Distinguishing between short and long range dependence: Finite sample properties of rescaled range and modified rescaled range

Ladislav Kristoufek

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Prague, Czech Republic, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

1. July 2009

Distinguishing between short and long range dependence
Finite sample properties of rescaled range and modified rescaled range

Ladislav Krištoufek

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Prague, Czech Republic
Opletalova 26, 110 00 Prague, Czech Republic
Phone: +420 607 684 417
E-mail: kristoufek@ies-prague.org

Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic

Abstract: *Mostly used estimators of Hurst exponent for detection of long-range dependence are biased by presence of short-range dependence in the underlying time series. We present confidence intervals estimates for rescaled range and modified rescaled range. We show that the difference in expected values and confidence intervals enables us to use both methods together to clearly distinguish between the two types of processes. The estimates are further applied on Dow Jones Industrial Average between 1944 and 2009 and show that returns do not show any long-range dependence whereas volatility shows both short-range and long-range dependence in the underlying process*

Keywords: *rescaled range, modified rescaled range, Hurst exponent, long-range dependence, confidence intervals*

JEL Classification: *C01, C49, G1, G10, G14*

PACS codes: *89.65.Gh, 05.40.Jc, 05.45.Pq, 05.45.Tp, 05.10.Ln*

1 Introduction

Long-range dependence in the financial time series has been discussed in plenty of research papers (e.g. [1], [2], [3], [4], [5], [6], [7], [8] and [9]). However, majority of the papers interpret the results on the basis of simple comparison of estimated self-similarity parameter – Hurst exponent H – with its asymptotic limit of 0.5. Hurst exponent of 0.5 indicates two possible processes – either independent [10] or short-range dependent process [11]. If $H > 0.5$, the process has significantly positive correlations at all lags and is said to be long-range dependent with positive correlations [12] or persistent [13]. On the other hand, if $H < 0.5$, it has similar properties to the previous case as it has significantly negative correlations at all lags and the process is said to be long-range dependent with negative correlations [12] or anti-persistent [14].

However, the estimates for pure Gaussian process N(0,1) can strongly deviate from the limit of 0.5 ([15] and [1]). Moreover, the most used methods are biased by short-range dependence ([16] and [17]) and there is only one method which is frequently used for the time series with short-range dependent processes present – modified rescaled range. Therefore, we present our original simulations and estimates for the confidence intervals and expected values of Hurst exponent for rescaled range analysis [18] and modified rescaled range [17] and apply the results on the time series of Dow Jones Industrial Average.

In Section 2, we present and describe both techniques in detail. In Section 3, we show results of Monte Carlo simulations for time series lengths from 512 to 131072 observations and uncover the significant difference between estimated results for each method which enables us to distinguish between short-range and long-range dependence in the underlying process. In Section 4, we show that returns of DJI between 1944 and 2009 did not show any long-range dependence. On the other hand, the measures of volatility – squared and absolute returns – show strong persistence even when cleared from the potential short-range dependence.

2 Hurst exponent estimation methods

Rescaled range analysis was developed by Edwin Hurst while working as an engineer in Egypt [18] and was later applied to financial time series by Mandelbrot [19]. In the procedure, one takes continuous (logarithmic) returns of the time series of length T and divides them into N adjacent sub-periods of length υ while $N * \upsilon = T$. Each sub-period is labeled as I_n with $n = 1,2,...,N$. Moreover, each element in I_n is labeled $r_{k,n}$ with $k = 1,2,...,\upsilon$. For each sub-period, one calculates new series of accumulated deviations from the arithmetic mean values (called profile) as

$$X_{k,n} = \sum_{i=1}^{k} \left(r_{i,n} - \bar{r}_n \right), \tag{2.1}$$

where $\bar{r}_n$ is an arithmetic mean of returns of sub-period I_n.

The procedure follows in calculation of the range, which is defined as a difference between maximum and minimum value of profile $X_{k,n}$, and standard deviation of the profile for each sub-period. Each range R_{I_n} is standardized by corresponding standard deviation S_{I_n} and forms the rescaled range as

$$\left(R/S\right)_{I_n} = \frac{R_{I_n}}{S_{I_n}}. \tag{2.2}$$

The process is repeated for each sub-period of length υ and we arrive at the average rescaled range

$$\left(R/S\right)_{\upsilon} = \frac{1}{N}\sum_{n=1}^{N}\left(R/S\right)_{I_n}. \tag{2.3}$$

The length υ is increased and the whole process is repeated. We use the procedure used in [15], so that we use the length υ equal to the power of a set integer value (the method is based on the theory of multiplicative cascades which are used for a construction of fractal time series, for more details see [20] and [21]). Thus, we set a basis *b* and a maximum power *pmax* so that we get sub-periods of length $\upsilon = b, b^2, ..., b^{p\max}$ and $b^{p\max} \leq T$. Moreover, we set a minimum power *pmin* so that we get $\upsilon = b^{p\min}, b^{p\min+1}, ..., b^{p\max}$.

We get average rescaled ranges $(R/S)_\upsilon$ for corresponding sub-interval lengths υ. Rescaled range then scales as

$$\left(R/S\right)_{\upsilon} \approx c * \upsilon^{H}, \tag{2.4}$$

where *c* is a positive finite constant independent of υ ([22] and [3]).

The linear relationship in double-logarithmic scale indicates the power scaling [15]. To uncover the scaling law, we use a simple ordinary least squares regression on logarithms of each side of the previous equation. We suggest using logarithm with basis equal to *b*. Thus, we get

$$\log_b\left(R/S\right)_{\upsilon} \approx \log_b c + H\log_b \upsilon, \tag{2.5}$$

where *H* is Hurst exponent.

As the R/S analysis is known for a long time, it has been a subject to a lot of testing and criticism. The method is mostly criticized for its problematic use for heteroskedastic time series [3] and for the series with short-term memory ([23] and [24]).

The complicated use for heteroskedastic time series which is due to use of sample standard deviation together with a filtration of a constant trend makes R/S analysis sensitive to non-stationarities in the underlying process. M-R/S presented by Lo [17] differs only slightly from the original R/S and that is in the calculation of S_{I_n}. Nevertheless, it deals with both heteroskedasticity and short-term memory by modified definition of standard deviation. The new equation is defined with a use of auto-covariance γ of the selected sub-interval I_n up to the lag ξ as follows

$$S_{I_n}^M = S_{I_n}^2 + 2\sum_{j=1}^{\xi}\gamma_j\left(1-\frac{j}{\xi+1}\right). \quad (2.6)$$

Thus, R/S turns into a special case of M-R/S with $\xi = 0$ [25]. The most problematic and also the crucial issue of the new standard deviation measure is the number of lags which are used for its estimation [26]. If the chosen lag is too low, it omits lags which may be significant and therefore still biases estimated Hurst exponent by the short-term memory in the time series. On the other hand, if the used lag is too high, the finite-sample distribution deviates significantly from its asymptotic limit [27].

