MUTUAL INFORMATION AND PERSISTENCE IN THE STOCHASTIC VOLATILITY OF MARKET RETURNS: AN EMERGENT MARKET EXAMPLE

BOGDAN DIMA[a1], Ştefana Maria DIMA[a]

[a] ECREB – East European Center for Research in Economics and Business, Faculty of Economics and Business Administration, West University of Timisoara, 16 J.H. Pestalozzi Street, 300115, Timisoara, Romania

ABSTRACT: This paper studies the volatility in financial markets' returns. First, the volatility is extracted from a stochastic volatility model. Second, various methods for persistence check are used. The results suggest that mutual information might be a valid alternative for persistence checking: significant deviations of mutual information from zero can be viewed as an evidence of long-run memory. We illustrate the case of Bucharest Stock Exchange's BET index, which displays a significant persistence in returns. The mutual information approach shows that volatility becomes more persistent during functional instability periods of the market. This result is consistent with the other methods applied.

KEYWORDS: Long-run memory**,** Stochastic volatility**,** Mutual information, Financial markets efficiency

JEL Classification: G120,G140,G150

[1] Corresponding author, Email address: bogdandima2001@gmail.com

1. Introduction

As the recent financial and real turmoil period has painfully pointed out, real world financial markets may be characterized by periods with substantial functional instability, when these are placing themselves ‘far from equilibrium’. The negative consequences of such periods are not only confined within market borders. Rather, by specific contagion mechanisms, local perturbations might be translated to other markets and / or sectors of the economic system.

Hence, as Poterba and Summers (1986:1142) argue: “explaining the dramatic intertemporal variation in real stock market prices is a fundamental problem of financial economics”.

Three main classes of approaches are usually involved in highlighting persistence (see also Rea et al., 2013). The first one is linked to the self-similarity parameter (Hurst exponent). The second one is related to the fractional integration parameter, *d*, in the generalization of the Box-Jenkins ARIMA (p,d,q) models. The third one is placed in the frequency domain (Geweke and Porter-Hudak, 1983; Reisen, 1994) here it can be included, as well, the wavelet-based estimators (Jensen, 1999; Whitcher and Jensen, 2000; Hsu, 2006; Boubaker and Péguin-Feissolle, 2013). However, as a methodological stand, we argue that ‘persistence’ is a multidimensional concept. Thus, one cannot rely, in the estimation of the persistence degree, on a single class of methods. Instead, several tools should be employed in order to capture the various features of information transmission underlying mechanisms.

In the study of persistence in financial markets’ evolutionary paths, these methods are usually applied at the level of financial assets’ prices or market indexes. Nevertheless, the various perturbations affecting these markets originate from both inside and outside. Hence, a more detailed analysis should distinguish between *persistence in long-run trends* (hysteresis effects and market rigidities, leading to misalignments from steady-state market dynamics) and, respectively, *persistence in volatility* (systematic deviations from the steady-state market path, under the impact of imperfect information and its use mechanisms). So, instead of assessing the presence of long-memory in the observed financial assets' prices, one may adopt a two-steps approach. As a preliminary step, the volatility can be extracted from these prices or, in a dynamic setting, from their associated returns. As a second step, various techniques for emphasising the persistence in the estimated volatility series can be applied.

The first type of long-memory might be justified by the occurrence of endogenous shocks connected to structural changes in the fundamental determinants of demand and supply, or to market's institutional and operational framework. Such changes may lead to a decoupling process between demand and supply, which can be shown by regime-shifts in long-run trends. The second type may be related to exogenous information shocks: if the market displays a certain degree of information inefficiency (the information is costly, asymmetrically distributed, incomplete or only partial relevant; the algorithms used by investors to deal with this information are only partial efficient), then the newly arrived information is not 'instantaneous' incorporated in prices. In such cases, the portfolio adjustment processes might take time and market might move in areas that are 'far from equilibrium'. If the market is not able to absorb information in the current period, then multi-periodic perturbations from the long-run equilibrium may occur and prices' volatility may become persistent (even if nothing changes at the level of demand and supply fundamentals and long-run trends).

Long-memory in volatility can be viewed as an evidence of market participants' inability to gather and use the available information (which itself is not necessary a 'perfect' one) and, hence, it can be linked to the issue of market (in)efficiency.

A standard approach in capturing volatility is represented by the ARCH-GARCH family models. However, empirical evidences such as Patton and Sheppard (2013) show that future volatility is more related to the volatility of past negative returns than to that of positive returns; this effect is stronger than that implied by standard asymmetric GARCH models. Also, Todorov and Tauchen (2011) find that jumps in volatility and price levels frequently occur together, are strongly dependent, and have an opposite sign. Other stylized facts include unexpected crashes, volatility clustering, 'smile' patterns in financial assets' prices or leverage effects.

Accounting for such market features, Duffie et al. (2000) advances a framework allowing for jumps in returns, stochastic volatility, as well as jumps in returns and volatility. Bayesian tests for stochastic volatility and jumps are proposed, for instance, by Ignatieva et al. (2009), Li et al. (2012), Yong and Zhang (2014), Liu and Li (2014). Such studies provide additional empirical support for persistence in volatility. Carnero (2004) finds that the relationship between kurtosis, persistence of shocks to volatility, and first-order autocorrelation of squares is different in GARCH and autoregressive stochastic volatility (ARSV) models. This difference can explain why the persistence estimated is usually higher in GARCH than in ARSV models and why Gaussian ARSV models seem to be adequate, whereas GARCH models often require leptokurtic conditional distributions. Also, the models with Markov regime changing state equations (SVMRS) proposed by Hwang et al. (2007) suggest that volatility is far less persistent and smooth than the conventional GARCH or stochastic volatility. Even more, a persistent short-run regime is more likely to occur when volatility is low, while far less persistence is possible to be observed in high volatility regimes. Messow and Krämer (2013) shows that structural changes in stochastic volatility models induce spurious persistence. Such implied persistence does not tend to unify with given size of the structural change and increasing sample size. In the case of foreign exchange market, Berger et al. (2009) finds that the time variation in the market's sensitivity to information is, at least, as relevant in explaining the persistence of volatility, as the rate of information arrival itself. Furthermore, one can argue that a financial market is not a closed system. Rather, through various financial and real flows, markets are interrelated and there might be significant volatility pass-through processes. For instance, Vo (2011) models the volatility of stock and oil futures' markets, by using the multivariate stochastic volatility structure and finds that there is a inter-market dependence in volatility (in the sense that innovations that hit either market can affect the volatility in the other market).

