RESEARCH ARTICLE

A New Simple Test Against Spurious Long Memory Using Temporal Aggregation

Heri Kuswanto*

Institut für Statistik, Leibiniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany
* Statistics Department, Institute Technology of Sepuluh Nopember, Indonesia

(Received 00 Month 200x; in final form 00 Month 200x)

We have developed a new test against spurious long memory based on the invariance of long memory parameter to aggregation. By using the local Whittle estimator, the statistic takes the supremum among combinations of paired aggregated series. Simulations show that the test performs good in finite sample sizes, and is able to distinguish long memory from spurious processes with excellent power. Moreover, the empirical application gives further evidence that the observed long memory in German stock returns is spurious.

Keywords: Local-Whittle method, Spurious long memory, Aggregation

AMS Subject Classification: 62E20; 62M10; 91B84

1. Introduction

Let \( x_t \) be a linear long memory process characterized mainly by the following condition

\[
\rho_k \sim C_\rho(k)k^{2d-1}, \quad \text{as } k \to \infty
\]

for \( d \in (0, 0.5) \), where \( \rho_k \) is the autocorrelation function of \( x_t \) and \( C_\rho(k) \) is a slowly varying function. We consider an aggregated long memory process defined as

\[
y_t = \sum_{j=0}^{m-1} x_{mt-j} = \sum_{j=0}^{m-1} B^j x_{mt}
\]

where \( B \) is backshift operator and \( m \) denotes the aggregation level. Chambers [1], Man and Tiao [2] and Souza [3] show that if \( x_t \) satisfies (1) with \( d < 0.5 \), then its aggregation process \( y_t \) also satisfies (1) with the same fractional integration order \( d \). This condition implies invariance of the memory parameter to aggregation.

Spurious long memory can arise in many cases, especially in stock market data. It still has been highly debated whether the observed long memory is real or spurious phenomena. Many studies found long memory in the volatility of stock returns ([4–6] among others). Lobato and Savin [7] and the references therein discuss the
real and spurious long memory properties of stock market data. They investigated
major causes of spurious long memory, such as aggregation, nonstationarity and
regime switching. By using LM type test of Lobato and Robinson [8], they estimated
the memory parameter and tested the significance of the parameter to conclude
whether the observed memory is real or spurious. However, it is well known that
several processes are able to create spurious long memory by generating a certain
degree of fractional integration (see [9–11] among others). Therefore, developing
a test which is able to distinguish long memory from spurious processes is still of
interest, which may lead to a proper model choice.

The fact that the memory parameter does not change with aggregation can be
used as a means to distinguish long memory from spurious processes. Ohanissian,
Russell and Tsay [12] estimate the memory parameter across several aggregation
levels and propose Wald type test to distinguish these two phenomena. They show
that the test is able to detect several spurious processes under alternative with
considerable power. Their results are based on the simulation study by examining
very large numbers of observations, meaning that it has good performance for
high frequency data and our initial study shows that the test loses the power
significantly under small and finite sample sizes. Furthermore, they use the GPH
method of Geweke and Porter-Hudak [13] to estimate the memory parameter and
the theoretical properties of the test have been well investigated. However, Teles
et al. [20] proved that using GPH estimator of aggregated series for testing long
memory has very serious consequences on the power of the test which may lead to
the wrong conclusion, especially by using bandwidth frequency $T^{0.5}$.

In this paper, we propose a new test against spurious long memory based on the
invariance principle, in line with the basic idea of [12]. Our test calculates a value
for every pair of aggregation levels and takes the maximum among the values. This
testing procedure has been previously applied by Beran and Terrin [15] for testing
change in memory parameter. Moreover, we estimate long memory parameter by
semi-parametric local Whittle maximum likelihood instead of by GPH estimator.
This estimation method has been proved to have the smallest bias estimate with
minimum standard deviation [16].

This paper is organized as follows. Section 2 discusses the main result including
the proposed test and its asymptotic distribution. Section 3 presents the results of
simulation study to assess the test performance in finite sample size. The empirical
application, i.e., the case of German stock returns is given in section 4. The proof
is given in the appendix.

2. Main Result

A stationary ARFIMA($p, d, q$) process $x_t$ has the following representation:

$$\phi(B)(1-B)^d x_t = \theta(B) \epsilon_t \quad t = 1, \ldots, N$$

where $B$ is the backshift operator, $\phi(B)$ and $\theta(B)$ are the AR and MA polynomials
respectively and $\epsilon_t$ is a white noise process with variance $\sigma^2$. The spectral density
of (3) satisfies

$$f_x(\omega, d) = C_f(\omega) |\omega|^{-2d} \quad \text{as} \quad \omega \to 0$$

We aggregate the process $x_t$ by a level of aggregation $m$ following (2), with
$m = 2, \ldots, M$. Under the aggregated series $y_t$, the series length becomes $n = N/m$.
Note that $m = 1$ corresponds to the original series $x_t$. The spectral density of $y_t$
with memory parameter \(d\) satisfies
\[
f_y(\lambda, d) \sim m^{2d+1}C_f(\lambda)|\lambda|^{-2d}, \quad \text{as } \lambda \to 0
\]
(5)
where \(\lambda = 2\pi jm/N = \omega m\) and the periodogram of \(y_t\) is given by
\[
I_{y(m)}(\lambda_j) = \frac{1}{2\pi n} \left| \sum_{j=1}^{n} (y_j - \bar{y}) \exp^{ij\lambda_j} \right|^2, \quad \bar{y} = \frac{1}{n} \sum_{j=1}^{n} y_j/n
\]
(6)