Majority of authors does not deal with the optimal lag choice and set several different lags which they use and examine the differences of the results (e.g. [28] and [29]). Nevertheless, there are two estimators of optimal lag suggested in the literature. The first one proposed by Lo [17] is the more complicated and still the most used one. The optimal lag is based on the first-order autocorrelation coefficient $\hat{\rho}(1)$:

$$\xi^* = \left[\left(\frac{3\upsilon}{2}\right)^{\frac{1}{3}}\left(\frac{2\hat{\rho}(1)}{1-(\hat{\rho}(1))^2}\right)^{\frac{2}{3}}\right] \quad (2.7)$$

The second one by Chin [30] is based on the length of the sub-interval only and sets the optimal lag as

$$\xi^* = \left[4\left(\frac{\upsilon}{100}\right)^{\frac{2}{9}}\right]. \quad (2.8)$$

Note that optimal lag ξ^* is recalculated for each length of specific sub-period υ. Optimal lags for different sub-period lengths are shown in Chart 2-1.

[CHART 2-1 AROUND HERE]

The method based on serial autocorrelations differs significantly with the changing correlations. In the case of low serial autocorrelations around 0.1, ξ^* is lower than the other method up to $\upsilon = 2^{16}$. However, if serial autocorrelation is doubled to 0.2 or even increased to 0.3, the differences between suggested lags ξ^* become significant. Autocorrelations of higher magnitude are rarely present in the financial time series [31]. Couillard & Davison [32] and Teverovsky, Taqqu & Willinger [27] show that M-R/S is biased towards rejecting any long-term memory in the process when high number of lags is used. Moreover, method of Lo [17] sets the optimal lag correctly only if the underlying process is AR(1) [33]. It implies that in the case of significant short-term memory in the process, the method of Lo would lead to biased estimates of *H*. Moreover, if the short-term memory is not significant or low, the method of Lo does not significantly differ from the method of Chin. It is visible from Chart 2-1 that for sub-period lengths up to 500, there is no difference between both methods with the autocorrelation of 0.2 and only a difference of one lag for the autocorrelation of 0.3. Therefore, the use of rather complicated version with serial autocorrelations does not differ significantly for the most used time series lenghts. Furthermore, for the purposes of simulations which are performed in Section 3.2, the use of the method of Lo would not lead us to strong results as the first order auto-correlation of an independent process is equal to zero, the suggested optimal lag would be zero as well and M-R/S would turn to R/S. Hence, we stick to the method of Chin [30].

3 Finite sample properties

3.1. R/S analysis

For the R/S analysis, we depict the results presented in recent research papers ([32], [15] and [1]) and then, we turn to the results of our simulations. Note that we provide such division for R/S solely as there are only several papers concerning with finite sample properties of M-R/S.

3.1.1. Recent results

R/S analysis has one significant advantage compared to the other methods – as it is known and tested for over 50 years, the methods for testing have been well developed and applied [34].

The condition for a time series to reject long-term dependence is that $H = 0.5$. However, it holds only for infinite samples and therefore is an asymptotic limit. The correction for finite samples is thoroughly tested in [32]. There are two methods used and both are based on estimating theoretical rescaled ranges for specific sub-intervals lengths.

The first method is the one of Anis & Lloyd [35], which we note AL76, and states the expected value of rescaled range as

$$E(R/S)_{\upsilon} = \frac{\Gamma\left(\frac{\upsilon-1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\upsilon}{2}\right)} \sum_{i=1}^{\upsilon-1} \sqrt{\frac{\upsilon-i}{i}} \,. \tag{3.1}$$

Peters [1] proposes “empirical correction”, which we note P94 and defines expected rescaled range as

$$E(R/S)_\upsilon = \frac{\upsilon - \frac{1}{2}}{\upsilon}\sqrt{\frac{2}{\upsilon\pi}}\sum_{i=1}^{\upsilon-1}\sqrt{\frac{\upsilon - i}{i}}. \tag{3.2}$$

The author of [1] argues that AL76 overestimates rescaled ranges for small υ. That is why $(2\upsilon - 1)/2\upsilon$ was added into equation to make it fit better the real data for small υ. Moreover, the gamma functions $\Gamma(\bullet)$ were substituted by $\sqrt{2/\upsilon}$ as when beta function $\mathrm{B}(\bullet)$ is used as a substitute of gamma function and Stirling’s approximation is applied, Peters obtains $\Gamma((\upsilon-1)/2)/\Gamma(\upsilon/2) \approx \sqrt{2/\upsilon}$. It is needed to mention that Peters used an approximation of an approximation when stating the equality. The exact application of Stirling’s approximation yields $\Gamma((\upsilon-1)/2)/\Gamma(\upsilon/2) \approx \sqrt{2/(\upsilon-1)}$ [36].

However, Couillard & Davison [32] tested the assertion and came up with different results – AL76 estimates rescaled range for small samples ($\upsilon < 500$) much more accurately and insignificantly underestimates rescaled range for large samples ($\upsilon > 500$) compared to P94. Authors also tested the asymptotic standard deviation of H (we use the same notation $\hat{\sigma}(H)$) which is essential for hypothesis testing. They argue that the Peters’ statement in [1] that $\hat{\sigma}(H) \approx 1/\sqrt{T}$ is again only an asymptotic limit and is significantly biased for finite number of observations and come to new estimate based on simulations up to $T = 10000$. The estimate states that standard deviation of H behaves as $\hat{\sigma}(H) \approx 1/e\sqrt[3]{T}$.

Unfortunately, Couillard & Davison [32] only tested the estimators up to $\upsilon = 1000$ and standard deviations up to $T = 10000$. However, the time series are often much longer. Therefore, we present the results of our original simulations in following subchapter.

3.1.2. Original results

We performed original tests for time series lengths from $T = 512 = 2^9$ up to $T = 131072 = 2^{17}$. The lengths of the time series were chosen with respect to the fact that the time series of lower lengths were shown to be rather volatile [15]. The need for an estimator of a standard deviation of Hurst exponent and the exponent itself is much more urgent for hypothesis testing than the estimators of rescaled ranges. Therefore, the simulations are performed for $E(H)$ and $\hat{\sigma}(H)$ only.

All steps of R/S analysis on 10000 time series drawn from standardized normal distribution $N(0,1)$ for $T = 2^9 = 512$ up to $T = 2^{17} = 131072$ were performed. $E_T(H)$ and $\hat{\sigma}_T(H)$ were estimated for each *T*. Hurst exponent was estimated by log-log regression according to the presented procedure. Averaged rescaled ranges applied in the regression were the ones for $2^4 \leq \upsilon \leq 2^{T-2}$. The logic behind this step is rather intuitive – very small scales can bias the estimate as standard deviations are based on very few observations; on the other hand, large scales can bias the estimate as outliers or simply extreme values are not averaged out [1].

Nonetheless, R/S estimators were tested against empirically obtained $E_T(H)$. We compared the simulated *H* with the ones estimated by AL76, P94 and corrected P94 procedure which is based on exact Stirling's approximation. We note the corrected procedure as P94c further on. AL76 contained gamma functions up to $\upsilon = 2^8 = 256$ and approximation for higher ones as gamma function for high values of υ can still cause problems to modern analytical softwares. $E_T(H)$ was obtained from rescaled ranges by log-log regression according to the power law of R/S analysis in Section 2.

The results for estimated *H* based on AL76, P94 and P94c are summed in Table 3-1.

[TABLE 3-1 AROUND HERE]

We can see that all estimates are converging to 0.50 with increasing *T* which is as expected. Note that we don't get very close to asymptotic *H* even for high *T*. However, we can't really say much about estimated Hurst exponents without the simulations.

The results for $E_T(H)$, $\hat{\sigma}_T(H)$ and corresponding descriptive statistics together with Jarque-Bera test [37] for normality are summed in Table 3-2, probability functions are showed in Chart 3-1a.

