periodogram of the tapered process $w_t X_t$ is defined by

$$I_{T,X}(j) = \frac{1}{2\pi \sum w_t^2} | \sum_{t=0}^{N-1} w_t X_t e^{-i\lambda_j t}|^2.$$

Here λ_j again denotes the j-th Fourier frequency and w_t denotes the taper. We use in this paper the full cosine bell taper given by

$$w_t = \frac{1}{2}[1 - \cos(\frac{2\pi(t+0.5)}{N})].$$

Velasco(1999) proves consistency and asymptotic normality of the tapered GPH-estimator. For a detailed discussion of tapering see Bloomfield(1992).

Thus, we have two consistent estimates for the memory parameter. Whereas the standard GPH-estimator is strongly biased in the case of major deterministic trends this bias is reduced by its tapered version. A test using this property is constructed in the following section.

3 The Test Statistic

From now on the model under test is the following

$$X_t = f(t) + Y_t, \tag{5}$$

where $f(t)$ is a deterministic trend specified later and Y_t is a noise process having zero mean and finite variance.

For defining the trend we follow Giraitis et al.(2001) and use their quite technical but weak assumptions. They include slowly decaying trends as well as change point models. We have the following assumptions for the trend $f(t)$:

Assumption T1: $[f^{(N)}(k)]_{k=1,\ldots,N}$, $N \geq 1$, is an array of real numbers for which there exists a positive sequence p_N and a function h on $[0,1]$, which is not identically zero, such that for $N \to \infty$

$$p_N^{-1} \sum_{k=1}^{[Nt]} f^{(N)}(k) \to h(t)$$

and

$$\frac{p_N}{N^{1/2}} \to a,$$

where $a \in [0, \infty]$. We further assume for the trend

Assumption T2: There exists a positive sequence $r_N \to \infty$ and numbers $0 < b, b^* < \infty$, such that as $N \to \infty$

$$r_N^{-1} \sum_{k=1}^{N} [f^{(N)}(k)]^2 \to b,$$

$$\sum_{k=1}^{N-1} |f^{(N)}(k) - f^{(N)}(k+1)| k^{1/2} = O(r_N^{1/2}),$$

$$\sum_{k=1}^{q_N} |f^{(N)}(k)|^2 = o(r_N)$$

for any $q_N = o(N)$,

$$|f^{(N)}(k)|^2 = O(r_N/N)$$

for $k \sim N$ and

$$\frac{p_N^2}{N r_N} \to b^* < \infty.$$

Assumption T1 describes the rate of decay of the trend function and assures that the trend is slowly decaying. Assumption T2 sets regulations to the variation of the trend. The trend cannot vary too much and poles are excluded.

Examples: (1) These assumptions cover structural breaks in the data. For the shift in mean model see Giraitis et al.(2001). Also generalizations of this model allowing for linear mean functions are covered.

Another function of great practical interest is the logistic regression function which is also considered in the simulations below. It is given by

$$f(t) = a + \frac{b}{1 + \exp(-\gamma(\frac{t}{N} - c))}, \tag{6}$$

where $a, b \in \mathbf{R}$, $\gamma > 0$ and $c \in [0, 1]$.

This function is appropriate for modelling changes in the mean and thus also structural breaks. Depending on the choice of parameters this function models a break from a to $a+b$ happening at the point cT. The parameter γ regulates how smooth the break happens. For small values of γ the function goes smoothly from a to $a+b$ for big parameter values we obtain a sudden shift in the mean.

This function fulfills the assumptions T1 and T2, too. Assumption T1 is fulfilled with $p_N = N^{1/2}$ because in the case of t/N being lower c the argument of the exponential function is positive and thus in the worst case the second term tends to zero and the function in its whole to a. If on the other hand t/N is greater or equal than c, the argument is lower or equal zero and thus the denominator tends to one and the function tends to $a+b$. Because the function is smooth in between assumption T1 is fulfilled.