Our statistic is constructed based on the semi-parametric local Whittle estimator proposed by Robinson [17]. Let us consider the Gaussian objective function for original series \(x_t\):
\[
Q(G, d) = \frac{1}{l} \sum_{j=1}^{l} \left[ \log(G \omega_j^{-2d}) + \frac{\omega_j^{2d}}{G} I_x(\omega_j) \right]
\]
(7)
by which discrete averaging is evaluated over a small bandwidth frequency \(l < N\). As \(G\) can be estimated by \(\hat{G} = \frac{1}{l} \sum_{1}^{l} \omega_j^{2d} I_x(\omega_j)\), then the memory parameter \(d\) can be estimated by minimizing the following objective function
\[
Q(d) = \log \left( \frac{1}{l} \sum_{1}^{l} \omega_j^{2d} I_x(\omega_j) \right) - 2d \frac{1}{l} \sum_{1}^{l} \log \omega_j
\]
(8)

Souza [16] discusses consistency of the estimator for aggregated series. It is worthwhile to summarize it as follows. Under the following regularity conditions:

1. \(f(\omega, d) \sim G_0 \omega^{-2d}\) as \(\omega \to 0^+\) where \(G_0 \in (0, \infty), -0.5 < \Delta_1 \leq d \leq \Delta_2 < 0.5\).
2. \(f(\omega, d)\) is differentiable near the origin such that
\[
\frac{d}{d\omega} \log f(\omega, d) = O(\omega^{-1}) \text{ as } \omega \to 0^+
\]
(3)

3. \(x_t - E[x_0] = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}, \sum_{j=0}^{\infty} \alpha_j^2 < \infty\)

where \(E(\epsilon_t | F_{t-1}) = 0, E(\epsilon_t^2 | F_{t-1}) = 1\) a.s., \(t = 0, \pm 1, \ldots\), in which \(F_t\) is the \(\sigma\)-field of events generated by \(\epsilon_s, s \leq t\), and there exists a random variable \(\epsilon_t\) such that \(E(\epsilon_t^2) < \infty\) and for all \(\eta > 0\) and some \(K > 0, P(|\epsilon_t| > \eta) \leq KP(|\epsilon_t| > \eta)\).
4. As \(N \to \infty, \frac{1}{N} \to 0\)
5. \(f(\omega, d)\) is bounded above, \(f'(\omega, d)\) exists and is finite in the vicinity of the non-zero Nyquist frequencies.
6. \(f(\omega, d) \sim G_0 \omega^{-2d}(1 + O(\omega^\beta))\) as \(\omega \to 0^+\) for some \(\beta \in (0, 2]\) where \(G_0 \in (0, \infty)\) and \(-0.5 < \Delta_1 \leq d \leq \Delta_2 < 0.5\).
(7) $\alpha(\omega)$ is differentiable near the origin such that
\[
\frac{d}{d\omega} \alpha(\omega) = O\left(\left|\frac{\alpha(\omega)}{\omega}\right|\right), \text{ as } \omega \to 0+
\]
where $\alpha(\omega) = \sum_{j=0}^{\infty} \alpha_j e^{ij\omega}$

(8) Condition 3 holds and $E(\epsilon_t^2|F_{t-1}) = \mu_3, a.s., E(\epsilon_t^4) = \mu_4, t = 0, \pm 1, ...$ for finite constant $\mu_3$ and $\mu_4$.

(9) There exists a $\beta$ satisfying Condition 6 such that
\[
\frac{1}{l} + \frac{1+2\beta}{l} (\log l)^2 \rightarrow 0, \text{ as } N \to \infty
\]

If conditions (1) to (5) hold for $x_t$, then it builds the consistency of the local Whittle estimator for aggregated time series $y_t$. Also, if conditions (5) to (9) hold for $x_t$, then the local Whittle estimator for $y_t$ is asymptotically normal such that
\[
\sqrt{l}(\hat{d} - d) \overset{D}{\rightarrow} N(0, 1/4)
\]

The readers are referred to [16] for the proof and the details of these conditions.

Now, consider two objective functions for two aggregated series $y^{(m_1)}$ and $y^{(m_2)}$ as follows:
\[
Q(n_1, d) = \log \left(\frac{1}{l} \sum_{j=1}^{l} \lambda_j^{2d} I_{y^{(m_1)}}(\lambda_j)\right) - 2d\frac{1}{l} \sum_{j=1}^{l} \log \lambda_j
\]
\[
Q(n_2, d) = \log \left(\frac{1}{l} \sum_{j=1}^{l} \lambda_j^{2d} I_{y^{(m_2)}}(\lambda_j)\right) - 2d\frac{1}{l} \sum_{j=1}^{l} \log \lambda_j
\]

where $Q(n_1, d)$ and $Q(n_2, d)$ denote the objective function of the aggregated series $y_t$ with level $m_1$ and $m_2$ respectively. From this, the local Whittle estimator $\hat{d}$ is defined by
\[
\hat{d}^{(m_1)} = \arg\min Q(n_1; \hat{d}), \quad \hat{d}^{(m_2)} = \arg\min Q(n_2; \hat{d})
\]

We will test the constancy of the estimated memory parameter among several aggregation levels to prove the invariance principle of the memory parameter to aggregation. The null hypothesis we attempt to test is that
\[
H_0: \hat{d}^{(m_1)} = \hat{d}^{(m_2)} = ... = \hat{d}^{(m_M)}
\]

The alternative hypothesis is therefore defined as any violation of the equalities in $H_0$, i.e at least one pair of aggregated levels, $m_i$ and $m_j$, $d^{(m_i)} \neq d^{(m_j)}$ where $i \neq j$.