Based on such arguments, we seek to contribute by a three-fold analysis. First, we consider the case of Bucharest Stock Exchange, an emergent market with fast institutional and structural changes and functional instability. For this market, there are strong reasons to expect endogenous and exogenous causes of regime-shifts, not only in trends, but also in volatility. Second, we model its BET index time-varying volatility by involving a stochastic volatility model. Third, we check for long-run memory, at the level of estimated volatility, based on some of the above mentioned approaches. Supplementary, we explore the possibilities of involving mutual information as an alternative approach to persistence analysis. One significant advantage of this is related to that, unlike the autocorrelation function, mutual information takes into account nonlinear correlations as well.

The paper is organized as follows. In Section 2, we describe the considered stochastic volatility model and several features of mutual information. Section 3 shows the main statistics of BET index market. Section 4 reports on time-varying volatility estimates and proceeds to check for persistence at its level,by involving the specific frameworks for Hurst analysis, Multi-Fractal De-trended Fluctuation analysis, the fractional ("memory") parameter and, respectively, the average mutual information index; while the last section concludes.

2. Methodology

2.1. A stochastic volatility model

As Jacquier et al. (2004) argues, the stochastic volatility models offer a 'natural alternative' to the GARCH family in capturing the time-varying volatility. One of their main advantages consists in that these allow the separation of error processes for the conditional mean and, respectively, conditional variance. Moreover, such models are able to reflect the volatility clustering processes and account for heavy-tailed dynamics. Indeed, as Bentes et al. (2008) and Chan and Hsiao (2014) note, long memory and volatility clustering are two stylized facts frequently related to financial markets. Such phenomenon can be directly linked to market's agent heterogeneity: institutional versus individual, informed versus non-informed, long-term traders versus short-term traders, risk takers versus risk aversion agents, autochthonous versus foreign agents and so on. It can be reasonable presumed that such heterogeneity is high in the case of an emergent market. Thus, we shall consider such model in order to describe the volatility behaviour.

Let y_t be the vector of Romanian capital market's index (BET) log-returns. In order to account for the potential presence of extreme values in data, we fit the data using a stochastic volatility model with *Student-t* errors. In addition, we allow for potential persistence through an *MA (1)* error process, such as:

$$\begin{aligned} y_t &= \mu + u_t \\ u_t &= \varepsilon_t + \psi \varepsilon_{t-1} \end{aligned} \tag{1}$$

Here $\varepsilon_t \sim N\left(0, \lambda_t e^{h_t}\right), \varepsilon_0 = 0, |\psi| < 1$.The log-volatilities evolve according to:

$$\begin{aligned} & h_t = \mu_h + \phi_h \left(h_{t-1} - \mu_h\right) + \varsigma_t, \\ & \varsigma_t \sim N\left(0, {\sigma_h}^2\right), |\phi_h| < 1, h_1 \sim N\left(\mu_h, {\sigma_h}^2 / \left(1 - {\phi^2}_h\right)\right) \end{aligned} \tag{2}$$

Meanwhile, the distribution for the scale mixture variables is given by:

$$\left(\lambda | \nu\right) \sim IG\left(\nu / 2, \nu / 2\right) \tag{3}$$

ν stands for the Student-t distribution degree of freedom parameter. Independent prior distributions are assumed for

$$\mu \sim N\left(\mu_0, V_\mu\right), \nu \sim \mathrm{U}\left(0, \bar{\nu}\right), \mu_h \sim N\left(\mu_{h0}, V_{\mu_h}\right), \phi_h \sim N\left(\phi_{h0}, V_{\mu_\phi}\right), {\sigma_h}^2 \sim IG\left(\nu_h, S_h\right)$$

Additionally, let H_ψ be a $T \times T$ lower triangular matrix with ones on the main diagonal, with

ψ_1 on first lower diagonal, ψ_2 on second lower diagonal, and so forth. Also, let $z = H_{\psi}^{-1} y$.

Then:

$$| \left(z \middle| \mathrm{h}, \lambda, \psi, \mu \right) \sim N\left(\mu H_{\psi}^{-1} 1, \sum_z \right), \sum_z = diag\left(\lambda_1 e^{h_1}, \ldots, \lambda_T e^{h_T} \right) \qquad (4)$$

A sampler procedure can be applied to z to sample the full conditional distributions of $\mathrm{h}, \lambda, \mu, \nu, \mu_h, \phi_h, \sigma_h^{\ 2}$. Such a sampler can be build based on band matrix algorithms and it can incorporate an evaluation of the likelihood function that exploits the band structure of the covariance matrix of *y* instead of involving the conventional methods based on the Kalman filter. Also, the density of *z* can be used for a *Metropolis-Hastings* step in order to simulate the full conditional distribution of ψ (see for more details [15 and Chan, 2015).

