Long Memory and Structural Change: New Evidence from German Stock Market Returns

Uwe Hassler and Maya Olivares

Goethe University Frankfurt

Preliminary version: December 17, 2007

Abstract

In this paper we analyze 11810 daily absolute returns of the German stock price index DAX. A modified local Whittle estimator allowing for mean shifts when estimating the long memory is applied. Our results only weakly support a significant break in the mean. We rather observe a change in persistence before and after October 1987. We discuss and apply a simple test for equality of the memory (no break in persistence) that builds on the difference between subsample estimators.

Keywords: absolute returns; break in mean and/or persistence; semiparametric estimation; testing significant changes

JEL classification: C14 (Semiparametric and Nonparametric Methods); C22 (Time-Series Models); G14 (Information and Market Efficiency; Event Studies)
1 Introduction

Absolute returns are considered as central to volatility forecasting, see Forsberg and Ghysels (2007) for a recent explanation. While stock market returns are accepted to be of short memory (see e.g. Lo (1991)), Ding et al. (1993) find considerable persistence in power transformations of absolute returns, and in particular in the absolute returns themselves, see also Granger and Ding (1996). Many authors have argued that the feature of long memory may be spurious and caused by breaks or regime shifts, see Banerjee and Urga (2005) for a recent overview. In particular, Lobato and Savin (1998) raise the question whether the long memory in volatility is due to deterministic shifts. If forecasting is the aim of study, the question whether the persistence arises from eventual shifts or not is of crucial importance.

In this paper we add three aspects to this literature. First, we analyze daily observations (absolute returns) of the German stock price index DAX until March 2007, thus extending a related study by Krämer et al. (2002). Second, we apply the modified local Whittle estimator by Hsu (2005) allowing for mean shifts while determining the persistence parameter. Simultaneously a break point is determined from the data (October 1987) and used to test for significance of this break applying the procedure by Hidalgo and Robinson (1996). Our results contrast Krämer et al. (2002) in that we do not find a significant break in the mean but rather long memory. Third, the sample is split and the memory parameters are estimated separately before and after October 1987. We discuss in line with Shimotsu (2006) a simple test for equality of the memory (no break in persistence) that builds on the difference between subsample estimators. A similar approach has recently been used ad hoc by Kumar and Okimoto (2007). Our results are highly indicative of a significant increase in persistence after October 1987.

The rest of the paper is organized as follows. The next section very briefly reviews the fractionally integrated model of long memory. Section 3 describes the (modified) local Whittle estimation and the way how we implement the procedure. The fourth section addresses the issue of testing for significance of
breaks and introduces in particular the new proposal. Section 5 is dedicated to the empirical application, while concluding remarks are collected in the last section.

2 Long memory

The most widely used model in econometrics to capture long memory is fractional integration. The fractionally integrated process $y_t$ is defined as

$$ (1 - L)^d(y_t - \mu) = x_t, \quad t = 1, \ldots, T, \quad |d| < 0.5, $$

where $x_t$ is a purely stochastic process without deterministic components, $L$ is the lag operator and the fractional differences $(1-L)^d$ are given by binomial expansion. If $x_t$ is a stationary and invertible autoregressive moving-average [ARMA] process, then $y_t$ is called an ARFIMA process, fractionally integrated of order $d$. The process is stationary as long as $d < 0.5$, but displays long memory for $d > 0$. Long memory implies a form of serial dependence and persistence that cannot be captured by traditional ARMA processes. The autocovariance function $\gamma_y(h)$ of a fractionally integrated process behaves as follows with lag $h$ being large:

$$ \gamma_y(h) \sim C h^{2d-1}, \quad \text{with} \quad h \to \infty. $$

We observe that $\gamma_y(h) \to 0$ as long as $d < 0.5$, but for $d > 0$ the rate of convergence is so slow that serial correlation coefficients are not summable:

$$ \sum_{h=0}^{H} \rho_y(h) \to \infty \quad \text{for} \quad H \to \infty \quad \text{as} \quad d > 0. $$

This long memory or persistence property translates into the frequency domain as follows. Define as spectral density of $y_t$:

$$ f_y(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_y(h) \cos(h\lambda), \quad \lambda \in [0, \pi]. $$
In case of fractional integration it holds around the origin:

\[ f_y(\lambda) \sim G \lambda^{-2d} \quad \text{as } \lambda \to 0, \tag{3} \]

where \( G \) is a positive finite constant\(^1\). If \( d = 0 \) there is no long memory, while \( d > 0 \) models a pole at the origin measuring the degree of persistence.