[TABLE 3-2 AROUND HERE]

[CHART 3-1 AROUND HERE]

The estimates of Hurst exponent are not equal to 0.5 as predicted by the asymptotic theory. Therefore, one must be careful when accepting or rejecting hypothesis about long-term dependence present in time series solely on its divergence from 0.5. This statement is most valid for short time series. Chart 3-2 presents the results together with estimations of *H* based on AL76, P94 and P94c. However, the Jarque-Bera test rejected normality of Hurst exponent estimates for time series lengths of 512, 65536 and 131072 and therefore, we should use percentiles rather than standard deviations for the estimation of confidence intervals [15]. Nevertheless, the differences for mentioned estimates not normally distributed are only of the order of the tenths of the thousandth and therefore, we present confidence intervals based on standard deviations for R/S.

[CHART 3-2 AROUND HERE]

[CHART 3-3 AROUND HERE]

In Chart 3-3, we present the estimated confidence intervals for 90%, 95% and 99% two-tailed significance level. From the chart, we can see that all shown confidence intervals are quite wide for short time series. Even if time series of 512 observations yields *H* equal to 0.65, we can't reject the hypothesis of an independent process even at 90% significance level. Specific values are presented in Table 3-3. The table shows that AL76 outperforms (measured by mean squared error - MSE) both P94 and P94c. Interestingly, P94c strongly outperforms P94. Nonetheless, we suggest AL76 for expected value of *H* for different *T* than we have tested here.

[TABLE 3-3 AROUND HERE]

Standard deviations of Hurst exponent $\hat{\sigma}_T(H)$ were also tested and compared with the estimations of Peters [1] and Couillard & Davison [32]. Just for reminder, authors propose that $\hat{\sigma}(H) \approx 1/\sqrt{T}$ and $\hat{\sigma}(H) \approx 1/e\sqrt[3]{T}$, respectively. Chart 3-4 shows the differences between predicted and simulated values.

[CHART 3-4 AROUND HERE]

Both estimators underestimate expected standard deviation $\hat{\sigma}_T(H)$. The estimator for infinite sample underestimates $\hat{\sigma}_T(H)$ more strongly. Therefore, we present new estimate of standard deviation, which is presented in Chart 3-4 as a solid line, as $\hat{\sigma}(H) \approx 1/\pi T^{0.3}$. Comparison of methods together with MSE is presented in Table 3-4.

[TABLE 3-4 AROUND HERE]

New method for estimation of expected standard deviation of Hurst exponent is three times more efficient than one of Couillard & Davison [32] and fifteen times more efficient than one of Peters [1] and therefore, we suggest it for estimation of $\hat{\sigma}_T(H)$ for any T from the tested interval and based on same procedure[1].

Moreover, we have shown that a combination of a minimum scale of 16 trading days with a maximum scale of a fourth of the time series length yields Hurst exponent value which is very close to all AL76, P94 and P94c methods with standard deviations almost twice lower than those of Weron [15]. Therefore, it implies that omitting of high scales is more important and efficient than omitting of scales of 16 and 32 trading days for R/S analysis.

As an implication, we propose AL76 method for an estimation of expected value of *H* with our estimate of standard deviation for a construction of confidence intervals for the application on time series. Let us follow with M-R/S.

3.2. M-R/S analysis

M-R/S analysis is rather different from R/S analysis when the applications are compared. R/S analysis is usually based on estimation of Hurst exponent itself [18]. On the other hand, only *V* statistics, defined as

$$V_\upsilon = (R/S)_\upsilon / \sqrt{\upsilon}\,, \tag{3.3}$$

is usually constructed for a specific investment horizon (scale in our case) and compared to critical values constructed by Lo [17] in the case of M-R/S. The same procedure is then applied in several research papers – e.g. [38], [39], [11] and [27]. However, the efficiency of critical values for *V* statistics of Lo [17] are criticized by Teverovsky, Taqqu & Willinger [27] as they tend to reject long-range memory more frequently than expected. Morever, the

[1] Different procedure can yield rather different results. For example, Weron [15] estimates Hurst exponent using rescaled ranges for scales of at least 50 trading days but does not restrict scales from the top which results in standard deviations almost twice a value of estimates presented in this thesis. Furthermore, the author proposes the estimates for the time series length of 256 and shows 95% confidence intervals which are almost equal extreme values of 0.2 and 0.8 for lower and upper confidence interval for the null hypothesis of independence. This implies that if the same procedure is used for the real world time series of a length of 256 trading days, the interpretation is very close to imposible.

estimates of *V* statistics and its critical values are hard to compare with estimates for *H* based on R/S. To make the results robust, we take the same path as for R/S and simulate the same random time series. We again simulated Hurst exponent for 10000 random time series drawn from standardized Gaussian distribution for minimum time series length of 2^9 and maximum one of 2^{17}. The minimum and maximum scales are set accordingly to R/S. Unfortunately, there are no theoretical estimates of modified rescaled range itself and therefore, we must stick to simulated estimates only. Note that we use the method of Chin [30], which was presented in Section 2, for estimation of optimal lag as it is the only method which bases the optimal lag on sub-period length only compared to the method of Lo [17] which is based on autocorrelations which would imply zero optimal lag and would turn M-R/S into R/S and the simulations would be of no additional information.

The descriptive statistics for simulated random time series are summed in Table 3-5. There are several interesting results. The estimates of *H* based on M-R/S are lower than those based on R/S. This finding suggests that one must be cautious when making conclusions based on comparison of Hurst exponents based on those two methods only. Moreover, standard deviations of estimates based on M-R/S are lower than the ones of R/S method and therefore, the estimates are more stable. On the other hand, distributions of Hurst exponent estimates are not normal for almost all lengths of the time series and therefore, we must stick to percentiles rather than standard deviation for the estimation of confidence intervals. The distributions are illustrated in Chart 3-5.

[TABLE 3-5 AROUND HERE]

[CHART 3-5 AROUND HERE]

As the simulated estimates are not normally distributed, we do not present any fits for estimated standard deviation as their use would not be of any help. Nevertheless, we present the confidence intervals based on percentiles together with average simulated values for Hurst exponent for M-R/S in Chart 3-6.

[CHART 3-6 AROUND HERE]

The most obvious result is the fact that the estimates of M-R/S are lower than those of R/S. The difference is more profound for upper confidence interval and is very broad at lower scales which in turn shows that R/S overestimates *H* much more than M-R/S while the statement is more valid for lower scales. For graphical comparison, we present Chart 3-7 which shows 95% confidence intervals for both R/S and M-R/S.

[CHART 3-7 AROUND HERE]

We provide expected value of *H* together with estimates for 90%, 95% and 99% two-tailed confidence intervals:

$$H_{UCI,90\%}(T)=\frac{0.6212}{T^{0.0668}} \qquad H_{LCI,90\%}(T)=\frac{0.4480}{T^{-0.0401}} \tag{3.4}$$

$$H_{UCI,95\%}(T)=\frac{0.6361}{T^{0.075}} \qquad H_{LCI,95\%}(T)=\frac{0.4480}{T^{-0.0519}} \tag{3.5}$$

$$H_{UCI,99\%}(T)=\frac{0.6623}{T^{0.0882}} \qquad H_{LCI,99\%}(T)=\frac{0.4151}{T^{-0.0824}} \tag{3.6}$$

$$E[H(T)]=\frac{0.5424}{T^{0.0189}} \tag{3.7}$$

Expected Hurst exponent decays rather slowly and does not reach a value of 0.50 up to very high time series lengths. Nevertheless, we propose the use of above mentioned estimates for the detection of significant long-term memory with short-term memory present as well and the usage of both R/S and M-R/S for comparison.