2.2. The mutual information based assessment of persistence

The mutual information between two random variables *X* and *Y* can be defined in terms of their joint probability distribution *p(X,Y)*as:

$$I\left[X;Y \right] \equiv \int\int p\left(X,Y \right) \log \frac{p\left(X,Y \right)}{p\left(X \right) p\left(Y \right)} dX dY \qquad (5)$$

Since $I\left[X;Y \right] = 0$ only when *p(X,Y)= p(X)p(Y)*, mutual information will be bigger than zero when *X* and *Y* exhibit any co-dependence, regardless of how nonlinear that dependence is. Stronger the mutual dependence, larger the value of $I\left[X;Y \right]$. In other words, mutual information measures how much the uncertainty of *Y* is reduced if *X* has been observed. If *X=Y*, then *I[X;Y]=H[X]*. Hence, entropy can be viewed as a measure of the 'self-information' contained by *X*. Of course, entropy does not capture the time information flow between current period *t* and previous periods *t-1, t-2,…, t-k*. It rather describes the information content of variable *X* at time *t*. For instance, time delayed mutual information was suggested by Fraser and Swinney (1986) as a tool to determine a reasonable time delay in phase-portrait reconstruction for time series data. We aim to use that specific capability of the mutual information as to -valuate the persistence in volatility. More exactly, we argue that such persistence can be revealed by significant values of mutual information between current and lagged series of volatility. If there are such significant levels of mutual information, then it can be presumed that information shocks are not absorbed on short-run and they propagate over specific lags.

However, as Kinney and Atwal (2014) notes, the estimation of mutual information from finite continuous data is a non-trivial task. The main difficulty lies in the estimation of the joint distribution from a finite sample of *N* observations. One of the simplest solutions is to "bin" the data. If $X_{-\tau}$ contains the delayed values of *X* (with lag τ), this approach implies to superimpose a rectangular grid on the $X, X_{-\tau}$ scatter plot and then assign each X value ($X_{-\tau}$ value) to the column bin (row bin) into which it falls. With this purpose, the following estimate is obtained:

$$I[X;LX] \approx MI(\tau) = \sum_{ij} p_\tau(i,j) \log \frac{p_\tau(i,j)}{p(i)p_\tau(j)}, where\ \tau \text{ the lag} \qquad (6)$$

Where *p(i)* is the probability the an observation of *X* falls into the i^{th} bin, $p_\tau(j)$ is the probability that an observation of $X_{-\tau}$ falls into the j^{th} bin, and respectively, $p_\tau(i,j)$ is the joint probability that an observation of X falls into the i^{th} bin and the lagged observation falls into the j^{th} bin. Let $n(i)$, $n_\tau(j)$ and $n_\tau(i,j)$ be the number of observations falling into the i^{th} bin, the number of lagged observations falling into the j^{th} bin and the number of observations in their intersection, respectively. Thus, the probabilities are obtained as follows: $p(i) \approx n(i)/N$, $p_\tau(j) \approx n_\tau(j)/N$ and $p_\tau(i,j) \approx n_\tau(i,j)/N$. Estimates of mutual information based on this approach are commonly called 'naive' estimates (Paninski, 2003). Such estimates might systematically overestimate the mutual information. Still, Kinney and Atwal (2014) emphasizes that this problem is less severe in large datasets, since the joint probability can be determined to arbitrary accuracy, as the number of observations increases to infinity and the width of the bins decreases to zero.

3. Bucharest Stock Exchange BET index data

The BET index reflects the 10 most liquid companies listed on Bucharest Stock Exchange's regulated market segment, excluding financial investment companies (SIFs). It is an index weighted by free float capitalization. The maximum weight of each share is 20%. The main selection criterion is the company's liquidity. Since 2015, several supplementary requirements of transparency, quality reporting and communication with investors have been imposed.

The analysis covers a time span between 10 November 1997 and 10 April 2015 (daily data; close values; 4352 observations). This span covers the 2007-2010 turmoil period as well as some key events, such as changes in the legal and supervisory framework, the merger with the Romanian Association of Securities Dealers Automated Quotation market (RASDAQ), the Bucharest Stock Exchange self-listing in 2010, the first dual listing (in the case of Erste Group Bank AG, which is listed on both Vienna Stock Exchange and Prague Stock Exchange), the launch of a structured financial instruments segment and of a new Alternative Trading System (ATS), so on (see Fig. 1.).

[Insert Fig. 1. about here]

Despite its relatively fast pace, the Romanian capital market remains, in terms of market capitalization, one of the smallest among Central and Easter European markets; reflecting the insufficient performances at the macroeconomic level (see Table 1). Also, the post 2007 period was characterised by an increasing financial integration with European Union's developed markets. Hence, Bucharest Stock Exchange became more vulnerable to exogenous shocks as the recent financial and real turmoil illustrates.

[Insert Table 1 about here]

The main statistics of BET levels and log-returns are reported in Table 2. The index is platykurtic distributed and some right fat-tails effects may be detected. The distribution of returns appears to be leptokurtic and skewness values indicate the presence of left-tails effects. The values of Jarque-Bera tests as well as the values of more formal Lilliefors test (an adaptation of Kolmogorov–Smirnov test), the Cramér–von Mises criterion, Watson test as well as the Anderson–Darling test (see D'Agostino and Stephens, 1986, for a presentation of these tests) are clearly rejecting the null of both levels and returns' normal distribution. The unit roots test of Zivot and Andrews (1992), which accounts for the existence of a potential structural break at the level of data, indicates that BET index is generated by a process with stationary increments.

We further focus on estimating the log-returns volatility into the described framework of the stochastic volatility model.