In this paper, the idea of the test is similar to testing change in long memory parameter (see [15, 18, 19]). To test the constancy of the long memory parameter between two aggregated levels $\{m_1 \neq m_2\}$, we propose the following statistic
\[
z_{m_1, m_2} = \sqrt{n_1 + n_2} \left(\frac{n_1n_2}{(n_1 + n_2)^2}\right) \left(\hat{d}^{(m_1)} - \hat{d}^{(m_2)}\right).
\]
The calculation of \( z_{m_1, m_2} \) involves two levels of aggregated series for all combinations of the paired \( m \). It means that for any choice of \( M \) aggregation level, we have \( 2^M \) values of \( z \). In this case, \( M \) is chosen such that the aggregated series can still be used for estimating the long memory parameter. The supremum value is proposed as the statistical test. Therefore, to test the constancy of parameter \( d \) among several aggregation levels, we suggest the statistic

\[
\chi_n = \max_{1 \leq i, j \leq M} |z_{m_i, m_j}|, \quad i \neq j.
\]

Let \( a_k \) is the coefficient of moving average representation of \( y_t \) defined as

\[
y_t = \sum_{i=0}^{\infty} a_i \epsilon_{t-i}, \quad b_k = \frac{1}{4\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f^{-1}(\lambda, d) d\lambda
\]

where \( f^{-1}(\lambda, d) \) is the inverse of spectral density function defined in (5), and \( c_k = b_0 a_k + 2 \sum_{i=1}^{k} b_i a_{k-i} \). The asymptotic distribution of the proposed test statistic is given in the following theorem.

**Theorem 2.1:** Assume \( 0 < d < 0.5 \) and the conditions (5), (6), (7), (8) and (9) are satisfied, then by the asymptotic normality of \( \hat{d} \) we have for \( m_1 \neq m_2 \)

\[
z_{m_1, m_2} \xrightarrow{D} \sigma B(t)
\]

in \( D[0,1] \) as \( T \to \infty \) where \( T = n_1 + n_2 \), and \( B(t), 0 \leq t \leq 1 \) is a Brownian bridge with \( D \) denotes convergence in distribution. Hence, the statistic \( \chi_n \) converges to

\[
\chi_n \xrightarrow{D} \sigma \sup_{0 \leq t \leq 1} |B(t)|, \quad i \neq j
\]

and the variance \( \sigma^2 \) is given by

\[
\sigma^2 = E(\epsilon_0^4) \left( \sum_{j=0}^{\infty} a_j c_j \right)^2 + \sigma^2 \sum_{i=1}^{\infty} \left( \sum_{j=0}^{\infty} \{a_j c_{j+i} + c_j a_{j+i}\} \right)^2
\]

Proof: see appendix

From the theorem above, we reject the null hypothesis for large values of \( \chi_n \). In principle, it is possible to generate the critical values from a sequence of Brownian bridge \( B(t) \) and variance \( \sigma^2 \) as written in the theorem. However, it seems that \( \sigma^2 \) has a very complicated form which leads to some difficulties. To avoid this, the critical values will be determined by using the simulated sampling distribution of \( \chi_n \).

### 3. Simulation

This section carries out simulation studies to obtain the critical values, as well as to assess the test performance in finite sample. As we pointed out above, the critical values are obtained by using the simulated sampling distribution of \( \max_{1 \leq i, j \leq M} |z_{m_i, m_j}| \). It is done by generating 50000 sample sizes and 10000 replications. The aggregation levels are set to be \( m = 2, 3, 4, 6, 8, 12 \), which are commonly used in empirical applications as suggested by [14, 20]. In the latter work, they studied the effect of the use of aggregate time series on the Dickey-Fuller test for unit root and a new unit root test based on aggregate time series was developed. In brief, the procedure to obtain the critical values can be described as follows:
we generate the process under null hypothesis i.e. ARFIMA(0, d, 0) with sample size of 50000
we apply the test statistic with each setting of aggregation level to obtain the statistic \(z_{m_1, m_2}\) as well as \(\chi_n\).
we do above steps 10000 times and therefore, we have 10000 values of \(\chi_n\).
we tabulate the sampling distribution of \(\chi_n\) to determine the quantile of the sampling distribution as the critical values. Having this sampling distribution, we do not need to generate \(\sigma\) and \(B(t)\) since we have that \(\chi_n \xrightarrow{D} \sigma \sup_{0 \leq t \leq 1} |B(t)|\).

It is an easy implemented procedure and commonly applied to simulate the critical value of tests whose nonstandard asymptotic distribution such as unit root Dickey Fuller test and the extensions.

Table 1. Quantile of the asymptotic distribution

<table>
<thead>
<tr>
<th>d</th>
<th>sign. level</th>
<th>Aggregation level (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>0.1</td>
<td>10%</td>
<td>0.4542</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.5258</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td>0.7098</td>
</tr>
<tr>
<td>0.2</td>
<td>10%</td>
<td>0.5005</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.5909</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td>0.8253</td>
</tr>
<tr>
<td>0.3</td>
<td>10%</td>
<td>0.5164</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.6347</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td>0.8195</td>
</tr>
<tr>
<td>0.4</td>
<td>10%</td>
<td>0.6111</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.7037</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td>0.8732</td>
</tr>
</tbody>
</table>

Table 1 provides the critical values of the test for \(d = 0.1, 0.2, 0.3, 0.4\). We see that the critical value increases with \(d\) and \(m\) through the \(\sigma\) in theorem 2.1. The idea of simulating the asymptotic distribution in this case is to overcome the complicated form of \(\sigma\). We see from the theorem that the asymptotic distribution of \(\chi_n\) is explained by the \(\sigma\) and Brownian bridge \(B(t)\). As can be seen from the definition of \(z_{m_1, m_2}\), the Brownian bridge is generated from the difference between a pair of long memory parameters, which is supposed to be constant for all pairs of \(m_1\) and \(m_2\). Therefore, any systematic behavior observed in the simulated critical values can only be explained through the \(\sigma\).