For further aspects of fractional integration we recommend the survey article by Baillie (1996).

3 Semiparametric memory estimation and testing

We are interested in estimating the memory parameter \( d \) that characterizes the spectral density \( f_y(\cdot) \). The considered estimators will be semiparametric in that they do not require the specification or estimation of a model for the short-run component \( x_t \).

3.1 Local Whittle [LW] estimation

The empirical counterpart to the spectral density is the periodogram \( I_y(\lambda_j) \) of \( \{y_t\}_{t=1,...,T} \), evaluated at the harmonic frequencies \( \lambda_j = \frac{2\pi j}{T} \),

\[
I_y(\lambda_j) = \left| \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} y_t \exp\{i\lambda_j t\} \right|^2, \quad i^2 = -1.
\]

Whittle estimation relies on a frequency domain approximation of the log-likelihood function and amounts to minimizing

\[
R(d) = G(d) - \frac{2d}{m} \sum_{j=1}^{m} \log(\lambda_j), \tag{4}
\]

with

\[
G(d) = \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I_y(\lambda_j). \tag{5}
\]

\(^1\)We may think of \( G = f_x(0) \) as the spectral density at frequency zero of a short-memory component \( x_t \).
Künsch (1987) proposed the local Whittle [LW] estimator, where summation is not over the whole frequency range. Instead we assume that $m$ diverges more slowly than the sample size $T$:

$$\frac{1}{m} + \frac{m}{T} \to 0, \quad T \to \infty.$$  

(6)

Now, the LW estimator of $d$ is defined as

$$\hat{d} = \arg \min R(d).$$

Robinson (1995) proves that under certain further assumptions

$$\mathcal{R}_d = 2 \sqrt{m} (\hat{d} - d) \overset{d}{\to} \mathcal{N}(0, 1),$$

(7)

for $d \in (-\frac{1}{2}, \frac{1}{2})$, where "$\overset{d}{\to}$" denotes convergence in distribution.

The LW estimator is often recommended in practice because it is more efficient than semiparametric competitors. Further, it is robust with respect to heteroskedasticity of a certain degree, see Robinson and Henry (1999) and Shao and Wu (2007).

### 3.2 Data-driven choice of $m$

In a semiparametric environment, the consistency of the estimated long-run behaviour depends crucially on the chosen bandwidth $m$. Choosing a very large (very small) bandwidth value is known to lead to a large bias (large variance) in the estimation. Hence, the optimal choice of the number of harmonic frequencies under consideration of (6) is essential for the estimation outcome.

Although the number of periodogram harmonics are often chosen deterministically, Henry (2001) proposes an approximate optimal-bandwidth-formula which is based on the minimization of the estimate’s mean squared error and is iterated until the convergence to an optimal bandwidth value is achieved:

$$\hat{\theta}^{(k)} = \arg \min R(\hat{m}^{(k)}, d)$$

$$\hat{m}^{(k+1)} = \left( \frac{3}{4\pi} \right)^{4/5} \left| \theta^* + \frac{\hat{d}^{(k)}}{12} \right|^{-2/5} T^{4/5},$$

(8)
for $-1/2 < d < 1/2$ and initial value $\hat{m}(0) = T^{4/5}$. Note that here the dependence on $m$ of the criterion function from (4) is spelled out explicitly: $R(d) = R(m, d)$.