Note that possibility of R/S analysis to be biased by short-term memory process is actually an advantage of the method since *H* equal to 0.5 can mean either independent or short-term dependent process. Therefore, *H* based on R/S which is out of confidence intervals only suggests that the process is dependent since short-term memory overestimates *H* [17]. If R/S analysis could not be biased by short-term memory process, it would be impossible to say whether an estimate of $H=0.5$ means independence or short-term dependence, it would only reject long-term memory of the process. However, the use of both R/S and M-R/S enables us to distinguish between the two types of memory. If both methods show significant dependence, the process is long-term dependent. If R/S analysis shows significant dependence and M-R/S analysis does not, the process is short-term dependent. If R/S analysis shows no significant dependence the process is independent. If we use R/S and M-R/S separately, we can arrive at ambiguous results as R/S analysis can only tell us that the time series is either not long-term dependent or not independent and M-R/S analysis can only tell us that either the time series is not long-term dependent or it is. However, the rejection of long-term dependence on the basis of M-R/S would still leave us with two very different options – independence or short-term dependence.

4 Application

We apply the estimated confidence intervals on the time series of Dow Jones Industrial Average (DJI). The time series lengths chosen range from 512 to 16384 trading days between 28.3.1944 and 28.5.2009. As daily returns do not ussually exhibit short-term memory, we test squared returns and absolute returns as measures of volatility as they were already shown to exhibit strong autocorrelations [31] and therefore are ideal data to test whether the underlying process is short-range dependent, long-range dependent or combination of both. The results are presented in Table 4-1; confidence intervals are shown in Chart 4-1, Chart 4-2 and Chart 4-3.

[TABLE 4-1 AROUND HERE]

[CHART 4-1 AROUND HERE]

[CHART 4-2 AROUND HERE]

[CHART 4-3 AROUND HERE]

Returns of DJI are independent for all time series lengths which was expected as it was shown in majority of research papers (e.g. [2] and [3]). However, measures of volatitlity show quite strong implications about long-range dependence. Even when the series were filtered from basic short-range dependence, the evidence of peristence remains. Nevertheless, the differences between estimates of R/S and M-R/S are higher than for the simulated time series which implies that there is short-range dependence present in the underlying process as well as long-range dependence. Moreover, the results suggest that dependence of any kind is more prevalent for the absolute returns when compared with squared returns as the measure of volatility.

References

[1] **Peters, E.:** *Fractal Market Analysis – Applying Chaos Theory to Investment and Analysis*, John Wiley & Sons, Inc., New York, ISBN 978-0471585244, 1994

[2] **Di Matteo, T., Aste, T. & Dacorogna, M.:** *Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development*, Journal of Banking & Finance, No. 29, pp. 827-851, 2005

[3] **Di Matteo, T.:** *Multi-scaling in Finance*, Quantitative Finance, Vol. 7, No. 1, pp. 21-36, 2007

[4] **Czarnecki, L., Grech, D. & Pamula, G.:** *Comparison study of global and local approaches describing critical phenomena on the Polish stock exchange market*, Acta Physica Polonica A, 2008

[5] **Grech, D. & Mazur, Z.:** *Can one make any crash prediction in finance using the local Hurst exponent idea?*, Physica A, Vol. 336, 2004

[6] **Carbone, A., Castelli, G. & Stanley, H.:** *Time-dependent Hurst exponent in financial time series*, Physica A, Vol 344, 2004

[7] **Matos, J., Gama, S., Ruskin, H., Sharkasi, A. & Crane, M.:** *Time and scale Hurst exponent analysis for financial markets*, Physica A, Vol. 387, No. 15, 2008

[8] **Vandewalle, N., Ausloos, M. & Boveroux, Ph.:** *Detrended Fluctuation Analysis of the Foreign Exchange Market*, Econophysic Workshop, Budapest, Hungary, 1997

[9] **Alvarez-Ramirez, J., Alvarez, J., Rodriguez, E. & Fernandez-Anaya, G.:** *Time-varying Hurst exponent for US stock markets*, Physica A, Vol. 387, 2008

[10] **Beran, J.:** *Statistics for Long-Memory Processes*, Monographs on Statistics and Applied Probability 61, Chapman & Hall, New York, ISBN 978-0412049019, 1994

[11] **Lillo, F. & Farmer, J.:** *The Long Memory of the Efficient Market*, Studies in Nonlinear Dynamics & Econometrics, Vol. 8, No. 3, 2004

[12] **Embrechts, P. & Maejima, M.:** *Selfsimilar Processes*, Princeton University Press, ISBN 978-0691096278, 2002

[13] **Mandelbrot, B. & van Ness, J.:** *Fractional Brownian motions, fractional noises and applications*, SIAM Review, Vol. 10, No. 422, 1968

[14] **Barkoulas, J., Baum, Ch. & Travlos, N.:** *Long memory in the Greek stock market*, Applied Financial Economics, Vol. 10, 2000

[15] **Weron, R.:** *Estimating long-range dependence: finite sample properties and confidence intervals*, Physica A, Vol. 312, 2002

[16] **Morariu, V., Buimaga-Iarinca, L., Vamoş, C. & Şoltuz, Ş.:** *Detrended Fluctuation Analysis of Autoregressive Processes*, arXiv:0707.1437, 2007

[17] **Lo, A.:** *Long-term memory in stock market prices*, Econometrica, Vol. 59, No. 5, 1991

[18] **Hurst, H.E.:** *Long term storage capacity of reservoirs*, Transactions of the American Society of Engineers, Vol. 116, 1951

[19] **Mandelbrot, B.:** *Analysis of long-run dependence in economics: the R/S technique*, Econometrica, Vol. 39, 1970

[20] **Lux, T.:** *Applications of Statistical Physics in Finance and Economics*, Economics Working Paper No 2007-05, Christian-Albrechts-Universität Kiel, 2007

[21] **Borland, L., Bouchaud, J.-P., Muzy, J.-F. & Zumbach, G.:** *The Dynamics of Financial Markets – Mandelbrot's Multifractal Cascades, and Beyond*, Science & Finance Working Papers Archive 500061, Capital Fund Management, 2005

[22] **Taqqu, M., Teverosky, W. & Willinger, W.:** *Estimators for long-range dependence: an empirical study*, Fractals, Vol. 3, No. 4, 1995

[23] **Lo, A. & MacKinlay, C.:** *A Non-Random Walk Down Wall Street*, Princeton University Press, New Jersey, ISBN 0-691-09256-7, 1999