The interesting part in assumption T2 is the second equation which assures that the decay of the function f is slower than with rate $N^{1/2}$. This is fulfilled for this function because around the point $k = [cT]$ the function $f(t)$ changes from a to $a+b$ whereas it is constant before and after this change. The exact duration to come from a to $a+b$ depends on the parameter γ. By this argument it is seen that also assumption T2 is fulfilled for the logistic regression function.

(2) Although the assumptions above and the theorems below are stated for deterministic trend functions the theory is also transmittable to stochastic components in the trend. In this example we consider the single change point model with a random breakpoint rather than a fixed deterministic. Of course the convergences in assumption T1 and T2 are now convergences in probability. We consider the trend function

$$f^{(N)}(k) = m_1, \quad 1 \le k \le S$$

and

$$f^{(N)}(k) = m_2, \quad S < k \le N.$$

Here S denotes a random variable with values between one and N. Thus, there exists a random variable τ with $0 < \tau < 1$ such that $S = [\tau N]$. Assume furthermore that $m_1 \neq m_2$. Then assumption T1 holds with $p_N = N, a = \infty$ and

$$h(t) = m_1 \min(t, \tau) + m_2(t - \min(t, \tau)).$$

Assumption T2 is satisfied with $r_N = N, b^* = 1$ and $b = m_1^2 \tau + m_2^2(1 - \tau)$.

For simplicity we showed that the assumptions are fulfilled for a random change point model with only one breakpoint. These considerations are easy to generalize to any finite number of breaks.

In the case of an infinite amount of breaks where the break times follow a power law Davidson/Sibbertsen (2002) showed that it is possible to construct long-memory processes by crosswise aggregation of independent copies of these processes. They also showed that processes constructed following this approach do not converge to fractional Brownian motion without aggregation but converging to a stable Levy motion in this case. Thus, the case of infinite breaks need extra consideration which is left for future work.

We have the test problem:

$$H_0 : f(t) \equiv 0$$

versus

$$H_1 : f(t) \text{fulfills assumption T1 and T2.}$$

Let us mention at this point that only trends fulfilling assumptions T1 and T2 are of interest in the alternative here. Major trends which do not fulfill assumptions T1 and T2 may cause technical problems, because the mean of the GPH-estimates can diverge in their presence. But again speaking in terms of applications these trends will hardly be misspecified as long-range dependence. Standard analysis will show up a non-stationary behaviour of the data rather than long memory. From this point of view assumptions T1 and T2 do not restrict the applicability of the method.

Denote from now on the standard GPH-estimator with $\hat{d}$ and the tapered GPH-estimator with $\hat{d}_T$. With m we denote the number of frequencies used for the estimation. Of course this number has to be equal for the standard and tapered estimator in our test. The test statistic will be defined as the squared difference of the standard and tapered GPH estimator. Before introducing the test statistic itself we prove that using this idea provides a method to distinguish trends and long-range dependencies.

Denote for this at first with

$$D := m^{1/2}(\hat{d} - \hat{d}_T)$$

the difference between both estimators.

To fix the notation denote from now on convergence in probability by $\xrightarrow{P}$ and convergence in distribution by $\xrightarrow{d}$.

To prove that the test is able to distinguish deterministic trends and long memory it has to be shown that $D \xrightarrow{P} 0$ under the null hypothesis and $D \xrightarrow{P} M(f)$, where $M(f)$ is a non-zero function depending on the trend function f, otherwise. This is done in the following.

Theorem 1 *Under H_0 we have $D \xrightarrow{P} 0$.*

Proof: It is a well known fact that both estimators are consistent for the true memory parameter d_0 of the underlying noise process (Robinson, 1995, Velasco, 1999 or Hurvich/Ray, 1995). Thus, their difference converge to zero in probability. This proves the theorem. $\diamondsuit$

The next theorem shows that the statistic is able to detect major trends.