In the case of an ARFIMA model, where $f_x(\lambda)$ is the spectrum of an ARMA process, the parameter $\theta^*$ can be approximated by $\theta^* = f''_x(0)/2 f_x(0)$, as Delgado and Robinson (1996) propose\footnote{Assuming that $f_x(\lambda)$ has first derivative $f'_x(0) = 0$ and second derivative $f''_x(0)$.}. The authors argue that a feasible approximation for this parameter is based on following least squares regression of the periodogram $I(\lambda_j)$ on the regressors $Z_{jk}(\hat{d}(0))$, $k = 0, 1, 2$:

$$I(\lambda_j) = \sum_{k=0}^{2} Z_{jk}(\hat{d}(0)) \bar{\varphi}_k + \bar{\varepsilon}_j, j = 1, \ldots, \hat{m}(0),$$

where $Z_{jk}(d) = |1 - \exp^{i\lambda_j}|^{-2d} \lambda_j^k/k!$. Thus, the regression to be conducted is

$$I(\lambda_j) = |1 - \exp^{i\lambda_j}|^{-2\hat{d}(0)} \left( \bar{\varphi}_0 + \bar{\varphi}_1 \lambda_j + \frac{\bar{\varphi}_2 \lambda_j^2}{2} \right) + \bar{\varepsilon}_j.$$

The estimates of $f_x(0)$ and $f''_x(0)$ are $\bar{\varphi}_0$ and $\bar{\varphi}_2$, respectively, so that the estimated $\hat{\theta}^*$ replacing $\theta^*$ in (8) is given by

$$\hat{\theta}^* = \frac{\bar{\varphi}_2}{2 \bar{\varphi}_0}. \tag{9}$$

### 3.3 Hsu’s modification

Hsu (2005) modifies the LW estimator allowing for a mean shift:

$$y_t = \begin{cases} \eta_t + \mu_1 & t \leq \lfloor T \tau \rfloor \\ \eta_t + \mu_2 & t > \lfloor T \tau \rfloor \end{cases},$$

where $\lfloor \cdot \rfloor$ stands for the integer part, and $\tau \in [0, 1]$ denotes the break fraction. The observed series is demeaned accordingly,

$$\hat{\eta}_t(\tau) = \begin{cases} y_t - \hat{\mu}_1(\tau) & t \leq \lfloor T \tau \rfloor \\ y_t - \hat{\mu}_2(\tau) & t > \lfloor T \tau \rfloor \end{cases},$$
with 
\[ \hat{\mu}_1(\tau) = \frac{1}{[T\tau]} \sum_{t=1}^{[T\tau]} y_t \quad \text{and} \quad \hat{\mu}_2(\tau) = \frac{1}{T-[T\tau]} \sum_{t=[T\tau]+1}^{T} y_t. \]

Let now \( I_\eta(\lambda_j, \tau) \) represent the periodogram computed from \( \hat{\eta}_t(\tau) \) instead of \( y_t \). The modified criterion function is
\[ R(d, \tau) = G(d, \tau) - \frac{2d}{m} \sum_{j=1}^{m} \log(\lambda_j) \tag{10} \]
with
\[ G(d, \tau) = \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I_\eta(\lambda_j, \tau). \tag{11} \]

In practice, one is rarely willing to assume a known break point \([\tau T]\). Hence, the conditional LW estimator \( \hat{d}(\tau) \) is obtained for given \( \tau \) in a first minimization, while a second one is necessary to find the change-point estimator \( \hat{\tau} \):
\[ \hat{\tau} = \arg \min_{\tau \in [\underline{\tau}, \overline{\tau}]} R(\hat{d}(\tau), \tau), \]
where \([\underline{\tau}, \overline{\tau}] \subset (0, 1)\). Consequently, the new estimator for the memory parameter \( d \) is found by localizing the point in \( \hat{d}(\tau) \) at which \( R(\hat{d}(\tau), \tau) \) achieves its minimum, i.e., \( \hat{d}(\hat{\tau}) \). Since the estimator \( \hat{\tau} \) converges to the true normalized change point \( \tau \), Hsu (2005) argues that the limiting normality is not affected:
\[ \mathcal{H}_d = 2\sqrt{m} (\hat{d}(\hat{\tau}) - d) \overset{d}{\rightarrow} N(0, 1). \tag{12} \]

### 3.4 Testing against a break in the mean
Significance testing against breaks in mean in a long memory environment is an unsolved problem as long as the break point is not known a priori. Krämer and Sibbertsen (2002) prove for CUSUM tests (standard as well as OLS-based) that both the rates of convergence and the limiting distributions depend on the memory parameter \( d \), see also Wright (1998). Working with the tabulated critical values for \( d = 0 \) will result in spurious detection of a
break in mean. A similar result holds for the least squares procedure by Bai (1994). Kuan and Hsu (1998) proved that the break point is only $T^{0.5-d}$ consistent for general I($d$) processes, and the limiting distribution depending on $d$ was derived by Horváth and Kokoszka (1997), compare also Sibbertsen (2004).