[24] **Alfi, V., Coccetti, F., Marotta, M., Petri, A. & Pietronero, L.:** *Exact Results for the Roughness of a Finite Size Random Walk*, Physica A, Vol. 370, No. 1, 2008
[25] **Dülger, F. & Ozdemir, Z.:** *Current Account Sustainability in Seven Developed Countries*, Journal of Economic and Social Research, Vol. 7, No. 2, 2005
[26] **Wang, W., Van Gelder, P., Vrijling, J. & Chen, X.:** *Detecting long-memory: Monte Carlo simulations and application to daily streamflow processess*, Hydrology and Earth System Sciences Discussions, Vol. 3, 2006
[27] **Teverovsky, V., Taqqu, M. & Willinger, W.:** *A critical look at Lo's modified R/S statistic*, Journal of Statistical Planning and Inference, Vol. 80, No. 1-2, 1999
[28] **Zhuang, Y., Green, Ch. & Maggioni, P.:** *The Great Rebound, The Great Crash, and Persistence in British Stock Prices*, Economic Research Paper No. 00/11, Loughborough University, 2000
[29] **Alptekin, N.:** *Long Memory Analysis of USD/TRL Exchange Rate*, International Journal of Social Sciences, Vol. 1, No. 2, 2006
[30] **Chin, W.:** *Spurious Long-range Dependence: Evidence from Malaysian Equity Markets,* MPRA Paper No. 7914, 2008
[31] **Cont, R.:** *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative Finance, Vol. 1, No. 2, 2001
[32] **Couillard, M. & Davison, M.:** *A comment on measuring the Hurst exponent of financial time series*, Physica A, Vol. 348, pp. 404-418, 2005
[33] **Andreou, A. & Zombanakis, G.:** *Computational Intelligence in Exchange-Rate Forecasting*, Working Paper No. 49, Bank of Greece, 2006
[34] **Peters, E.:** *Chaos and Order in the Capital Markets – A New View of Cycles, Prices, and Market Volatility*, John Wiley & Sons, Inc., New York, ISBN 978-0471533726, 1991
[35] **Anis, A. & Lloyd, E.:** *The Expected Value of the Adjusted Rescaled Hurst Range of Independent Normal Summands*, Biometrika, Vol. 63, 1976
[36] **Boisvert, R., Clark, Ch., Lozier, D. et al.:** *Digital Library of Mathematical Functions*, N.I.S.T., 2008 (available at http://dlmf.nist.gov/)
[37] **Jarque, C. & Bera, A.:** *Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence*, Economics Letters, Vol. 7, No. 4, 1981
[38] **Eitelman, P. & Vitanza J.:** *A Non-Random Walk Revisited: Short- and Long-Term Memory in Asset Prices*, International Finance Discussion Papers, Board of Governors of the Federal Reserve System, No. 956, 2008
[39] **Berg, L. & Lyhagen, J.:** *Short and long run dependence in Swedish stock returns*, Applied Financial Economics, Vol. 8, 1998

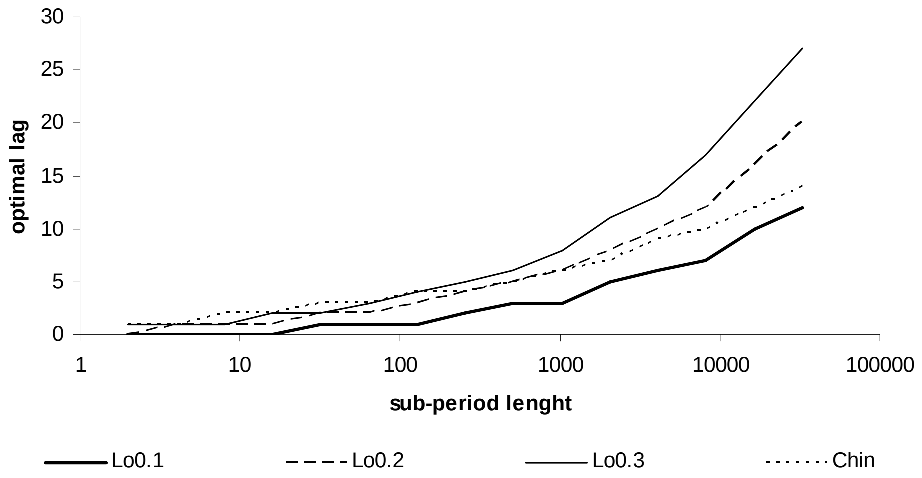


Chart 2-1 Comparison of different optimal lags for M-R/S

"Lo0.1", "Lo0.2" and "Lo0.3" stand for the first method with serial autocorrelations 0.1, 0.2 and 0.3, respectively, and "Chin" stands for the second method.

Chart 3-1 Distributions of simulated Hurst exponents

(a) (b)

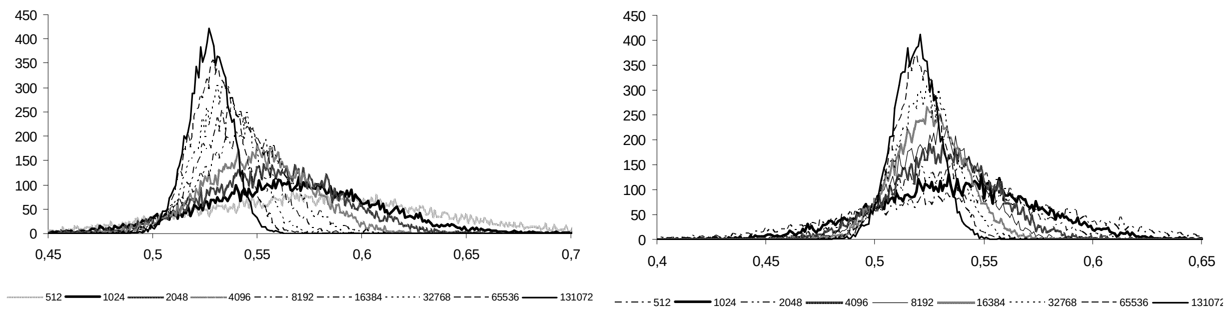


Chart 3-1: (a) R/S, (b) M-R/S: Charts show distributions of simulated Hurst exponents. For each time series length, 10000 simulations have been run with minimum scale of 16 and maximum scale of one fourth of the time series length.