Theorem 2 *Under the alternative H_1 it holds*

$$D \xrightarrow{P} M(f),$$

where $M(f)$ is a non-zero function depending on the trend function f.

Proof:

To prove the theorem it is enough to show

$$I_X(j) - I_{T,X}(j) \xrightarrow{P} \tilde{M}(f), \tag{7}$$

where $I_X(j)$ denotes again the periodogram based on X and $I_{T,X}(j)$ denotes the tapered periodogram based on X, the process X is as defined in (5) and $\tilde{M}(f)$ is another non-zero function. Proving (7) is enough because the stochastic behaviour of the GPH-estimates depends only on the behaviour of the periodogram. Because all other terms are equal anyway for the tapered and the non-tapered estimator showing that the stochastic part is different is enough to prove that both estimators are not equal. Because this is the only point of interest in this theorem we do not have to care for the exact representation of the error terms in the regression equation (4) defining the GPH-estimator.

Analytically the following shows that from proving (7) it follows that both estimators are not equal. From the definition of the GPH-estimator and the tapered GPH-estimator we have:

$$\begin{aligned} \hat{d} - \hat{d}_T &= \frac{-2\log\lambda_j \log I_X(\lambda_j)}{4\log^2\lambda_j} + \frac{2\log\lambda_j \log I_{T,X}(\lambda_j)}{4\log^2\lambda_j} \\ &= \frac{2\log\lambda_j}{4\log^2\lambda_j}\left(\log I_{T,X}(\lambda_j) - \log I_X(\lambda_j)\right) \end{aligned}$$

Because the logarithm is a monotonous function it is clear that (7) implicates that $\hat{d} - \hat{d}_T$ is nonzero.

In respect of the results of Hurvich et al.(1998) it is enough to consider the difference of the periodograms itself rather than the difference of the logarithm of the periodograms what would be indicated by the form of the estimator. Considering the logarithm would not lead to any further difficulties for the purpose of this proof even not for low frequencies close to zero. Thus, for simplicity of the presentation we do the proof by considering the differences of the raw periodograms.

In what follows λ_j denotes again the $j-$th Fourier frequency. Let us first consider:

$$
\begin{aligned}
I_X(j) - I_{T,X}(j) &= \\
\frac{1}{2\pi N}|\sum_{t=1}^{N}(Y_t + f(t))e^{-it\lambda_j}|^2 - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}|\sum_{t=1}^{N}(Y_t + f(t))w_t e^{-it\lambda_j}|^2 &= \\
\frac{1}{2\pi N}(\sum_{s=1}^{N}\sum_{t=1}^{N}(Y_sY_t + Y_sf(t) + Y_tf(s) + f(s)f(t))e^{-i(t-s)\lambda_j}) &- \\
\frac{1}{2\pi\sum_{t=1}^{N} w_t^2}(\sum_{s=1}^{N}\sum_{t=1}^{N}(Y_sY_t + Y_sf(t) + Y_tf(s) + f(s)f(t))w_sw_te^{-i(t-s)\lambda_j}) &= \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sY_te^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sY_tw_sw_te^{-i(t-s)\lambda_j} &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sf(t)e^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sf(t)w_sw_te^{-i(t-s)\lambda_j} &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)Y_te^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)Y_tw_sw_te^{-i(t-s)\lambda_j} &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)f(t)e^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)f(t)w_sw_te^{-i(t-s)\lambda_j} &= \\
I_Y(j) - I_{T,Y}(j) &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sf(t)e^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}Y_sf(t)w_sw_te^{-i(t-s)\lambda_j} &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)Y_te^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)Y_tw_sw_te^{-i(t-s)\lambda_j} &+ \\
\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)f(t)e^{-i(t-s)\lambda_j} - \frac{1}{2\pi\sum_{t=1}^{N} w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N}f(s)f(t)w_sw_te^{-i(t-s)\lambda_j}.
\end{aligned}
$$

Let us now denote the first difference with A, the second with B, the third with C and the last with D.