Size experiment is done by evaluating the performance of the test in finite sample size. In this case, we generate 1000 time series of 5000 sample sizes. The rejection rate is calculated based on the critical values in table 1. The data generating process (DGP) is pure stationary long memory with a certain degree of fractional integration \(d = 0.1, 0.2, 0.3, 0.4\). Therefore, the DGP does not account for short dependencies \(\phi\) and \(\theta\) written in (3). The model can be rewritten as

\[(1 - B)^d x_t = \epsilon_t \quad t = 1, ..., N.\]

Table 2 presents mean and standard deviation of the estimated long memory parameter for several aggregation levels. It is useful to assess the performance of the local Whittle estimator.
As expected, the estimated memory parameters are very close to the original value. For instance, under ARFIMA(0,0.1,0), the estimated memory parameters range from 0.0999 to 0.1053. Also, under DGP ARFIMA(0,0.2,0), the estimated memory parameters range from 0.2007 to 0.2131 and so for ARFIMA(0,0.3,0) and ARFIMA(0,0.4,0). It indicates that the local Whittle estimator is a good approximation for our test. In line with [16], the standard deviation of the estimated memory parameter increases with the aggregation level. The following table presents the result of size experiment.

From table 3, it is obvious that the rejection rate is very close to the nominal value although some values indicate size distortion, meaning that the test is correctly sized under the null of long memory process.

The power experiment is carried out by generating several processes which are able to create spurious long memory, ie. Markov switching, STOP-BREAK and random level shift process. These models can be described as follows:

- **Markov-switching process**
  \[ x_t = \begin{cases} 
  \phi_1 x_{t-1} + \epsilon_t & \text{if } s_t = 1 \\
  \phi_2 x_{t-1} + \epsilon_t & \text{if } s_t = 2 
  \end{cases} \]
  with \( \epsilon_t \sim N(0,1) \), \( s_t \) is state of the Markov process with the state transition probability \( p_{00} \) and \( p_{11} \).

- **STOP-BREAK process**
  \[ x_t = \mu_t + \epsilon_t, \quad \mu_t = \mu_{t-1} + \frac{\epsilon_{t-1}^2}{\gamma + \epsilon_{t-1}^2} \epsilon_{t-1} \]
  with \( \epsilon_t \sim N(0,1) \).
- Stationary random level shift process

\[ x_t = \mu_t + \epsilon_t, \quad \mu_t = (1 - j_t)\mu_{t-1} + j_t\epsilon_t \]

with \( j_t \) is IID Bernoulli(\( p \)), \( \epsilon_t \sim \text{iidN}(0, \sigma^2_\epsilon) \) and \( \epsilon_t \sim \text{iidN}(0, \sigma^2_\epsilon) \).

- Nonstationary random level shift process

\[ x_t = \mu_t + \epsilon_t, \quad \mu_t = \mu_{t-1} + j_t\epsilon_t \]

with \( j_t \) is IID Bernoulli(\( p \)), \( \epsilon_t \) and \( \epsilon_t \) are defined as in the stationary random level shift process.

These models are strong candidates which can easily mislead the properties of long memory [9, 21–24]. We call them as model 1, model 2, model 3 and model 4 respectively hereafter. Basically, they are short memory processes with zero integration order. Therefore, any degree of fractional integration more than zero observed from these processes are spurious results. For each model, the considered parameters as well as the result of power experiment can be seen in tables 4, 5, 6 and 7.

In this part, we generate data with two different sample sizes, \( N = 2000 \) and \( N = 5000 \) with 1000 replications. Note that for \( N = 2000 \), it is considered a very small sample in practice, especially in the context of volatility modeling. Meanwhile, \( N = 5000 \) is a reasonable sample size for this case. Moreover, aggregating 5000 sample size with level of 12 results on big enough sample required to estimate the memory parameter. In the table, we present mean value of the fractional integration order obtained from 5000 sample sizes. Smaller bias is observed for smaller sample size. However, we omit the results for the reason of space.

Table 4. Power experiment

<table>
<thead>
<tr>
<th>Model 1</th>
<th>( p_{00} = p_{11} = 0.90 )</th>
<th>( p_{00} = p_{11} = 0.90 )</th>
<th>( p_{00} = p_{11} = 0.90 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_1 = -\phi_2 = 0.8 )</td>
<td>( \phi_1 = -\phi_2 = 0.5 )</td>
<td>( \epsilon_1 = N(1, 1), \epsilon_2 = N(-1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( m )</td>
<td>( \text{mean}(d) )</td>
<td>( \text{reject freq.} )</td>
<td>( \text{mean}(d) )</td>
</tr>
<tr>
<td>( \text{mean}(d) )</td>
<td>( \text{reject freq.} )</td>
<td>( \text{mean}(d) )</td>
<td>( \text{reject freq.} )</td>
</tr>
<tr>
<td>( N = 2000 )</td>
<td>( N = 5000 )</td>
<td>( N = 2000 )</td>
<td>( N = 5000 )</td>
</tr>
<tr>
<td>1</td>
<td>0.3470 ( (0.0251) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0.2115 ( (0.0319) )</td>
<td>0.989 ( 1.000 )</td>
<td>0.0450 ( (0.0225) )</td>
</tr>
<tr>
<td>3</td>
<td>0.1567 ( (0.0318) )</td>
<td>0.994 ( 1.000 )</td>
<td>0.0366 ( (0.0271) )</td>
</tr>
<tr>
<td>4</td>
<td>0.1178 ( (0.0376) )</td>
<td>0.998 ( 1.000 )</td>
<td>0.0253 ( (0.0296) )</td>
</tr>
<tr>
<td>6</td>
<td>0.0988 ( (0.0370) )</td>
<td>0.999 ( 1.000 )</td>
<td>0.0200 ( (0.0309) )</td>
</tr>
<tr>
<td>8</td>
<td>0.0610 ( (0.0455) )</td>
<td>1.000 ( 1.000 )</td>
<td>0.0120 ( (0.0407) )</td>
</tr>
<tr>
<td>12</td>
<td>0.0411 ( (0.0495) )</td>
<td>1.000 ( 1.000 )</td>
<td>0.0072 ( (0.0518) )</td>
</tr>
</tbody>
</table>
Dealing with the ability of the processes to resemble long memory, we see that all data generating processes are able to generate fractional integration orders which lie in long memory range. It can be seen from the mean values of the long memory data generating processes are able to generate fractional integration orders which in this paper is not focused on whether the models are able to create spurious long