[Insert Table 2 about here]

4. Results and comments

4.1. Time-varying volatility estimation

We use 100000 draws from the posterior distribution, after a burn-in period of 10000. Table 3 reports the posterior means, standard deviations and quantiles of the model parameters. The average daily return over the sample period is estimated to equal 0.030%, while its true value is 0.021%.The posterior mean of the MA(1) coefficient ψ is 0.133 with a 90% credible interval (0.108,0.159) indicating some persistence in the errors.

[Insert Table 3 about here]

Also, the degree of freedom parameter ν is estimated to be about 27.271. Such value suggests that the error distribution displays heavier tails than those specific to a Gaussian distribution.

The left panel of Fig. 2 depicts the estimate of the marginal density $p(\psi | y)$.This density plot indicates that values greater than 0.2 are highly unlikely. The right panel shows the density plot of $p(\nu | y)$. Its largest portion ranges between 20 and 40. This may be seen as an additional evidency for the relevance of an Student-t error distribution.

[Insert Fig. 2. about here]

Fig. 3 depicts the posterior means and quantiles of the time-varying standard deviation $\exp(h_t / 2)$. This indicates that the estimated volatility is characterized by a significant time-variation. Particularly, it displays peaks in November 1998 (when it goes around 1.19%), October 2010 (with a fluctuation around 2.21%), January-June 2009 (when it stays around 1.24%) or May-June 2010 (when it goes beyond 1.39%). These peaks are associated to structural, institutional and functional changes of the Romanian capital market, after the re-opening of the Bucharest Stock Exchange in 1995, as well as to the 2007-2010 financial and real turmoil.

[Insert Fig. 3. about here]

Table 4 reports the main statistics for the estimated time-varying volatility. The distribution parameters show that the estimated volatility does not display a normal distribution. The unit roots (with single structural break) test of Zivot and Andrews (1992) rejects the null of unit root with a structural break in both intercept and trend.

[Insert Table 4 about here]

As one cannot rely on a single unit root test, Table 5 reports various unit root tests (see for a description of such tests: Banerjee et al., 2003; Said and Dickey, 1984; Schwert, 1989) performed by using the R interface developed by Wuertz (2014).

The null hypothesis of the presence of unit roots in volatility is rejected by Augmented Dickey–Fuller, Elliott–Rothenberg–Stock, Phillips–Perron, and Schmidt–Phillips tests, while Kwiatkowski, Phillips, Schmidt and Shin test accept the null of stationarity. Whole, it appears that the volatility series can be reasonable be viewed as *I(0)* type process. Hence, a preliminary evidence supporting persistence in volatility is obtained.

[Insert Table 5 about here]

Since the Zivot and Andrews test rejects the null of a unit root process with drift, excluding exogenous structural change - possible in the favor of the alternative hypothesis of trend stationary process with a break in the intercept -, we check for the potential existence of several breaks. For this purpose, we perform a Bai and Perron test of multiple breakpoints (Bai, 1997; Bai and Perron, 1998; Bai and Perron, 2003). The test is implemented by Zeileis et al. (2015), following the ideas of Zeileis et al. (2003) and Zeileis et al. (2010).

Fig. 4 reports on the selection of the optimal number of breaks, based on Bayesian Info Criterion (BIC) as well as on the residual sum of squares (RSS), for a segment starting at observation *i* and ending at *j*, by looking up the corresponding element in the triangular RSS matrix.

[Insert Fig. 4. about here]

Both criteria are significantly dropping with the shift from zero breaks to one break; and are reaching minimal levels for a number of breaks equal with four. These breaks are placed on the 28.05.2001, 16.12.2004, 18.12.2007 and, respectively, 02.08.2010. The 97.5% confidence intervals are relatively non-symmetric, equaling [22.05.2001, 13.06.2001], [08.11.2004, 13.01.2005], [27.11.2007, 21.12.2007], and [30.07.2010, 10.08.2010]. This implies that the ends of a volatility regime can be easier predicted than the beginning of the subsequent regime.

We further investigate the distribution of stochastic volatility series in greater details, by fitting on it an alpha-stable distribution. Thus, we involve two approaches, the former proposed by Koutrouvelis (1980), Koutrouvelis (1981), and the latter by McCulloch (1986), as are these implemented in MATLAB by Veillette (2012). Both approaches are providing a parameter providing the shape of the distribution $\beta = 1$,specific for a distribution which is heavily skewed to the right ('extremal stable'- Zolotarev, 1986); for both approaches, the characteristic exponent $\alpha < 2$. Hence, extreme events in the volatility series are more probable than in the case of a Gaussian distribution (the tails are asymptotically equivalent to a Pareto law, i.e. they exhibit a power-law behaviour) (for more details, see Samoradnitsky and Taqqu, 1994; Barunik and Kristoufek, 2010).

Moreover, we fit a *Tweedie distributions* model for the time-varying volatility. This is an exponential dispersion model (EDM) (a two-parameter family of distributions, consisting of a linear exponential family with an additional dispersion parameter) with power mean–variance relationships. More exactly, if $\exp(h_t / 2)$ follows an EDM distribution with mean μ and variance function *V ()*, then $\mathrm{var}(\exp(h/2)) = \varnothing V(\mu), V(\mu) = \mu^p, p \geq 0$ (with $\varnothing$ being the dispersion parameter) (see Dunn and Smyth, 2008; Kendal and Jørgensen, 2011). This class includes the normal (p = 0), Poisson (p = 1), gamma (p = 2) and the inverse Gaussian (p = 3) distributions. If p>3, there is a continuous with strictly positive support stable distribution. Fig. 5 indicates that the last type of distribution fits $\exp(h_t / 2)$ (the estimation was done by using the 'tweedie' R package developed by Dunn, 2014).