A has no impact because it is the difference of the periodogram and the tapered periodogram of the noise process only. The noise process is only a short- or long-memory process containing no disturbance by any trend. Thus, the difference of both of them tends to zero.

D is non-stochastic and thus the expression here is the square of the classical Fourier transform of the function f and of those after applying the smoothing taper w_t to f. Because $0 \leq w_t \leq 1$, for all t, and $w_t \neq 0$ and $w_t \neq 1$ for at least one t both functions are different. From assumption T2 we have that the Fourier transform of $f(t)$ converges. Hence the difference is non-zero.

The mixed terms B and C are remaining. For those we obtain:

$$\frac{1}{2\pi N}\sum_{s=1}^{N}\sum_{t=1}^{N} Y_s f(t) e^{-i(t-s)\lambda_j} - \frac{1}{2\pi \sum w_t^2}\sum_{s=1}^{N}\sum_{t=1}^{N} Y_s f(t) w_s w_t e^{-i(t-s)\lambda_j} =$$
$$\frac{1}{2\pi N}\sum_{t=1}^{N} f(t) \sum_{s=1}^{N} Y_s e^{-i(t-s)\lambda_j} - \frac{1}{2\pi \sum w_t^2}\sum_{t=1}^{N} f(t) \sum_{s=1}^{N} Y_s e^{-i(t-s)\lambda_j} =$$
$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N} f(t) \frac{1}{2\pi\sqrt{N}} \sum_{s=1}^{N} Y_s e^{-i(t-s)\lambda_j} - \frac{1}{\sqrt{N}}\sum_{t=1}^{N} f(t) \frac{8}{6\pi\sqrt{N}} \sum_{s=1}^{N} Y_s w_s w_t e^{-i(t-s)\lambda_j}$$

For the last equality we use the property of the cosine bell taper that $\sum w_t^2 = \frac{3}{8}N$.

Assumptions T1 and T2 give that $1/\sqrt{N}\sum_{t=1}^{N} f(t) \to h(t)$ with a function $h(t)$ as in the assumptions. Because Y_s was a mean zero random variable with finite variance the terms under the other two sums fulfill the Lindeberg condition. From the limit theorem of Lindeberg-Feller we obtain

$$\frac{1}{2\pi\sqrt{N}} \sum_{s=1}^{N} Y_s e^{-i(t-s)\lambda_j} \xrightarrow{d} \xi$$

and

$$\frac{8}{6\pi\sqrt{N}} \sum_{s=1}^{N} Y_s w_s w_t e^{-i(t-s)\lambda_j} \xrightarrow{d} \tilde{\xi},$$

where ξ and $\tilde{\xi}$ are standard normal random variables. Hence we have

$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N} f(t)\frac{1}{2\pi\sqrt{N}}\sum_{s=1}^{N} Y_s e^{-i(t-s)\lambda_j} - \frac{1}{\sqrt{N}}\sum_{t=1}^{N} f(t)\frac{8}{6\pi\sqrt{N}}\sum_{s=1}^{N} Y_s w_s w_t e^{-i(t-s)\lambda_j} \xrightarrow{d} h(t)(\xi - \tilde{\xi}).$$

Because $h(t)$ was a smooth function the mixed terms converge to zero in probability.

Therefore, altogether the difference of the periodograms is a non-zero function and hence the difference of the estimates is non-zero because the stochastic behaviour of the estimates is determined by the behaviour of the periodograms in this behalf. This proves the theorem. $\diamondsuit$

Remarks:(1) We restrict ourselves in this paper to the case of stationary long memory because of simplicity of the presentation. But all these results do hold as well for the case of non-stationary long memory. However, at the end the question of interest is to distinguish a stationary time series from a non-stationary series misspecified as being stationary.