Chart 3-2 Predicted and simulated Hurst exponent for R/S

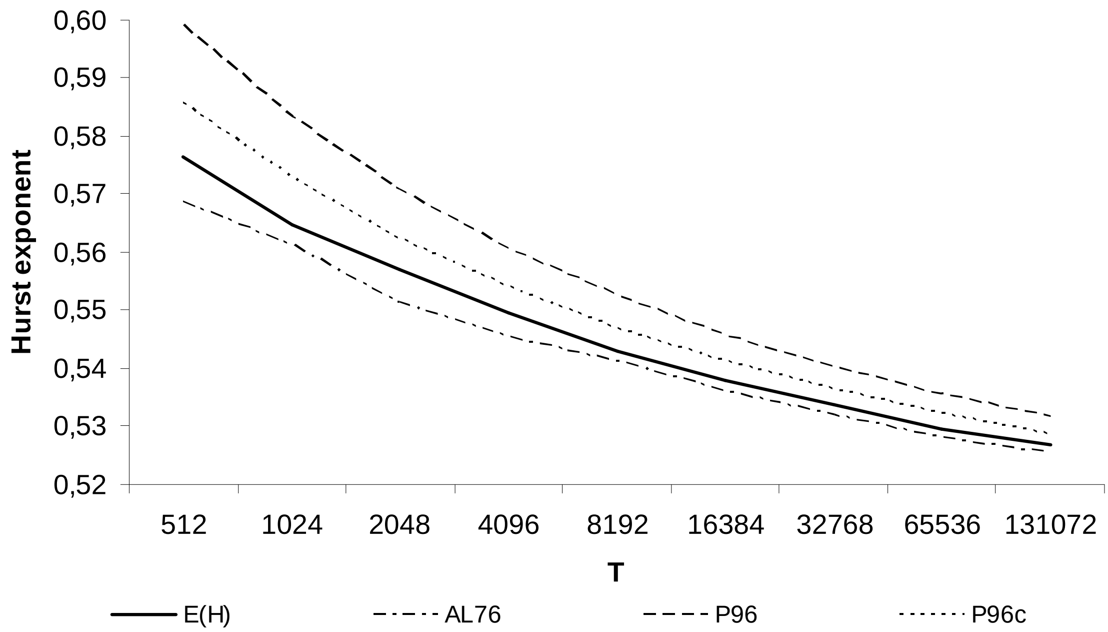
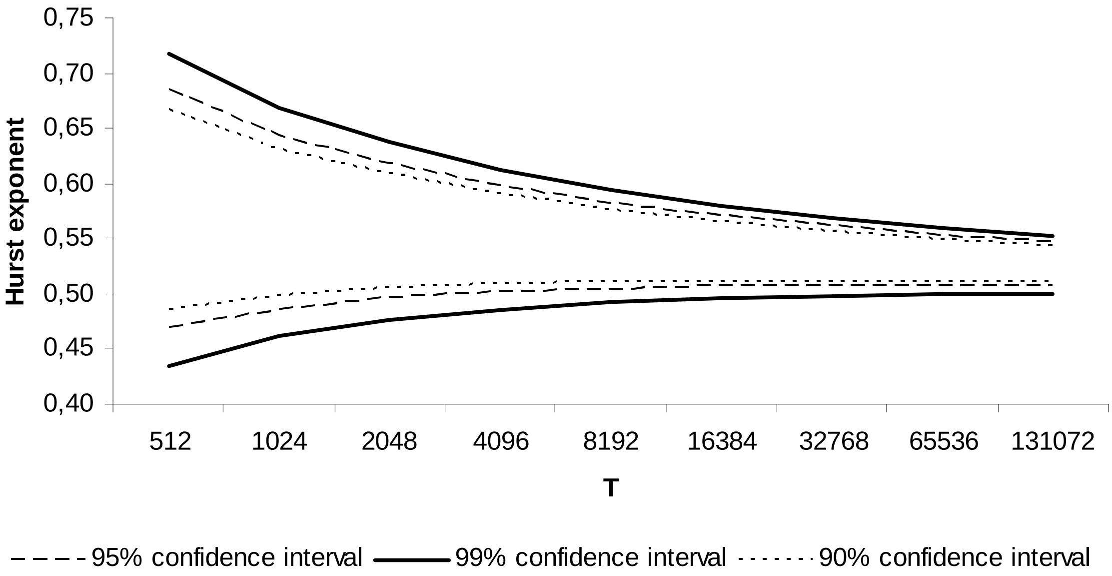


Chart 3-3 Confidence intervals for R/S



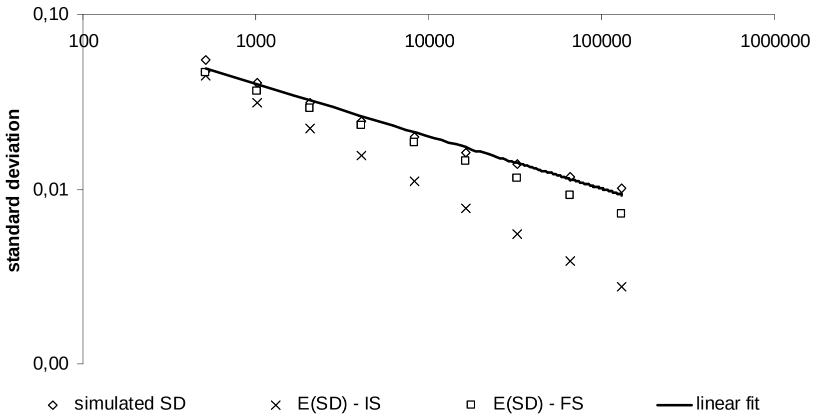


Chart 3-4 Standard deviation of Hurst exponent for R/S

"simulated SD" marks standard deviations based on our simulations, "E(SD) – IS" stands for an expected standard deviation of an infinite sampe from Peters (1994), "E(SD) – FS" stands for an expected standard deviation of a finite sample from Couillard & Davison (2005) and "linear fit" marks the ordinary least squares fit on double logarithmic scale.

Chart 3-5 Confidence intervals and expected values of simulated Hurst exponents based on M-R/S

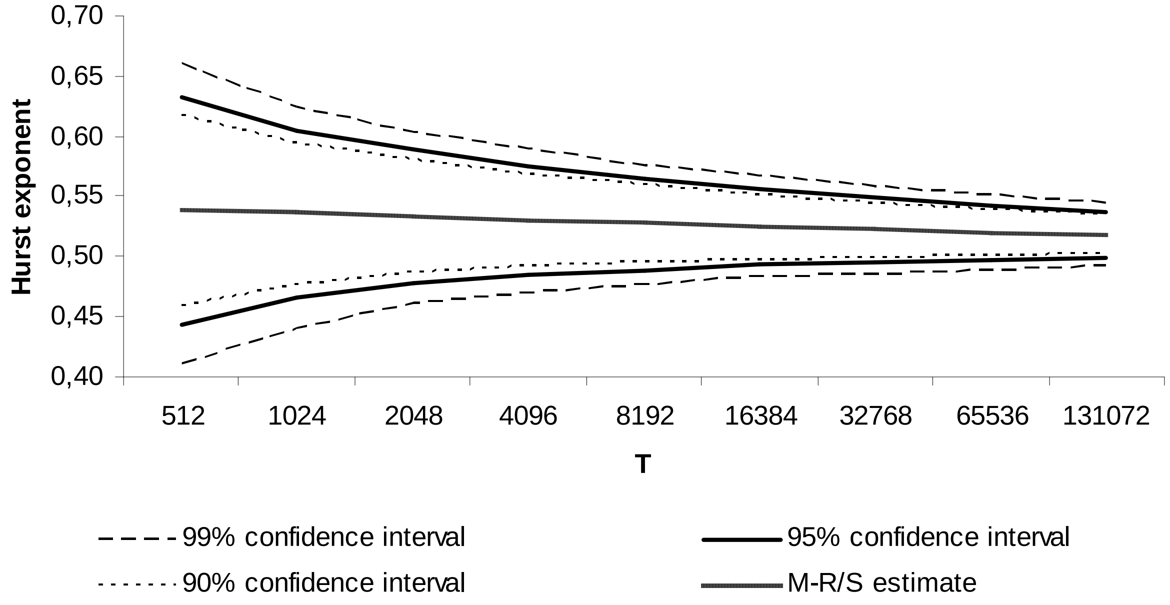


Chart 3-6 Comparison of 95% confidence intervals of simulated Hurst exponents based on R/S and M-R/S