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2290</td>
<td>-</td>
<td>-</td>
<td>0.3409</td>
<td>-</td>
<td>-</td>
<td>0.4709</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0.2842</td>
<td>0.608</td>
<td>0.989</td>
<td>0.4055</td>
<td>0.809</td>
<td>0.985</td>
<td>0.5660</td>
<td>(0.0645)</td>
<td>(0.0645)</td>
</tr>
<tr>
<td>3</td>
<td>0.3553</td>
<td>0.794</td>
<td>1.000</td>
<td>0.4589</td>
<td>0.928</td>
<td>0.996</td>
<td>0.6276</td>
<td>(0.0606)</td>
<td>(0.0606)</td>
</tr>
<tr>
<td>4</td>
<td>0.3577</td>
<td>0.735</td>
<td>1.000</td>
<td>0.5025</td>
<td>0.957</td>
<td>0.995</td>
<td>0.6708</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td>0.4005</td>
<td>0.779</td>
<td>1.000</td>
<td>0.5586</td>
<td>0.985</td>
<td>1.000</td>
<td>0.7407</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>6</td>
<td>0.4458</td>
<td>0.823</td>
<td>1.000</td>
<td>0.6003</td>
<td>0.987</td>
<td>1.000</td>
<td>0.7815</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>7</td>
<td>0.5011</td>
<td>(0.8863)</td>
<td>0.821</td>
<td>0.6817</td>
<td>0.987</td>
<td>1.000</td>
<td>0.8417</td>
<td>(0.0569)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 6. Power experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2596</td>
<td>(0.0919)</td>
<td>-</td>
<td>0.4911</td>
<td>(0.0738)</td>
<td>-</td>
<td>0.6581</td>
<td>(0.2208)</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0.3370</td>
<td>0.553</td>
<td>0.951</td>
<td>0.5845</td>
<td>0.955</td>
<td>1.000</td>
<td>0.7238</td>
<td>(0.2748)</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>0.3747</td>
<td>0.625</td>
<td>0.965</td>
<td>0.6419</td>
<td>0.986</td>
<td>1.000</td>
<td>0.8070</td>
<td>(0.2468)</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.4047</td>
<td>0.647</td>
<td>0.963</td>
<td>0.6881</td>
<td>0.985</td>
<td>1.000</td>
<td>0.8022</td>
<td>(0.2980)</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0.4606</td>
<td>0.664</td>
<td>0.968</td>
<td>0.7563</td>
<td>0.992</td>
<td>1.000</td>
<td>0.8770</td>
<td>(0.3275)</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>0.4926</td>
<td>0.668</td>
<td>0.978</td>
<td>0.8106</td>
<td>0.983</td>
<td>1.000</td>
<td>0.8797</td>
<td>(0.2782)</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>0.5076</td>
<td>(0.1617)</td>
<td>0.634</td>
<td>0.8554</td>
<td>0.990</td>
<td>1.000</td>
<td>0.8747</td>
<td>(0.3131)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7. Power experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2802</td>
<td>(0.9911)</td>
<td>-</td>
<td>0.4927</td>
<td>(0.0681)</td>
<td>-</td>
<td>0.7185</td>
<td>(0.0587)</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0.3374</td>
<td>0.553</td>
<td>0.941</td>
<td>0.5950</td>
<td>0.963</td>
<td>1.000</td>
<td>0.8266</td>
<td>(0.0486)</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>0.3875</td>
<td>0.560</td>
<td>0.964</td>
<td>0.6496</td>
<td>0.996</td>
<td>1.000</td>
<td>0.8806</td>
<td>(0.0407)</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.4064</td>
<td>0.618</td>
<td>0.972</td>
<td>0.7110</td>
<td>0.999</td>
<td>1.000</td>
<td>0.9124</td>
<td>(0.0361)</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0.4587</td>
<td>0.683</td>
<td>0.973</td>
<td>0.7656</td>
<td>1.000</td>
<td>1.000</td>
<td>0.9483</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>0.5258</td>
<td>0.640</td>
<td>0.975</td>
<td>0.8118</td>
<td>1.000</td>
<td>1.000</td>
<td>0.9665</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>0.5757</td>
<td>(0.1398)</td>
<td>0.626</td>
<td>0.8744</td>
<td>(0.0607)</td>
<td>1.000</td>
<td>0.9827</td>
<td>(0.0470)</td>
<td>-</td>
</tr>
</tbody>
</table>

Dealing with the ability of the processes to resemble long memory, we see that all data generating processes are able to generate fractional integration orders which lie in long memory range. It can be seen from the mean values of the long memory parameter under m = 1, which corresponds to the original series. Therefore, the examined parameters are correctly specified. However, the point of consideration in this paper is not focused on whether the models are able to create spurious long
memory or not, since it has been proved in the aforementioned references. Through the power experiment, we assess the behavior of the estimated memory parameter to aggregation and the ability of our test to specify these models into their class, which is spurious long memory. Since our test involves a pair of aggregation levels, thus we cannot obtain any value for $m = 1$. We denote it with "-" in the table.