[Insert Fig. 5. about here]

4.2. Hurst exponent analysis

Based on the work of Hurst (1951) and Mandelbrot and Van Ness (1968), an extended literature deals with the Hurst ('self-similarity' parameter), inclusively in the context of financial markets (Di Matteo, 2007; Ellis, 2007; Alvarez-Ramirez et al., 2008; Grech and Pamuła, 2008; Matos et al. 2008; Barunik and Kristoufek, 2010; Rea et al., 2013).

A Hurst exponent with a value equal to 0.5 indicates two possible generate processes: 1) an independent one (Beran, 1994) or, alternatively, 2) a short-range dependent one (Lillo and Farmer, 2004). If the Hurst exponent is higher than 0.5, this may reflect persistence in data.

Correlatively, if Hurst exponent is less than 0.5, this may point towards an anti-persistent process. However, the sampling properties of the Hurst exponent estimates change with the fat tails effects in distribution (Barunik and Kristoufek, 2010).

In order to estimate the presence of long-run persistence in volatility, Table 6 reports several estimators for the Hurst ('self-similarity') parameter (for detailed discussions on these methods see Taqqu et al., 1995; Montanari et al., 1999; Rea et al., 2013). The reason of not relying on a single estimation method is related to the presence of heavy tails at the level of volatility data. Such tails may affect the performance of individual estimates leading to wider confidence intervals (Barunik and Kristoufek, 2010).

[Insert Table 6 about here]

For the 'aggregated variance method', the original series is divided into blocks of size *m* and the sample variance is computed within each block. The slope $\beta = 2H - 2$ from the least square fit of the logarithm of the sample variances versus the logarithm of the block sizes is providing an estimate of the Hurst exponent. However, Giraitis et al. (1999) show that this method is asymptotically biased by the order $1/\ log\ N$ (with N being the number of observations). The 'differenced aggregated variance method' aims to distinguish jumps and slowly decaying trends from long-range dependence. It differences the sample variances of successive blocks and it uses the slope β from the least square fit of the logarithm of the differenced sample variances versus the logarithm of the block sizes, in order to provide an estimate for the Hurst exponent. The 'aggregated absolute value/ moment method' evaluates the Hurst exponent from the moments *moment*=*M* of absolute values of an aggregated FGN or FARIMA time series process. The 'Higuchi (fractal dimension) method' resembles the previous method, but, instead of blocks, a sliding window is used to compute the aggregated series (Higuchi, 1988; Bhattacharya et al., 1983) shows that the rescaled range statistic (the 'R/S method') is not robust to the deviations from stationarity. Hence, for a short memory process with slowly decaying deterministic trend, this method will provide a Hurst estimate which will spuriously suggest the presence of long-memory. The 'Whittle estimator' is a semi-parametric maximum likelihood estimator working with a part of spectrum near the origin (Beran, 1994; Robinson, 1995; Horvath and Shao, 1999; Kristoufek and Vosvrda, 2014). For this estimator, it is interesting to note that Rea et al. (2013) analyses the properties of twelve estimators for the Hurst exponent and the fractional differencing parameter (absolute value, aggregated variance, boxed periodogram, differenced variance, Higuchi or Peng periodogram, rescaled range, wavelet, Whittle, GPH and Haslett-Raftery). Rea et al. (2013) concludes that only the Whittle (frequency domain) and Haslett-Raftery (time domain) estimators produce acceptable statistical properties, with sufficiently narrow confidence intervals, for time series with fewer than 4,000 observations.

All these tests are implemented in the R package 'fArma' (Wuertz, 2013). In addition, we consider the so-called 'Generalized Hurst Exponent' (Di Matteo, 2007). This approach is based on scaling of *q-th* order moments of the increments of a time series *X(t)* (with $t = \iota, 2\iota, ..., k\iota, ..., T$) and it is based on the statistic:

$$K_q(\tau) = \frac{\left\langle \left| X(t+\tau) - X(t) \right|^q \right\rangle}{\left\langle \left| X(t) \right|^q \right\rangle} \qquad (7) \cdot$$

Here the time interval τ can vary between ι and $\tau_{\max}$. This statistic scales $K_q(\tau) \approx \left(\frac{\tau}{\iota} \right)^{qH(q)}$

(for q=2, this statistic is proportional to the autocorrelation function). Two cases can be identified: (1) the case in which *H(q)=H* is constant and independent of *q* ('mono-fractal processes', for which scaling behaviour is determined from a unique constant *H* coinciding with the Hurst coefficient) and, respectively, (2) a process with *H(q)* not constant ('multi-fractal processes' different exponents characterize the scaling of different *q-moments* of the distribution). The estimations are done for *q=1,2, 3* by using the MATLAB code provided by

(Aste, 2011). One can note that a value of q equal with 1 is associated with the scaling behaviour of the absolute values of the increments, while a value of 2 is associated with the scaling of the autocorrelation function being related to the power spectrum (Di Matteo, 2007). The values of the tests point towards the presence of long-memory persistence in volatility. The lowest estimated values correspond to the Generalized Hurst method, while the highest are provided by the Higuchi method. Nevertheless, all estimates are greater than 0.7 and, so, they are consistent with a significant persistence in volatility.