If the break fraction is known the situation is different. Hidalgo and Robinson (1996, eq. (11)) propose the following test statistic:

$$\mathcal{HR} = T^{0.5-d} \frac{\hat{\mu}_1(\tau) - \hat{\mu}_2(\tau)}{\sqrt{\Omega}},$$

where

$$\Omega = \frac{C}{d(2d+1)} \frac{\tau^{2d} + (1-\tau)^{2d} - 1}{\tau(1-\tau)}$$

(14)

with $C$ from (2). Under the null hypothesis of no break in mean,

$$H_0 : \mu_1 = \mu_2,$$

they prove a limiting standard normal distribution. In practice, $d$ and $C$ have to be estimated consistently. This can be achieved through the LW estimator, since Hidalgo and Robinson (1996) relate $C$ to $G$ from (3) as follows (where $\Gamma(\cdot)$ denotes the Gamma function):

$$C = G^2 \Gamma(1-2d) \cos((0.5-d)\pi).$$

Having a consistent estimator of $d$, a consistent estimator of $G$ is readily available from (5) or (11), $\hat{G}(\hat{d})$ and $\hat{G}(\hat{d}, \hat{\tau})$, which allows for consistent estimation of $C$ and ultimately $\Omega$.

Note from (13) that the rate of convergence to limiting normality decreases as $d$ approaches 0.5. Still, the test should be applicable in our context, where $T$ is very large.

4 Testing against a break in $d$

There is a growing literature on tests against breaks in persistence that are embedded in the integer paradigm of changes from I(0) to I(1) or the other
way round. The first contributions were by Kim (2000), Kim et al. (2002) and Busetti and Taylor (2004). The power of these tests when breaking from \( I(0) \) to \( I(d) \) has recently been studied by Hassler and Scheithauer (2007). Further, Hassler and Scheithauer (2007, Prop. 2) prove that the limiting distribution under no break in \( d \) depends on \( d \), so that critical values are not readily available.

However, a simple procedure based on subsample estimation can be straightforwardly used to detect a possible change in \( d \). We split the sample into non-overlapping subsamples, \( t \in \{1, \cdots, \lceil \tau T \rceil \} \) and \( t \in \{ \lceil \tau T \rceil + 1, \cdots, T \} \). For given break fraction \( \tau \) we compute LW estimators from both subsamples (\( \hat{d}_1(\tau) \) and \( \hat{d}_2(\tau) \)), and follow Kumar and Okimoto (2007) in that we test the difference against significant deviations from zero. Let \( m_1 \) and \( m_2 \) denote the bandwidths from the respective subsamples. Following the lines of Shimotsu (2006, proof of Lemma 1), joint normality as well as asymptotic independence\(^3\) of the estimators can be derived, so that

\[
\sqrt{m_i} (\hat{d}_i(\tau) - d_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{4}\right), \quad i = 1, 2. 
\]

Hence, we approximate under \( H_0 \):

\[
\frac{\hat{d}_1(\tau) - \hat{d}_2(\tau)}{\sqrt{\text{Var}\left(\hat{d}_1(\tau) - \hat{d}_2(\tau)\right)}} \xrightarrow{d} \mathcal{N}(0,1),
\]

where the asymptotic variance is

\[
\text{Var}\left(\hat{d}_1(\tau) - \hat{d}_2(\tau)\right) = \frac{1}{4m_1} + \frac{1}{4m_2} = \frac{m_1 + m_2}{4m_1m_2}.
\]