(2) The idea of this test is related to Hausman tests introduced by Hausman (1978). The idea there is to compare two estimators which behave similar under the null hypothesis but one of them behaves badly under the alternative. Our situation is slightly different by having two estimators with similar properties under the null but both behaving badly in alternative situations. In our case both estimators go under the alternative in different directions and thus the alternative situation can be distinguished from the null.

Now the test statistic is defined as

$$T := m^{1/2}(\hat{d} - \hat{d}_T)^2. \tag{8}$$

From Theorem 1 and 2 we obtain that the test (6) can distinguish major deterministic trends and long-range dependence. But the limiting distribution depends on the true memory parameter of the noise process. Thus, the limiting distribution of the test statistic should be estimated from the data by using bootstrap. This is done in the next section.

4 Estimating the limiting distribution

In the last section we developed a test statistic which is able to distinguish long-range dependencies and major trends. In this section we discuss the limiting distribution of this test under the null of no major trend.

In the case of a normally distributed data generating process both estimators are normally distributed. The variance of the tapered estimator is greater than that of the standard GPH-estimator and its variance depends on the chosen taper. For the full cosine bell taper used in this paper the variance can be three times as big as for the non-tapered estimator. For a detailed discussion of the limiting distribution of each of these estimators we refer to Robinson(1995) and Velasco(1999).

But from this discussion we can see that our test statistic is asymptotically χ_1^2 distributed after standardization whenever the data generating process is Gaussian. It should be mentioned that normality of the data generating process is not needed for the following discussion and that the test statistic remains of use if this is not the case. Anyway, for the case of a Gaussian process we can state the following theorem:

Theorem 3 *Under the null of no major trend and if the error process Y_t is Gaussian, the test statistic T is asymptotically χ_1^2 distributed with one degree of freedom after standardizing with the standard deviation depending on the memory parameter d_0 of the process Y_t.*

Proof: The test statistic T is given in (8) by

$$T := m^{1/2}(\hat{d} - \hat{d}_T)^2.$$

Asymptotic normality of $m^{1/2}(\hat{d} - \hat{d}_T)$ follows directly from the asymptotic normality of each of these estimators. The mean zero follows because of the consistency of both estimators and the discussion above. It remains the variance. Because both estimators are not independent the asymptotic variance is given by

$$\mathsf{Var}(T) = \mathsf{Var}(\hat{d}) + \mathsf{Var}(\hat{d}_T) - 2\mathsf{Cov}(\hat{d}, \hat{d}_T).$$

Here $\mathsf{Var}()$ denotes the asymptotic variances of each term. But of course the covariance of both estimators depends on d_0. Thus, after standardization the statistic itself is χ^2_1 distributed. $\diamondsuit$

Therefore, from the covariance term it turns out that the variance depends on the true memory parameter of the data generating process. This makes it impossible to compute critical values direct from the asymptotic distribution without knowledge of the true memory parameter what is the problem under test. We renounce computing the exact form of the variance term what is rather complicated and does not support the goal of this paper.

Instead of this we estimate the asymptotic distribution of the test by employing bootstrap methods. Because of this step the stated Gaussianity of the test statistic is not crucial for us. The test is still applicable even if the data is not normal because the asymptotic distribution and thus critical values for the test statistic have to be estimated in any way.

Bootstrap is a resampling technique which allows the estimation of an estimator or test statistic depending asymptotically on an unknown parameter. For a detailed discussion about the bootstrap and its applications in econometrics see Horowitz(2000) or Davidson(2002).

The problem in our situation is that bootstrap techniques apply only for independent data. But if $d_0 > 0$ this is not the case here. We have strongly dependent data. The bootstrapping idea for this data is based on the ARFIMA-representation (3) of a long-memory process. It says that differencing the process appropriately results in a white noise process which then can be bootstrapped.