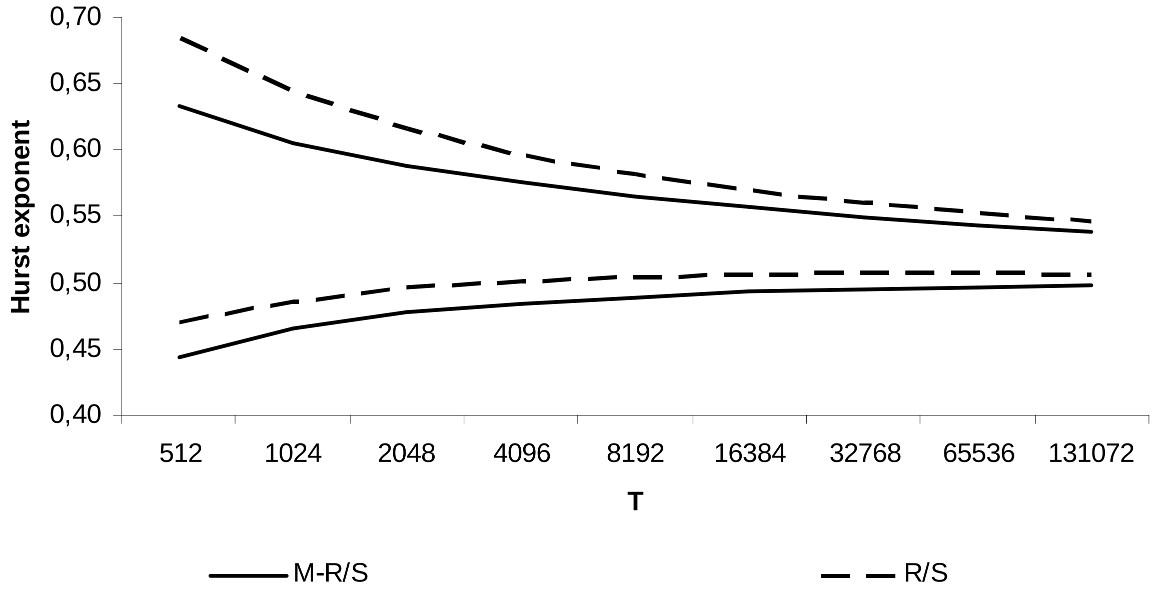


Chart 4-1 Results for DJI returns for R/S and M-R/S with 90% confidence intervals

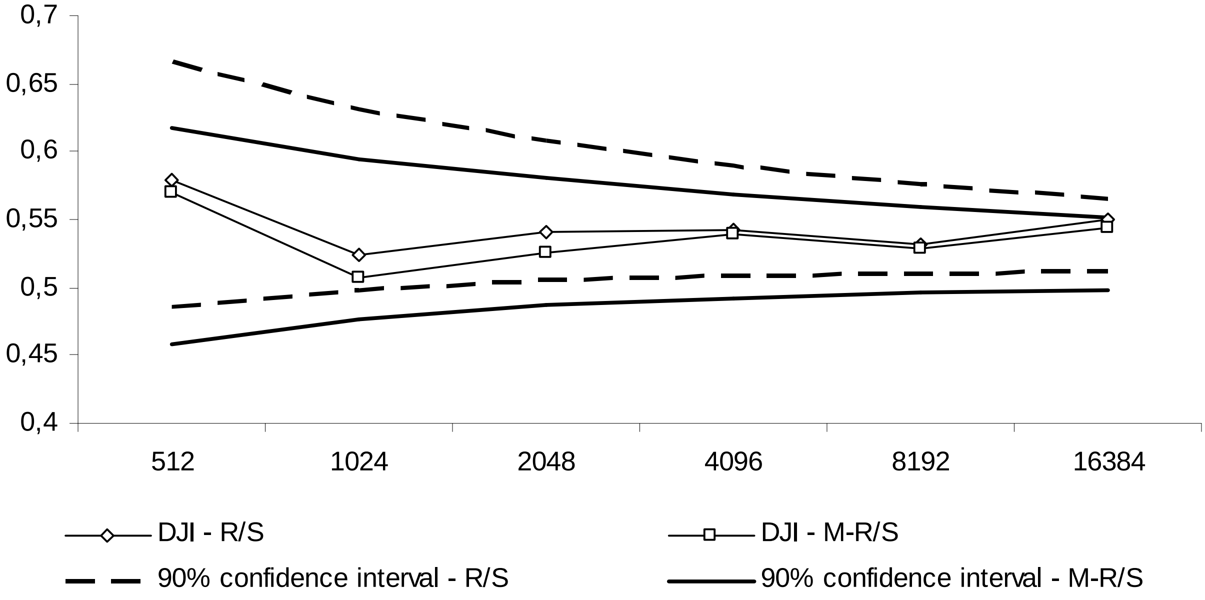


Chart 4-2 Results for DJI squared returns for R/S and M-R/S with 99% confidence intervals

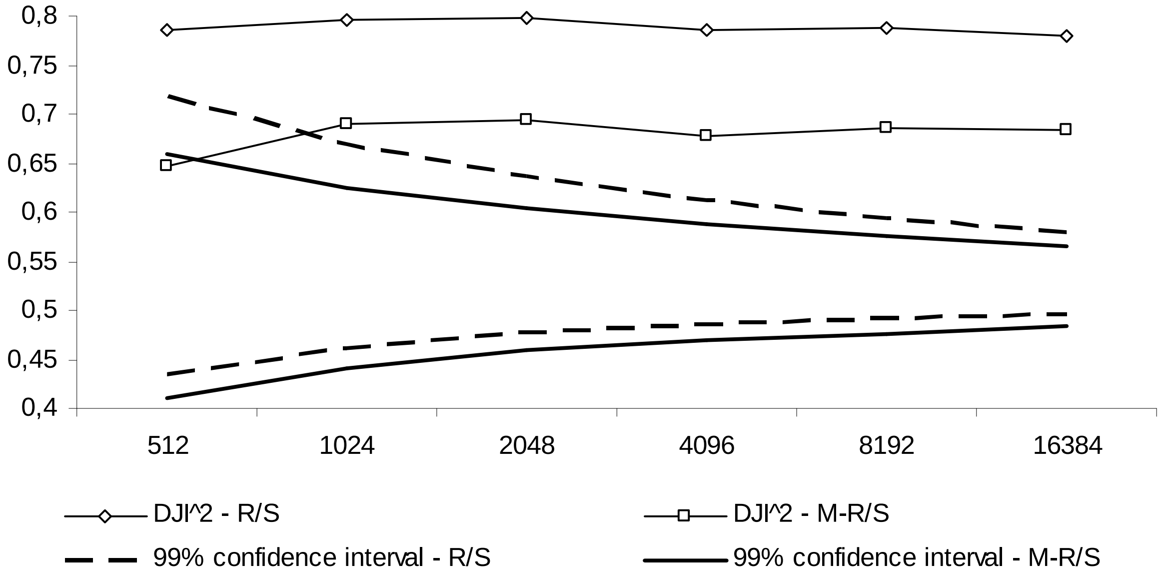


Chart 4-3 Results for DJI absolute returns for R/S and M-R/S with 99% confidence intervals