Let us consider Markov switching processes in table 4. The choice of the transition probabilities mainly refers to previous works which found that the higher the transition probability $p_{ii}$, the longer the process is expected to remain in state $i$ and the process becomes more persistent. Under this condition, the process will easily be confused with long memory (see [11, 25] for intensive simulation results). The first two parameter settings in model 1 are general Markov switching processes and the last is Markov switching with iid regimes (MS-IID) and therefore, $\phi_1 = -\phi_2 = 0$. From table 4, under the defined parameter settings, the test is able to specify the Markov switching processes as spurious long memory process with high power. Only two cases have power lower than 0.5. The power increases with sample size and shows no monotonic tendency regarding the level of aggregation. However, we can see that most cases have higher power with higher aggregation level.

Now, we discuss the results for model 2. The STOP-BREAK model was introduced by Engle and Smith [26]. Similar results as Markov switching are observed for this case. Under the three different parameter settings defined in table 5, the test is able to detect the model as spurious long memory with satisfying power, both in small and medium sample size. Especially when $N = 5000$, the power reaches almost one for all cases. For random level shift processes, either stationary or nonstationary, the test also performs very well. Under small probability of Bernoulli distribution, the estimated fractional integration parameters are biased toward stationary long memory. For $p = 0.1$, the memory parameter is biased toward nonstationary long memory. It indicates that higher probability leads to a more persistent process. Since our test is derived under stationary long memory condition, therefore, this case (nonstationary long memory with $d \geq 0.5$) is out of consideration and the power of the test cannot be presented. The considered random level shift processes in this paper were firstly introduced by Chen and Tiao [27]. Further conditions about the possibility of these models to resemble long memory has been clearly investigated by Breidt and Hsu [28].

Our results in this experiment are consistent with the test proposed by [12]. Their test is also able to distinguish long memory from the spurious processes with extremely high power by setting $N = 610304$. Since [12] uses Wald type test, it is well known that the test will tend to have full power for infinite sample size. However as we pointed out before, their test looses the power significantly in finite sample size. Therefore, our test fills this gap by having good performance in finite sample size.

4. Empirical Application

The dataset used in this study consists of daily absolute and squared returns for 9 German stock price series, listed in the DAX30. The examined cases are Allianz, BASF, BAYER, BMW, Commerz Bank, Continental, Deutsche Bank, Siemens and Volkskagen (VW) spanning from the period of January 1973 to December 2007. Therefore we have 9132 observations for each stock. Several previous studies have considered German stock returns and found long memory in the considered cases [29, 30], based on the fact that several estimation procedures such as GPH, Whittle estimator or Wavelet estimator give a fractional integration order within
long memory interval. Again, it becomes crucial since several processes are able to create spurious long memory by having a certain degree of fractional integration as discussed in the previous section. Hassler and Olivares [31] independently study the daily absolute returns of the German stock price index DAX and found a significant break in mean, which might be one source of the spurious long memory.

Figure 1. ACF plot of absolute returns

Figure 1 and 2 depict the autocorrelation function (ACF) of absolute and squared returns of the considered stocks respectively. For sake of space, we present only the ACF of four stocks. The ACF of other five stocks have a similar pattern as those four. We plot the autocorrelations up to 300 lags, and the figures show that the autocorrelations of both absolute and squared returns are strongly correlated until long lags.

Figure 2. ACF plot of squared returns
They decay slowly with hyperbolic rates and show the property of long memory process. Again, having this property does not provide enough evidence that the processes are long memory. Kuswanto and Sibbertsen [11] demonstrate that several nonlinear processes under specific parameter settings may produce the same feature of autocorrelation function as long memory. This similarity holds also for the spectrum of both processes. Therefore, using only this information may lead to the wrong conclusion.

We apply our new test as a formal procedure to detect whether long memory observed in the German stocks is real or spurious. The results of the test are presented in table 8 and table 9, for absolute and squared returns respectively. In the tables, we provide the estimated long memory parameter of the aggregated series under several aggregated levels \( m \). The value in the last column is statistic \( \chi_n \) obtained from applying the test with \( m \) up to 12. This choice is based on the simulations which suggest that the test tends to have more power for high aggregation level. Table 8 presents the results of the test for absolute returns.

<table>
<thead>
<tr>
<th>stock</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>( \chi_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allianz</td>
<td>0.1959</td>
<td>0.2170</td>
<td>0.2363</td>
<td>0.2426</td>
<td>0.2587</td>
<td>0.2883</td>
<td>0.3272</td>
<td>1.0166*</td>
</tr>
<tr>
<td>BASF</td>
<td>0.2365</td>
<td>0.2945</td>
<td>0.3201</td>
<td>0.3201</td>
<td>0.2982</td>
<td>0.3070</td>
<td>0.3475</td>
<td>1.7279*</td>
</tr>
<tr>
<td>BAYER</td>
<td>0.2491</td>
<td>0.2880</td>
<td>0.3373</td>
<td>0.3640</td>
<td>0.3872</td>
<td>0.3963</td>
<td>0.4189</td>
<td>1.9809*</td>
</tr>
<tr>
<td>BMW</td>
<td>0.2437</td>
<td>0.3015</td>
<td>0.3569</td>
<td>0.3730</td>
<td>0.3894</td>
<td>0.3942</td>
<td>0.4050</td>
<td>2.3434*</td>
</tr>
<tr>
<td>Commerz Bank</td>
<td>0.2705</td>
<td>0.3142</td>
<td>0.3534</td>
<td>0.3795</td>
<td>0.3962</td>
<td>0.4335</td>
<td>0.4806</td>
<td>1.8642*</td>
</tr>
<tr>
<td>Continental</td>
<td>0.2060</td>
<td>0.2280</td>
<td>0.2460</td>
<td>0.2455</td>
<td>0.2499</td>
<td>0.2763</td>
<td>0.3068</td>
<td>0.8276**</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>0.2701</td>
<td>0.3398</td>
<td>0.3986</td>
<td>0.3966</td>
<td>0.3898</td>
<td>0.4367</td>
<td>0.4876</td>
<td>2.5551*</td>
</tr>
<tr>
<td>Siemens</td>
<td>0.2951</td>
<td>0.3480</td>
<td>0.3766</td>
<td>0.3404</td>
<td>0.4233</td>
<td>0.4709</td>
<td>0.5167</td>
<td>2.3127*</td>
</tr>
<tr>
<td>VW</td>
<td>0.2278</td>
<td>0.2829</td>
<td>0.3097</td>
<td>0.3473</td>
<td>0.3440</td>
<td>0.3582</td>
<td>0.3623</td>
<td>0.0418*</td>
</tr>
</tbody>
</table>