4.3. Multi-Fractal De-trended Fluctuation analysis

Supplementary, we run a Multi-Fractal De-trended Fluctuation Analysis (MFDA) for the stochastic volatility. Proposed by Kantelhardt et al. (2002) this method can be used in order to estimate the multi-fractal spectrum of power law exponents. It proceeds by dividing the time series into sub-periods. The profile of each period is described similar to the rescaled range analysis (R/S). The polynomial fit of order l, is estimated for each sub-period. Since we set $l=1$, we involve a linear de-trending. The de-trended signal $Y(t,\mathrm{i})$ which is constructed for sub-period $i = 1, \ldots, N$ is then used to define, for each sub-period of length υ, the fluctuation $F^2{}_{DFA,q}(\upsilon,i) = \left(1/\upsilon \sum_{i=1}^{\upsilon} Y(t,i)^2\right)^{\frac{q}{2}}$. Furthermore, this is averaged over N sub-periods of length υ, for different values of q, giving $F_{DFA,q}(\upsilon) = \left(1/N \sum_{i=1}^{\upsilon} F^2{}_{DFA,q}(\upsilon,i)\right)^{\frac{1}{q}}$. $F_{DFA,q}(\upsilon)$ scales as $F_{DFA,q}(\upsilon) \approx c\upsilon^{H(q)}$ (with c being a constant,which is independent of υ and $H(q)$ is the generalized Hurst exponent) (see for a more detailed presentation, Kantelhardt et al., 2002; Barunik and Kristoufek, 2010; Ihlen, 2012). The results are displayed in Fig. 6 based on the MATLAB implementation by Ihlen (2012).

[Insert Fig. 6. about here]

Typical for a multi-fractal time series, the slopes $H(q)$ are *q*-dependent. Furthermore, as commonly seen in the literature, $H(q)$ is converted to the *q*-order mass exponent, which thereafter is converted to the *q*-order singularity exponent $(h(q))$ and to *q*-order singularity dimension$(D(q))$. The plot of $h(q)$ versus $D(q)$ reflects the multi-fractal spectrum. As in the case for stochastic volatility, a multi-fractal time series has a mass exponent with a curved *q*-dependency and, consequently, a decreasing singularity exponent. The corresponding multi-fractal spectrum is characterised by the difference between maximum and minimum levels of $h(q)$ (the multi-fractal spectrum width). Since the multi-fractal spectrum displays a left truncation, it can be presumed that the volatility has a multi-fractal structure, which is insensitive to the local fluctuations with large magnitude (Ihlen, 2012).

We also use a version of this implementation, where the multi-fractal spectrum is directly estimated from local fluctuation. As Ihlen (2012) notes: "the width and shape of the multifractal spectrum reflect the temporal variation of the local Hurst exponent or, in other words, the temporal variation in the local scale invariant structure of the time series". The temporal variations of local Hurst exponent are synthetized by a histogram representing the probability distribution, while the multi-fractal spectrum is defined as the log-transformation of the normalized probability distribution.

The local Hurst exponent is able to discriminate between periods with small and, respectively, large fluctuations. The relatively large width of probability distribution and multi-fractal spectrum are specific to multi-fractal series.

[Insert Fig. 7. about here]

4.4. The fractional ("memory") parameter

To complete the preliminary analysis, we account for the fact that the autoregressive fractionally integrated moving average process, *ARFIMA (p, d, q)*, has widely been used in order to represent a time series with long memory properties (Beran, 1994). One of the key issues in the estimation of such model is represented by the estimation of the fractional parameter *d*. More exactly, a simple *ARFIMA (p, d, q)* takes the following form:

$$\Phi(B)(1-B)^{d} = \Theta(B)\varepsilon_{t} \tag{8}$$

Here ε_t is a white noise process and B is the back-shift operator. The polynomials $\Phi(B) = 1-\Xi_1 B - ... - \Xi_p B^p$, $\Theta(B) = 1-\Theta_1 B ... - \Theta_s B^s$ have orders p and, respectively, s with all their roots outside the unit circle (Reisen et al., 2001).

The fractional ("memory") parameter *d* was estimated by using the R package 'fracdiff' (Frailey et al., 2012). The estimates are based on Geweke and Porter- Hudak (1983): estimator and, respectively, on Reisen estimator (Reisen, 1994; Reisen et al., 2001; Reisen et al., 2001). The results are reported in Table 7.

[Insert Table 7 about here]

These estimators are providing large (and in almost all cases statistical significant) values for the fractional parameters. The Reisen method for estimating *d* is more accurate than the Geweke and Porter- Hudak estimator,but both methods are indicating the presence of long-memory in stochastic volatility.

4.5. Mutual information analysis

4.5.1. Shannon's index of diversity (Chao and Shen estimator)

As a preliminary stage in the mutual information analysis, we consider the dynamic of volatility entropy over overlapping windows (with a pre-determined length of 50 days - approximately a two trading month span). With this purpose, we take into account the estimator proposed by Chao and Shen (2003), as implemented in R language by Hausser and Strimmer (2014). This is based on unequal probability sampling and account for the nonuniform distribution of data. The estimations are displayed by Fig. 8.

[Insert Fig. 8. about here]

As these estimations suggest, the volatility is characterized by shifts from low-entropy regimes to high-entropy regimes (with significant peaks in the last months of 2003 and, respectively, 2004 as well as in the first quarter of 2011). Also, the wavelet power spectrum (rectified according to Liu et al. (2007) and implemented in the MATLAB codes provided by Ng and Chan (2012) and Grinsted et al. (2004) shows that the Chao and Shen estimator for volatility is high and statistically significant especially for frequencies up to 32 days. Hence, there is room for a more detailed analysis of such flows, which can be based on mutual information between current and lagged values of volatility. We further focus on the estimation of the mutual information index.

4.5.2. Average mutual information index

The estimates of this index up to a lag equal with 250 days are reported in Fig. 9. The estimations are performed by using the R package 'tseriesChaos' (di Narzo, 2013).

[Insert Fig. 9. about here]

For comparison purposes, we show, in the left panel, the average mutual information for a fractional Gaussian noise (fGn) process with the same number of observations (4352) and an Hurst exponent equal to the average estimates (0.883) as for the time-varying volatility series. For such process, the mutual information decline fast and stays around zero for different lags.