Consequently, the asymptotically valid test \( D(\tau) \) is defined as

\[
D(\tau) = 2 \cdot \frac{\sqrt{m_1m_2}}{\sqrt{m_1 + m_2}} \left(\hat{d}_1(\tau) - \hat{d}_2(\tau)\right) \xrightarrow{d} \mathcal{N}(0,1).
\]

\(^3\)This improves the upper bound argument in Kumar and Okimoto (2007) who assumed in their footnote 18 that \( \text{Cov}(\hat{d}_1, \hat{d}_2) > 0 \), such that \( \text{Var}(\hat{d}_1 - \hat{d}_2) \leq \text{Var}(\hat{d}_1) + \text{Var}(\hat{d}_2) \).
In particular, Shimotsu (2006) proposes bandwidths

\[ m_1 = \tau m \quad \text{and} \quad m_2 = (1 - \tau) m, \]

where \( m \) satisfies the usual requirements, such that

\[
\text{Var} \left( \hat{d}_1(\tau) - \hat{d}_2(\tau) \right) = \frac{1}{4 m \tau (1 - \tau)}.
\]

(16)

From a practical point of view the break fraction \( \tau \) in (15) has to be chosen somehow. Firstly, one may fix \( \tau \) a priori as in Shimotsu (2006) (e.g. \( \tau = 0.5 \)) without attaching a break-point meaning. In other words, an arbitrary point where the sample is split may be selected. This is clearly suboptimal if a break indeed occurs under the alternative hypothesis, where an adequate split-point would be desirable to increase power. Secondly, one may estimate \( \tau \) from the data, \( \hat{\tau} \). However, under the null hypothesis, where no break exists, \( \hat{\tau} \) will not converge to a constant but rather to a random variable, which, in turn, will affect the limiting normality in (15).

5 Empirical results

The presence of long memory components in asset returns has been the subject of several studies, considering the important implications that persistent autocorrelation structures can have on stock returns. Among the most important works analyzing this statistical issue are Greene and Fielitz (1977), Lo (1991) and Ding et al. (1993). While the first authors find significant evidence of long-run dependence using the rescaled range statistic (or “R/S” statistic) on the returns of 200 common stocks listed in the New York Stock Exchange, Lo (1991) proposes a modified R/S statistic to account for short-term dependence and finds in contrast to Greene and Fielitz (1977) little evidence of long memory in U.S. stock market returns. Ding et al. (1993) investigate the long-term components of the daily returns on the Standard & Poor’s 500 price index\(^4\) (S&P 500) and find strong evidence of long memory

\(^4\)From January 3, 1928 to August 30, 1991.
in the squared and absolute returns. By using a semiparametric procedure that is robust to the presence of short-term dependence, Lobato and Savin (1998) find evidence in favor of long memory for squared daily returns of the S&P 500, confirming thus the results of Ding et al. (1993), and they argue that this evidence is not spurious. Following that line, the semiparametric LW estimation is applied to German data, where Krämer et al. (2002) did not find (much) long memory in squared returns.

5.1 The data

The data to be analyzed are the daily returns of the German stock price index DAX, which comprises the 30 largest German companies in terms of traded volume and market capitalization. The data set is composed of 11,811 observations from January 4, 1960 to March 30, 2007. If the price index is denoted \( p_t \), then the returns \( r_t \) are defined as

\[
r_t = \ln \left( \frac{p_t}{p_{t-1}} \right),
\]

for \( t = 1, \ldots, 11810 \). Figures 1, 2 and 3 present the plots of \( p_t, r_t \) and \( |r_t| \). The price index presents an upward trend until its substantial fall in September 2001 caused by the terrorist attack against the World Trade Center in New York. Being down by around 17% on September 18, the DAX only begins to recover in March 2003. In contrast, a sudden drop in prices is known to have occurred in the Black Monday’s stock market crash of 1987, where the DAX Index fell by approximately 9.9%, as well as on the “Friday the 13th mini-crash” in October 1989, where the German price index lost 13.7% of its

---

5Its calculation starts at 9.00 a.m. and ends with the prices from the closing auction of the electronic trading system Xetra at 5.30 p.m. The index was obtained from the Deutsche Börse AG and www.Markt-Daten.de.