The idea is as follows. We differentiate the data with the smaller of both of the estimated memory parameters obtained from the standard and the tapered GPH-estimation because this is the less biased estimator. The resulting process is being bootstrapped and these bootstrap samples are integrated again with the estimated memory parameter. For the so generated data we compute the test statistic. Repeating this procedure M times estimates the empirical distribution function of our test statistic. From this empirical distribution function p-values for the true value of the test statistic using the original data can be computed.

5 Monte Carlo Results

In the Monte Carlo study we focus on the logistic regression (6) and on a sinus trend. The logistic regression function is a useful way for modelling shifts in the regime. Depending on the choice of parameters rapid breaks can be modelled as well as smooth changes. For this reason the logistic regression function is very popular in economic modelling. The sinus trend simulates a periodic behaviour as it occurs in seasonal data. It is considered to show that the test can deal also with those structures. That is why we concentrate our studies in this paper on these functions.

The actual simulations are in each case based on $N = 1000$ repetitions. The actual distribution of the test statistic is in each case estimated by $M = 1000$ Bootstrap replications.

Let us first consider the logistic regression function. We compute the power of the test statistic for various parameter choices concerning the memory parameter as well as the time of the break and the smoothness. The parameter a and b describing the value of the function before and after the break and thus the size of the break are fixed for the whole study with $a = 0$ and $b = 1.5$. The noise process used is a Gaussian $ARFIMA(0, d, 0)$-process on which the logistic regression function is added.

Table 1 gives the power of the test for a rather smooth changeover of the regimes by choosing the parameter $\gamma = 10$ as well as for a classical structural break in the mean by choosing $\gamma = 1000$. The power is computed for various memory parameters. We consider the case of an independent noise process meaning $d = 0$, of a memory parameter in the middle of the stationary long-memory range by $d = 0.2$ and we consider strong long memory for $d = 0.4$. The actual break point is located at $10, 20, 50, 80, 90\%$ of the data given by a value of $c = 0.1, 0.2, 0.5, 0.8, 0.9$. Thus, we consider breaks as well at the beginning of the observation period as in the middle and the end of the data.

Table I *Power of the test for the logistic regression*

	$d = 0$		$d = 0.2$		$d = 0.4$	
	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10$	$\gamma = 1000$
c = 0.1	0.813	0.971	0.93	1	0.997	1
c = 0.2	0.919	0.864	0.989	0.9	1	0.974
c = 0.5	0.971	0.964	0.98	0.993	1	1
c = 0.8	0.986	0.994	1	1	1	1
c = 0.9	0.976	1	1	1	1	1

As it can be seen from the table we have a good power of mostly above 90% for all values of d. The power increases with d but is still high for $d = 0$. There are also no differences between a rather smooth change in mean ($\gamma = 10$) and an abrupt structural break ($\gamma = 1000$). The power of the test does also not depend on the break time. We obtain a high power for breaks at the beginning of the observation period ($c = 0.1$) as well as in the middle ($c = 0.5$) and at the end ($c = 0.9$).

Let us now consider the sinus trend given by

$$f(t) = \frac{\sin(t)}{t}.$$

This trend fulfills the conditions T1 and T2. The idea of considering this sinus trend is to simulate a periodic behaviour as it occurs for example in seasonal or cyclical data. We consider $\sin(t)/t$ rather than $\sin(t)$ directly to have a decaying trend which makes it even harder to distinguish the trend from long-range dependence. The power for this trend is given in table II.

Table II *Power of the test for* $\sin(t)/t$.

d	Power
d = 0	0.945
d = 0.2	1
d = 0.4	1

Again we observe a very good power for all values of d. This shows that the test behaves well not only for structural breaks but also for smooth decaying functions. It is able to detect also seasonal effects.

Table III shows that the test keeps its levels. The level of the test is almost reached in all cases. Still the test is rather conservative. This emphasizes the good properties of the test.