Still, as this figure suggests, the mutual information index for stochastic volatility declines relatively fast up, till a lag equal with 50 days (approximately two trading months). However, it does not converge to zero beyond this lag and it even starts to slowly increase for lags greater than 162 (approximately two trading quarters). The global maximum is reached in mutual information for the first lag, while other local maxima are placed around lags equal with 25, 125 and, respectively, 200 days.

In order to investigate this, in greater details, we apply a Monte Carlo simulation based on 10000 draws from a fGn process with the same characteristics as volatility. These draws are generated accordingly to the procedure described in (Beran, 1994). For each draw, we estimate the mutual information up to a lag equal with 250 observations.

Fig. 10 displays the average squared deviations of volatility specific mutual information from the fGn processes mutual information at the same lag: $\sum_{i=1}^{10000}\left[MI(\tau)-MI^{fgn}(\tau)\right]^2/10000,\ \tau=1,\ldots,250$ (where $MI^{fgn}(\tau)$ is the mutual information of the fGn processes up to a lag equal with τ). It appears that such difference is fast declining for small lags in order to increase again for lags around 125 days and, respectively, for lags greater than 200 days (Fig. 10).

[Insert Fig. 10. about here]

Overall, the mutual information analysis reveals the existence of both medium and long-run memory in the specific BET market volatility generative processes. This outcome is robust to various choices for the number of considered bins (not reported here).

Possible explanations for such findings may be related to: the low market capitalization of $33.5 billion as of January 2015 (even if, between January 2012 and January 2015, the domestic market capitalization increased approximately by 70% mainly due to the listing of new companies on the regulated market); the small set of traded financial instruments (including only shares, bonds, fund units and certificates); the limited number of listed companies (83 in January 2015); the frequency of thin trade cases on different market components; and the fact that (with the exception of a small segment including only 20 securities) short trades are not allowed.

There are fewer arbitrage opportunities on the market and the investors seldom involve in high frequencies trades. Thus, the investors feel constrained to adopt more ‘passive’ trading strategies, with larger horizons for portfolios holding. 'Buy and hold' monthly or even quarterly trading strategies are easier to be applied in comparison with intra-day scalping strategies. Alternatively, these results can reflect lower information efficiency: due to both information asymmetry and moral hazard phenomena, the new information shocks are slowly absorbed by the market, being translated over several trading periods.

In order to provide more insights, we 'zoom-in' by splitting the stochastic volatility series in overlapping windows with a pre-determined length of 500 days (two trading years) and we compute the mutual information inside each individual window for lags up till 50 days. Fig.10 displays the minimal values of these estimates. If such values are significantly different from zero, then it can be argued that mutual information is capturing persistence at the level of the stochastic volatility series.

It is interesting to note that the results resemble the estimates of local Hurst exponent, with larger deviations of minimal local mutual information from zero appearing almost in the same data positions as the large Hurst exponent values (during 2004-2005 as well as for the time span of 2008-2010). After the peak of the instability period, local mutual information starts slowly to decline, but without fully converging to zero. It is likewise interesting to note that the breaks identified by Bai and Perron tests are associated as well with noticeable values of mutual information (Fig. 11).

[Insert Fig. 11. about here]

Overall, the mutual information approach shows that volatility becomes more persistent during functional instability periods for the market, while in normal conditions it declines without completely vanishing. Such outcome is consistent with the results of the other involved methods.

5. Conclusions and discussions

This paper deals with persistence in financial markets returns' volatility. We argue that such persistence might arise as a consequence of the imperfect nature of the information available on the markets as well as of market agents' imperfect tools of gathering and processing such information.

We extract the time-varying volatility from returns, based on a stochastic model as an alternative to ARCH-GARCH approach. We further apply several methods in order to check for persistence (the Hurst exponent analysis, the Multi-Fractal De-trended Fluctuation analysis, and the fractional ("memory") parameter). We also suggest that the mutual information might be a valid alternative method for persistence checking: significant deviations of mutual information from zero (alternatively, from the levels of mutual information specific to a fractional Gaussian noise with the same Hurst exponent as the empirical series) at higher order lags can be interpreted as an evidence of long-memory. A rolling estimation of mutual information can provide more insights (with the length of the rolling window acting as a 'zoom-in' parameter). We argue that one major advantage of applying the mutual information consists in that it is a 'model-free' procedure: no special requirements about the underlying mechanisms of returns and their volatilities are needed on an ex-ante basis. The results from the mutual information approach are consistent with those derived from other methods. Moreover, mutual information appears able to identify the regime-shift areas.

As the case of Bucharest Stock Exchange BET index illustrates, markets can displays persistence in returns' volatility. Such finding is in the line with the other findings in the literature. Such empirical findings seriously challenge the Efficient Market Hypothesis (EMH). For instance, the persistence in volatility can lead to the existence of some arbitrage opportunities. The existence of such opportunities contradicts a strong requirement of EMH, according to which discrete time markets are efficient, if and only if, there are no arbitrage opportunities (Jarrow and Larsson, 2012). Comprehensively, the assumption of 'weak-form' market (in) efficiency is directly tied to the martingale property of security price processes. Kristoufek and Vosvrda (2013) argues that this feature is primarily related to uncorrelated price time series. Hence, the existence of such long-memory can contradict the core assumptions of EMH. Thus, a more realistic framework able to accommodate the empirical evidences against EMH should be developed. Such framework should not necessarily be build-up on the assumption that EMH if fully invalid. Rather, it may argue that EMH is valid but only as a local property of the market: the market paths might include both areas of high efficiency (in which EMH holds) and areas with low efficiency in which EMH is no longer valid.

References

[1] Aste, T. (2011). Generalized Hurst exponent, MATLAB code http://www.mathworks.com/matlabcentral/fileexchange/30076-generalized-hurst-exponent.