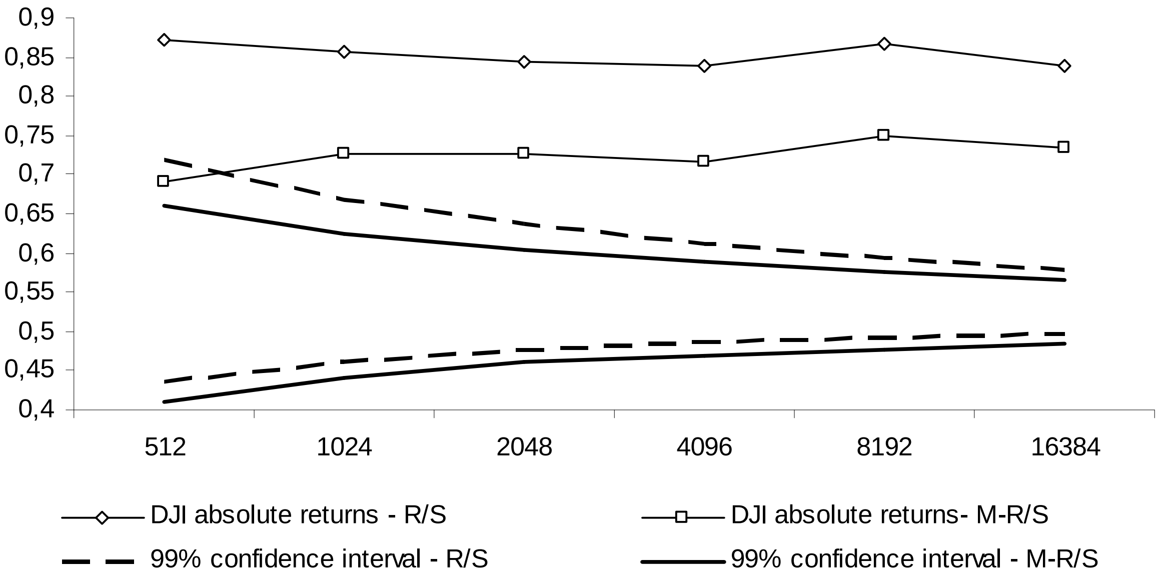


Table 3-1 Comparison of Anis & Llloyd's and Peters' formula for long series

T	*AL76*	*P96*	*P96c*
512	0,5686	0,5992	0,5858
1024	0,5611	0,5833	0,5729
2048	0,5513	0,5708	0,5624
4096	0,5455	0,5607	0,5540
8192	0,5411	0,5525	0,5470
16384	0,5361	0,5458	0,5412
32768	0,5318	0,5402	0,5363
65536	0,5282	0,5356	0,5322
131072	0,5254	0,5316	0,5287

Table 3-2 Descriptive statistics of simulated of *H* for R/S

	512	*1024*	*2048*	*4096*	*8192*	*16384*	*32768*	*65536*	*131072*
mean	0,5763	0,5647	0,5570	0,5494	0,5430	0,5380	0,5338	0,5296	0,5267
SD	0,0551	0,0404	0,0310	0,0246	0,0199	0,0162	0,0138	0,0118	0,0102
skewness	0,0104	0,0003	-0,0231	-0,0316	-0,0223	-0,0331	-0,0329	0,0068	-0,0762
excess kurtosis	-0,1316	0,0730	-0,0595	-0,0567	0,0220	-0,0271	0,0136	-0,1108	0,0237
JB statistic	7,4569	2,1800	2,3895	3,0314	1,0196	2,1440	1,8737	5,2405	9,9080
p-value	0,0240	0,3362	0,3028	0,2197	0,6006	0,3423	0,3919	0,0728	0,0071

Table 3-3 Simulated Hurst exponents compared with predicted ones for R/S

	512	*1024*	*2048*	*4096*	*8192*	*16384*	*32768*	*65536*	*131072*	*MSE*
E(H)	0,5763	0,5647	0,5570	0,5494	0,5430	0,5380	0,5338	0,5296	0,5267	
Upper CI	0,6843	0,6438	0,6178	0,5977	0,5820	0,5698	0,5608	0,5528	0,5466	
Lower CI	0,4684	0,4856	0,4962	0,5011	0,5040	0,5062	0,5068	0,5065	0,5069	
AL76	0,5686	0,5611	0,5513	0,5455	0,5411	0,5361	0,5318	0,5282	0,5254	0,000015
P96	0,5992	0,5833	0,5708	0,5607	0,5525	0,5458	0,5402	0,5356	0,5316	0,000160
P96c	0,5858	0,5729	0,5624	0,554	0,547	0,5412	0,5363	0,5322	0,5287	0,000028

Table 3-4 Comparison of standard deviations for R/S

	512	***1024***	***2048***	***4096***	***8192***	***16384***	***32768***	***65536***	***131072***	***MSE***
mean	0,5763	0,5647	0,5570	0,5494	0,5430	0,5380	0,5338	0,5296	0,5267	
SD	0,0551	0,0404	0,0310	0,0246	0,0199	0,0162	0,0138	0,0118	0,0102	
E(SD) – IS	0,0442	0,0313	0,0221	0,0156	0,0110	0,0078	0,0055	0,0039	0,0028	0,000077
E(SD) – FS	0,04598	0,0365	0,02897	0,02299	0,01825	0,01448	0,0115	0,00912	0,00724	0,000015
E(SD) - AFS	0,04899	0,03979	0,03232	0,02625	0,02132	0,01732	0,01407	0,01143	0,00928	0,000005

Table 3-5 Descriptive statistics of simulated of H for M-R/S

	512	***1024***	***2048***	***4096***	***8192***	***16384***	***32768***	***65536***	***131072***
mean	0,5393	0,5365	0,5337	0,5304	0,5278	0,5245	0,5223	0,5198	0,5182
SD	0,0485	0,0360	0,0284	0,0233	0,0192	0,0161	0,0139	0,0117	0,0101
skewness	-0,1088	-0,1048	-0,0393	-0,0693	-0,0824	0,0061	-0,0619	-0,0077	-0,0317
excess kurtosis	0,1919	0,0933	-0,0930	0,1823	0,0068	-0,0187	0,0282	0,1272	-0,0317
JB statistic	34,9582	21,8861	6,2216	21,7428	11,3207	0,2170	6,7039	6,7651	2,1094
p-value	0,0000	0,0000	0,0446	0,0000	0,0035	0,8972	0,0350	0,0340	0,3483

Table 4-1 Results for returns, squared returns and absolute returns of DJI

<table>
<tr><th rowspan="2">time series length</th><th rowspan="2">date</th><th>R/S</th><th>M-R/S</th><th>R/S</th><th>M-R/S</th><th>R/S</th><th>M-R/S</th></tr>
<tr><th colspan="2">DJI</th><th colspan="2">DJI^2</th><th colspan="2">|DJI|</th></tr>
<tr><td>512</td><td>17.5.2007 - 28.5.2009</td><td>0,57959</td><td>0,57008</td><td>0,78575</td><td>0,64679</td><td>0,8726</td><td>0,69091</td></tr>
<tr><td>1024</td><td>4.5.2005 - 28.5.2009</td><td>0,52382</td><td>0,50653</td><td>0,79564</td><td>0,69061</td><td>0,85555</td><td>0,72549</td></tr>
<tr><td>2048</td><td>4.4.2001 - 28.5.2009</td><td>0,54155</td><td>0,52541</td><td>0,79852</td><td>0,69347</td><td>0,84432</td><td>0,72605</td></tr>
<tr><td>4096</td><td>24.2.1993 - 28.5.2009</td><td>0,54208</td><td>0,53943</td><td>0,78574</td><td>0,67812</td><td>0,83834</td><td>0,71652</td></tr>
<tr><td>8192</td><td>10.12.1976 - 28.5.2009</td><td>0,53114</td><td>0,5285</td><td>0,78726</td><td>0,68649</td><td>0,8658</td><td>0,74884</td></tr>
<tr><td>16384</td><td>28.3.1944 - 28.5.2009</td><td>0,55062</td><td>0,54396</td><td>0,77972</td><td>0,68452</td><td>0,83769</td><td>0,73325</td></tr>
</table>

bold italics values show significance at 99% level of significance, **bold** values show significance at 95% level of significance