6Black Monday is the name given to Monday, October 19, 1987, when the Dow Jones Industrial Average Index fell dramatically by 22.6%, and on which similar enormous drops occurred across the world. The Black Monday decline was the second largest one-day percentage decline in stock market history.
value in one day\footnote{The “mini-crash” was apparently caused by a reaction to a news story of a US$6.75 billion leveraged buyout deal for UAL Corporation, the parent company of United Airlines, which did not materialize.}. The high market volatility caused by the September 11 attacks can be clearly distinguished in Figures 1 and 2, whereas the sudden fall of the DAX index in 1987 and 1989 is accentuated by the large absolute returns observed in Figure 3.

![Figure 1. DAX daily price index $p_t$, from January 1960 to March 2007.](image_url)

### 5.2 Long memory in absolute returns

It is a well established fact that, according to the efficient market hypothesis, stock market returns are uncorrelated and exhibit no memory at all. However, as Ding \textit{et al.} (1993) note, this does not imply that they are independently identically distributed, as the theory assumes. In fact, Fama (1970) observes that many return series display a positive (though very small) first order autocorrelation, and concludes that the random walk theory does not
hold strictly for prices. In spite of this fact, Ding et al. (1993) find no evidence of long-range dependence in the returns. Contrary to $r_t$, however, they find persistent positive serial correlation in the squared values of the returns, and observe that the autocorrelations are even greater for $|r_t|$. The respective autocorrelations of the German price index present the same pattern, and are plotted in Figure 4.

As expected, a positive first order autocorrelation can be observed in the returns, as well as in its transformations. This confirms that $r_t$ cannot be an iid process, although the remaining lags of the autocorrelation function seem to be predominantly inside the 95% confidence interval. In contrast to $r_t$, the transformed returns $|r_t|$ and $r_t^2$ are all outside the confidence interval, and it is clear that the sample autocorrelations for absolute returns are greater than for squared returns. As evidence of long memory is much stronger for $|r_t|$ and Granger and Ding (1996) find that absolute returns have the properties of an $I(d)$ process with $\hat{d} = 0.474$ using stock market data, the subsequent analysis of long memory will also focus on $|r_t|$.
Table 1 reports the results of the Whittle estimator for absolute returns of the DAX between January 1960 and March 2007. The long memory parameter $d$ is estimated using a grid of values of $m$ from $m = 299$ to $m = 599$. The first value $m = 299$ corresponds to the bandwidth obtained applying the iterative (and feasible) procedure defined in (8) and (9). Additionally, $m$ is increased by steps of 100 to check that the estimation outcome remains robust.

<table>
<thead>
<tr>
<th>$m$</th>
<th>299</th>
<th>399</th>
<th>499</th>
<th>599</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>0.4239</td>
<td>0.4272</td>
<td>0.4116</td>
<td>0.4234</td>
</tr>
<tr>
<td>$\mathcal{R}_0$</td>
<td>14.660</td>
<td>17.067</td>
<td>18.389</td>
<td>20.725</td>
</tr>
</tbody>
</table>

Note: LW estimator with $\mathcal{R}_0$ from (7).

In line with the results of Granger and Ding (1996), strong evidence in favor of
long-range dependence is found in the absolute returns. In fact, the estimates vary with $m$ fairly robust around $\hat{d} \approx 0.42$. However, as Lobato and Savin (1998) note, spurious long memory can possibly be caused by nonstationarity, i.e., a change in the mean of absolute returns.

### 5.3 Allowing for a break in mean

As important stock market crises are indeed known to have occurred during the observed time period, the modified LW estimator proposed by Hsu (2005) is applied to estimate $d$, in order to confirm that the evidence in favor of long memory is real. As Table 2 shows, the proposed method locates the break in October 1987 (the month where the Black Monday stock market crash occurred), and, taking into account this possible structural change, the long memory property of $|r_t|$ remains stable and significantly different from
zero (again $\hat{d} \approx 0.42$).