Table III *Level of the test.*

d	*Level*		
	0.01	0.05	0.1
d = 0	0.012	0.05	0.083
d = 0.2	0.005	0.049	0.085
d = 0.4	0.005	0.045	0.088

It is worth mentioning at this place that although the distribution of the test statistic depends on the true memory parameter d_0 simulated critical values for $d = 0, 0.2$ and 0.4 have been very close to each other.

6 Application

There is evidence of long memory in several economic data sets as volatilities of stock returns or inflation rates. Sibbertsen(2002b) showed in an empirical study by employing the above ideas that there is strong evidence of long memory in the volatilities of the returns of seven German stocks.

Here we re-analyze the long-memory behaviour of inflation rates. There is an intensive discussion whether inflation rates contain a unit root. In recent years several empirical analysis found evidence of long memory in inflation rates. For an overview of this discussion see Baillie(1996). Bos et al.(1999) discussed whether these findings are due to level shifts. They modelled exogenous level shifts during the oil price crises. They compared models with no, two and four exogenous shifts for the G7 countries. For testing of significance of the shift they employed the LM- and Wald-test. They found that the estimated memory parameter clearly reduces in all cases when level shifts are introduced to the model. Still this can not be seen as a proof because tests on structural breaks do misspecify long memory as structural breaks (Krämer/Sibbertsen, 2002). But we still support the thesis of level shifts in inflation rates by applying the test (8) to the inflation rates of three industrialized countries, namely the US, UK and Germany. Using the monthly consumer price index (CPI) for all of these countries from January 1957 to March 2002[2] we obtain the inflation rate I_t at time t for each country by $I_t = \log(CPI_t/CPI_{t-1})$. Thus, we have 543 observations for each country meaning that we use $m = N^{0.8} = 154$ frequencies for the estimation of the memory parameter. The results of the test is given in table IV.

Table IV *Test results for three inflation rates*

	$\hat{d}$	$\hat{d}_T$	$\sqrt{154}(\hat{d} - \hat{d}_T)^2$
US	0.39	0.53	0.246
UK	0.344	0.442	0.118
Germany	0.17	0.31	0.238

[2]Data obtained from Datastream.

Thus, the null of no major trend can be rejected for the US and Germany at the 95% and even the 99% level. For the UK the hypothesis still can be rejected to the 90% level whereas it cannot be rejected to the 95% level.

These results clearly reject the hypothesis of no trends or structural breaks in inflation rates and support the thesis that to some extent the long-range dependence effects are the results of misspecification of major trends such as level shifts as long memory. Our estimation results reproduce the previous empirical findings of long-range dependence by using log-periodogram based techniques. The smaller of both of our estimators gives almost those results Bos et al. (1999) estimated after introducing four level shifts. Only for the UK we obtain a slightly higher value.

Comparing the standard GPH-estimator and its tapered counterpart show that there is evidence of structural breaks in inflation rates. Whether there are also some long memory effects present in inflation rates has to be considered in future research.

7 Conclusion

In this paper we constructed a test for distinguishing long-range dependence and major trends such as structural breaks. The idea of the test is to compare the standard GPH-estimator with the tapered GPH-estimator. Both estimators are consistent under the null of no major trend but behave different under the alternative of major trends or structural breaks. This idea is similar to Hausman tests. It is proven that this test can distinguish long memory and major trends. The asymptotic distribution of the test statistic which depends on the memory parameter of the underlying noise process is estimated by using bootstrap. It turns out that the test performs well by having a high power. It is also shown that the test behaves well for different types of trends by considering structural breaks with the logistic regression function as well as periodic trends by considering a sinus trend. In the last section the test is applied to inflation rates of the US, UK and Germany. By rejecting the null of no major trend to the 99% level for the US and Germany and to the 90% level for the

UK the test clearly rejects the empirical findings of long-range dependence in inflation rates and supports the hypothesis that these findings are caused by level shifts in the data.