The * and ** sign represent significance under 5% and 10% level respectively.

From the table, by 5% level of significance the test rejects almost all cases, except for Continental. Since we have under the alternative hypothesis that there is a violation to the invariant condition of the estimated memory parameters, then to reject the null hypothesis means that the observed long memory is spurious. Continental is the only case which seems to have real long memory. It is quiet natural if we look at the \( d \) values under several aggregation levels, they are very close to each other. For this, we are only able to reject the null of long memory by 10% level of significance. Now we analyze the results for squared returns, which are given in the following table.

<table>
<thead>
<tr>
<th>stock</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>( \chi_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allianz</td>
<td>0.1470</td>
<td>0.1713</td>
<td>0.2020</td>
<td>0.2244</td>
<td>0.2467</td>
<td>0.2675</td>
<td>0.2763</td>
<td>1.6631*</td>
</tr>
<tr>
<td>BASF</td>
<td>0.2378</td>
<td>0.2673</td>
<td>0.2783</td>
<td>0.2629</td>
<td>0.2319</td>
<td>0.2389</td>
<td>0.2739</td>
<td>0.8362**</td>
</tr>
<tr>
<td>BAYER</td>
<td>0.1422</td>
<td>0.1501</td>
<td>0.1912</td>
<td>0.2911</td>
<td>0.3087</td>
<td>0.2995</td>
<td>0.2601</td>
<td>2.4222*</td>
</tr>
<tr>
<td>BMW</td>
<td>0.1994</td>
<td>0.2486</td>
<td>0.3128</td>
<td>0.3268</td>
<td>0.3227</td>
<td>0.3290</td>
<td>0.3191</td>
<td>2.3460*</td>
</tr>
<tr>
<td>Commerz Bank</td>
<td>0.2385</td>
<td>0.3029</td>
<td>0.3362</td>
<td>0.3622</td>
<td>0.3646</td>
<td>0.3601</td>
<td>0.4076</td>
<td>2.1151*</td>
</tr>
<tr>
<td>Continental</td>
<td>0.2028</td>
<td>0.2290</td>
<td>0.2615</td>
<td>0.2646</td>
<td>0.2555</td>
<td>0.2762</td>
<td>0.3037</td>
<td>1.2861*</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>0.2326</td>
<td>0.3109</td>
<td>0.3698</td>
<td>0.3631</td>
<td>0.3505</td>
<td>0.3281</td>
<td>0.3399</td>
<td>2.8387*</td>
</tr>
<tr>
<td>Siemens</td>
<td>0.2469</td>
<td>0.2842</td>
<td>0.3215</td>
<td>0.3812</td>
<td>0.4020</td>
<td>0.4202</td>
<td>0.4285</td>
<td>2.2959*</td>
</tr>
<tr>
<td>VW</td>
<td>0.1757</td>
<td>0.2454</td>
<td>0.2724</td>
<td>0.3049</td>
<td>0.2987</td>
<td>0.2991</td>
<td>0.2968</td>
<td>2.2086*</td>
</tr>
</tbody>
</table>

The * and ** sign represent significance under 5% and 10% level respectively.

In line with the result for absolute returns, the test rejects the null of real long memory. By 5% level of significance, it fails to reject the null only for BASF case. Therefore, we may say that long memory observed in most of the German stock
returns is spurious process, both in absolute and squared returns. The existence of this spurious process could be the result of non-stationarity, regime switching, mean shift, aggregation, etc. These results thus give new evidence about the behavior of German stock returns dealing with long memory.

5. Conclusion

This paper contributes to the literature on spurious long memory tests by providing a simple procedure to detect the spurious long memory based on the invariance principle of the estimated memory parameter under several aggregation levels. The test performs well in finite sample size. The empirical application gives evidence of spurious long memory in the absolute and squared German stock returns.

6. Appendix

This session gives the proof of theorem 2.1. We start the proof by showing that the following holds

\[ Q(n) - \sigma W(n) = O(n^{1/2-\varepsilon}) \quad a.s \quad << A1 >> \]

where \( \{W(t), 0 \leq t < \infty\} \) is a Wiener process and \( \varepsilon > 0 \). By theorem 1.1 of [18], condition \( << A1 >> \) is satisfied if we can show that there exists \( \varsigma > 0, \tau > 0, \vartheta > 0 \) satisfying \( \varsigma + \tau > 1/2 \) and \( \vartheta + 2\varsigma > 1 \), such that

\[
(i). \ a_k = O(|k|^{-\frac{1}{2}-\varsigma}), \quad (ii). \ b_k = O(|k|^{-\frac{1}{2}-\vartheta}), \quad (iii). \ c_k = O(|k|^{-\frac{1}{2}-\tau})
\]

where \( a_k, b_k, c_k \) are defined in the previous section.