| Table 2. Break in mean of $|r_t|$ at $[\hat{\tau} T]$ |
|---|---|---|---|---|
| $m$ | 299 | 399 | 499 | 599 |
| $\hat{d}(\hat{\tau})$ | 0.4223 | 0.4253 | 0.4085 | 0.4205 |
| $\hat{\tau}$ | 0.5863 | 0.5864 | 0.5864 | 0.5865 |
| $H_0$ | 14.604 | 16.990 | 18.250 | 20.583 |

| Test for no break in mean at $[T \hat{\tau}]$ |
|---|---|---|---|---|
| $HR$ | -1.2030 | -1.1867 | -1.2837 | -1.2130 |
| p-value | 0.2290 | 0.2354 | 0.1992 | 0.2262 |

Note: Modified LW with $H_0$ from (12). $HR$ is defined in (13), two-sided p-values against $\mu_1 = \mu_2$.

It is, however, important to notice that the modification proposed by Hsu (2005) will determine a break in the data irrespective of whether it really took place or not. Thus, although it seems plausible to accept that a break may have occurred in October 1987, the simple comparison of the estimation results obtained applying the LW estimator (Table 1) and its modified version (Table 2) shows that the possible break may not be relevant, given that the estimates of $d$ are quite similar: If a break occurs and $d$ is estimated without taking it into account, $\hat{d}$ usually displays a significant upward bias, which cannot be observed with our data.

In order to check formally the significance of the break date found by Hsu’s method, the test proposed by Hidalgo and Robinson (1996) was applied at this point to check if the no-break hypothesis is indeed rejected for $\hat{\tau}$. The lower panel of Table 2 reports the outcome of the test for the “known” break point $\tau = 0.58$ (which corresponds to October 1987 in the sample). The results confirm that there is no dominant change in mean in the sample,
since the null cannot be rejected at 20% significance (except for marginally with \( m = 499 \)).

5.4 Allowing for a break in \( d \)

Although no significant break in mean can be confirmed at \( \tau = 0.58 \), it is interesting to study why the modified LW procedure produced this particular estimate. Has this happened simply by chance, or is this choice indicative of some true break in the data? To find an acceptable answer to this question, the sample is split into two periods, taking October 1987 as the theoretical break point. The LW estimator is then computed for both periods January 1960-September 1987 and October 1987-March 2007. First, the presence of long memory is again confirmed in both samples, see Table 3. Second, another plausible cause for the findings reported in Table 2 becomes evident: A break in the order of fractional integration. The estimation outcome for the first period varies around \( \hat{d}_1 \approx 0.30 \), while the second sample yields much stronger persistence with \( \hat{d}_2 \approx 0.45 \). The first choice of \( m_1 \) and \( m_2 \) (294 and 169, respectively) was determined again from (8) and (9). Variation of the corresponding bandwidths shows that, during the first period, only moderate long memory is found in the absolute returns, while strong long-range persistence dominates the second one. The results are reported in Table 3.

The empirical approach discussed in Section 4 provides a simple significance test for a change in the order of fractional integration. As expected, the last row of Table 3 shows that a highly significant increase in \( d \) is found after October 1987, which remains significant for all chosen bandwidth values.
Table 3. Split sample

<table>
<thead>
<tr>
<th></th>
<th>294</th>
<th>394</th>
<th>494</th>
<th>594</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_1$</td>
<td>0.3044</td>
<td>0.3045</td>
<td>0.2930</td>
<td>0.2758</td>
</tr>
<tr>
<td>$R_0$</td>
<td>10.439</td>
<td>12.088</td>
<td>13.024</td>
<td>13.444</td>
</tr>
<tr>
<td>$m_2$</td>
<td>169</td>
<td>269</td>
<td>369</td>
<td>469</td>
</tr>
<tr>
<td>October 1987 - March 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_2$</td>
<td>0.4606</td>
<td>0.4875</td>
<td>0.4369</td>
<td>0.4136</td>
</tr>
<tr>
<td>$R_0$</td>
<td>11.976</td>
<td>15.991</td>
<td>16.785</td>
<td>17.914</td>
</tr>
</tbody>
</table>

Significance test for $d_1 = d_2$

| $D(0.58)$ | -3.236 | -4.627 | -4.183 | -4.462 |

Note: LW estimation from subsamples with $R_0$ from (7), $D(0.58)$ is defined in (15).
6 Conclusions
References