Acknowledgements

Research was done during a visit of the first author at Cardiff Business School, Cardiff University. Research supported by Volkswagenstiftung and Deutsche Forschungsgemeinschaft under SFB 475. The computational assistance of Klaus Nordhausen is gratefully acknowledged.

References

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

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **74**, 3 - 30.

Beran, J. (1994): *Statistics for long-memory processes.* London: Chapman & Hall.

Bollerslev, T., Mikkelsen, H. O. (1996): "Modeling and pricing long memory in stock market volatility." *Journal of Econometrics* **73**, 151 - 184.

Bos, C. S., Franses, P. H., Ooms, M. (1999): "Long memory and level shifts: Re-analyzing inflation rates." *Empirical Economics* **24**, 427 – 449.

Davidson, J. (2002): "A model of fractional cointegration, and tests for cointegration using the bootstrap." *Journal of Econometrics* **110**, 187 – 212.

Davidson, J., Sibbertsen, P. (2002): "Generating schemes of long-memory processes." *Technical Report 46/2002*, SFB 475, University of Dortmund.

Ding, Z., Engle, R.F. and Granger, C.W.J. (1993): "A long memory property of stock market returns and a new model." *Journal of Empirical Finance* **1**, 83 – 106.

Geweke, J., Porter-Hudak, S. (1983): "The estimation and application of long-memory time series models." *Journal of Time Series Analysis* **4**, 221 – 237.

Giraitis, L., Kokoszka, P. and Leipus, R. (2001): "Testing for long memory in the presence of a general trend." *Journal of Applied Probability* **38**, 1033 - 1054.

Hausman, J. A. (1978): "Specification Tests in Econometrics." *Econometrica* **46**, 1251 – 1271.

Hurst, H. E. (1951): "Long-term Storage of Capacity of Reservoirs." *Transactions of the American Society of Civil Engineers* **116**, 770 - 799.

Hurvich, C., Deo, R. and Brodsky, J. (1998): "The Mean Squared Error of Geweke and Porter-Hudak's estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **19**, 19 – 46.

Hurvich, C., Deo, R. (1999): "Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long-memory time series." *Journal of Time Series Analysis* **20**, 331 –341.

Hurvich, C., Ray, B. (1995): "Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes." *Journal of Time Series Analysis* **16**, 17 – 42.

Krämer, W., Sibbertsen, P. (2002): "Testing for structural change in the presence of long memory." *International Journal of Business and Economics* **1**, 235 - 242.

Krämer, W., Sibbertsen, P. and Kleiber, C. (2002): "Long memory versus Structural Change in Financial Time Series." *Allgemeines Statistisches Archiv* **86**, 83 - 96.

Lobato, I.N. and Savin, N.E. (1998): "Real and spurious long-memory properties of stock market data (with discussion and reply)." *Journal of Business and Economic Statistics* 16, 261 – 283.

Lohre, M. and Sibbertsen, P. (2001): "Persistenz und saisonale Abhängigkeiten in Abflüssen des Rheins." *Hydrology and Water Resources Management* **46**, 166 - 174.

Sibbertsen, P. (1999): *Robuste Parameterschätzung im linearen Regressionsmodell.* VWF, Berlin.

Sibbertsen, P. (2002a): "Long memory versus Structural Breaks: An overview." *forthcoming in Statistical Papers.*

Sibbertsen, P. (2002b): "Long memory in the volatilities of German stock returns." *forthcoming in Empirical Economics.*

Sibbertsen, P. (2003): "Log-periodogram estimation of the memory parameter of a long-memory process under trend." *Statistics and Probability Letters* **61**, 261 - 268.

Velasco, C. (1999): "Non-Stationary Log-Periodogram Regression." *Journal of Econometrics* **91**, 325 – 371.

Willinger, W., Taqqu, M. S. and Teverovsky, V. (1999): "Stock market prices and long-range dependence." *Finance and Stochastics* **3**, 1 - 13.