Suppose that the original series \( x_t \) has the following infinite moving average representation:

\[ x_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i} \]

where \( \epsilon_t \) is mean zero, independent and identically distributed random variable and having variance \( \sigma_{\epsilon}^2 \). Now, equation (2) can be written as

\[ y_t = \sum_{j=0}^{m-1} B^j \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i} \]

\[ = \sum_{i=0}^{\infty} a_i \epsilon_{t-i} \]

where \( a_i = \sum_{j=i-m+1}^{i} \alpha_j \) and \( \alpha_j = 0 \) for \( j < 0 \). Before we examine \( (i) \), we firstly need to show that \( a_i \) converges in mean square. This condition has been previously examined by [20]. Nevertheless, let us describe it in brief here since it is very important for the test.
Let us define

\[(1 + B)^d = \sum_{j=0}^{\infty} \varphi_j B^j\] (14)

where for non-integer \(d\), it has been shown in proposition 2.2 of [39] and proposition 1 of [20] that \(\varphi_j = (-1)^j \binom{d}{j} = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}\) and satisfies \(\varphi_j \sim (\Gamma(-d))^{-1} j^{-(d+1)}\) as \(j \to \infty\) (see references [35] and [36] for details of the asymptotic of \(\varphi_j\)). From the definition of aggregated long memory \(y_t\) in (2) and the theorem 1 of [20],

\[(1 + B + \ldots + B^{m-1})^d = \prod_{j=1}^{m-1} (1 + \zeta_j B)^d\] (15)

\[
= \prod_{j=1}^{m-1} \left[ \sum_{k=0}^{\infty} \varphi_k \zeta_j^k B^k \right]
\] (16)

therefore, for \(d > -0.5\), \(\prod_{j=1}^{m-1} \left[ \sum_{k=0}^{\infty} |\varphi_k \zeta_j^k|^2 \right] < \infty\) and this implies that \(\sum_{j=0}^{\infty} a_j^2 < \infty\), which is the basic condition allowing to develop test statistic by using aggregated long memory. As shown in [37, 38], the condition that \(a_t\) converges in mean square and \(y_t\) has the spectral density in (5) imply

\[a_k \sim C(k)k^{d-1}\] (17)

as \(k \to \infty\) for some \(C\) slowly varying at infinity. Wang and Wang [40] examined (i) by showing that \(a_k\) decays at hyperbolic rate with an order of \(k^{H-3/2}\), where \(H\) is Hurst parameter. Given the fact that \(H = d + 0.5\), therefore (17) satisfies (i) for \(0 < d < 1/2\).

To examine (ii), let us define \(b_k = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} e^{ik\lambda} f^{-1}(\lambda, d) d\lambda\) and assume that \(f(\lambda, d)\) and \(f^{-1}(\lambda, d)\) are continuous at all \(\lambda\) and \(d\) [32] such that

\[\frac{\partial f^{-1}(\lambda, d)}{\partial d} \approx O(|\lambda|^{-2d})\] (18)

Recall the covariance of \(y_t\) as follow

\[\text{E}(y_j y_k) = \sigma_j^2 \rho(j - k) = \int_{-\pi}^{\pi} f(\lambda, d) e^{i(j-k)\lambda} d\lambda\] (19)

Define a Toeplitz matrix \(\mathcal{R}_{n\times n}\) with the \(j, k\)-th entry \(\rho(j - k)\) and a matrix \(\mathcal{A}_{n\times n}\) with the \(j, k\)-th entry \(b_{j-k}\). Then, by condition (1) and Parseval relation, \(\mathcal{A}_{n\times n}\) can be defined as an inverse of the covariance matrix \(\mathcal{R}_{n\times n}\) [33, 34] written as

\[\mathcal{R} \left( \frac{1}{4\pi^2} f^{-1}(\lambda, d) \right)\] (20)

By this relation, we intend to get the asymptotic of \(b_k\). Furthermore, by proposition
1 of [3], the autocovariance of \( y_t \) is given by
\[
\gamma_y(k) \sim m^2 \sigma^2 C_p(k)k^{2d-1} + O(k^{2d-3}), \quad \text{as } k \to \infty
\]  
(21)

From this, it is sufficient to write (21) as
\[
\gamma_y(k) \sim C(k)|k|^{2d-1}, \quad \text{as } k \to \infty
\]
and therefore for \( 0 < \delta < 1/2 - d \)
\[
|b_k| = O(|k|^\delta - 1), \quad \text{as } k \to \infty
\]
(23)

Further details about the autocovariance function of \( y_t \), the readers are referred to [3].

From (17) and (23), it is sufficient to have as \( n \to \infty \),
\[
|c_k| = O(C(k)k^{2d-1} + O(C(k)k^{2d-1+\delta})\beta(\delta, d) = O(C(k)k^{2d-1+\delta})
\]
(24)

where \( \beta(\delta, d) \) is beta function defined as \( \beta(\delta, d) = \int_0^1 y^\delta (1 - y)^{2d-1} dy \).

Now, the condition \( \ll A1 \gg \) is satisfied and we can define a sequence of Brownian bridges \( B_n(t), 0 \leq t \leq 1 \) such that
\[
\max_{0 \leq s_1, s_2 \leq 1} T^{1/2} s_1 s_2 \left\{ \frac{Q(n_1, d)}{n_1} - \frac{Q(n_2, d)}{n_2} \right\} \xrightarrow{D} \sup_{0 \leq t \leq 1} \sigma |B(t)|
\]
(26)

and
\[
\sup_{0 \leq s_1, s_2 \leq 1} \left| T^{1/2} s_1 s_2 \{d(m_1) - d(m_2)\} - \sigma B_n(t) \right| = O_p(T^{-1/2})
\]
(27)

with \( s_1 = \frac{n_1}{n_1 + n_2}, \quad s_2 = \frac{n_2}{n_1 + n_2} \) and \( T = n_1 + n_2 \) and theorem 1 is proved.

References