[2] Alvarez-Ramirez, J., Alvarez, J., Rodriguez, E. & Fernandez-Anaya, G. (2008). Time-varying Hurst exponent for US stock markets. *Phys. A Stat. Mech. Its Appl.* 387, 6159–6169.

[3] Bai, J. (1997). Estimation of a Change Point in Multiple Regression Models. *Rev. Econ. Stat.* 79, 551–563.

[4] Bai, J., & Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66, 47–78.

[5] Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *J. Appl. Econom.* 18, 1–22.

[6] Banerjee, A., Dolado, J.J., Galbraith, J.W., & Hendry, D. (2003). *Co-integration, error-*

correction, and the econometric analysis of non-stationary data. Oxford University Press.

[7] Barunik, J., & Kristoufek, L. (2010). On Hurst exponent estimation under heavy-tailed distributions. *Phys. A Stat. Mech. Its Appl.* 389, 3844–3855.

[8] Bentes, S.R., Menezes, R., & Mendes, D.A. (2008). Long memory and volatility clustering: Is the empirical evidence consistent across stock markets?. *Phys. A Stat. Mech. Its Appl.* 387: 3826–3830.

[9] Beran, J. (1994). *Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Book 61)

[10] Berger, D., Chaboud, A., & Hjalmarsson, E. (2009). What drives volatility persistence in the foreign exchange market? *J. Financ. Econ.* 94, 192–213.

[11] Bhattacharya, R., Gupta, V.K., & Waymire, E. (1983). The Hurst effect under trends. *J. Appl. Probab.* 20, 649–662.

[12] Boubaker, H., & Péguin-Feissolle, A. (2013). Estimating the Long-Memory Parameter in Nonstationary Processes Using Wavelets. *Comput. Econ.* 42(3), 291-306.

[13] Carnero, M.A. (2004). Persistence and Kurtosis in GARCH and Stochastic Volatility Models. *J. Financ. Econom.* 2, 319–342.

[14] Chan, J.C.C. (2015). The Stochastic Volatility in Mean Model with Time-Varying Parameters: An Application to Inflation Modeling. CAMA Working Paper, 7(May), available at SSRN: http://ssrn.com/abstract=2579988.

[15] Chan, J.C.C., & Hsiao, C.Y.L. (2014). Estimation of Stochastic Volatility Models with Heavy Tails and Serial Dependence. In I. Jeliazkov & X.S. Yang (Eds.), *Bayesian Inference in the Social Sciences* (pp. 159-180). New York: John Wiley & Sons; Matlab code available at: http://people.anu.edu.au/joshua.chan.

[16] Chao, A., & Shen, T.J. (2003). Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environ. Ecol. Stat.* 10, 429–443.

[17] D'Agostino, R.B., & Stephens, M.A. (1986). *Goodness-of-Fit Techniques (Statistics: textbooks and monographs, Volume 68).* New York: Marcel A.Deckker.

[18] Di Matteo, T. (2007). Multi-scaling in finance. *Quant. Financ.*, 7, 21–36.

[19] Duffie, D., Pan, J., & Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-diffusions. *Econometrica*, 68, 1343–1376.

[20] Dunn, P.K. (2014). Tweedie exponential family models, R package 'tweedie' Version 2.2.1, Available at: http://cran.r-project.org/web/packages/tweedie/tweedie.pdf,2014.

[21] Dunn, P.K., & Smyth, G.K. (2008). Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Stat. Comput.*, 18, 73–86.

[22] Ellis, C. (2007). The sampling properties of Hurst exponent estimates. *Phys. A Stat. Mech. Its Appl.*, 375, 159–173.

[23] di Narzo, A.F. (2013). Analysis of nonlinear time series. R Package 'tseriesChaos', version 0.1-13, Available at: http://cran.r-project.org/web/packages/tseriesChaos/tseriesChaos.pdf.

[24] Fraser, A.M., & Swinney, H.L. (1986). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A.*, 33, 1134–1140.

[25] Frailey, C., Leisch, F., Maechler, M., Reisen, V., & Lemonte, A. (2012). Fractionally differenced ARIMA aka ARFIMA(p,d,q) models. R package 'fracdiff', http://cran.r-project.org/web/packages/fracdiff/fracdiff.pdf.

[26] Geweke, J., & Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models. *J. Time Ser. Anal.*, 4, 221–238.

[27] Giraitis, L., Robinson, P.M., & Surgailis, D. (1999). Variance-type estimation of long memory. *Stoch. Proc.Appl.*, 80, 1–24.

[28] Grech, D., & Pamuła, G. (2008). The local Hurst exponent of the financial time series in the vicinity of crashes on the Polish stock exchange market. *Phys. A Stat. Mech. Its Appl.*, 387, 4299–4308.

[29] Grinsted, A., Moore, J.C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.*, 11, 561–566. MATLAB code: http://www.glaciology.net/wavelet-coherence.

[30] Hausser, J., & Strimmer, K. (2014). Estimation of Entropy, Mutual Information and Related Quantities, R package 'entropy'. version 1.2.1., http://cran.r-project.org/web/packages/entropy/entropy.pdf.

[31] Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Phys. D Nonlinear Phenom.*, 31, 277–283.

[32] Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civ. Eng.*, 116, 770–808.

[33] Horvath, L., & Shao, Q.-M. (1999). Limit Theorems for Quadratic Forms with Applications to Whittle's Estimate. *Ann.Stat.*, 9(1), 146–187.

[34] Hsu, N.-J.J. (2006). Long-memory wavelet models. *Stat. Sin.*, 16, 1255–1271.

[35] Hwang, S., Satchell, S.E., & Valls Pereira, P.L. (2007). How Persistent is Stock Return Volatility? An Answer with Markov Regime Switching Stochastic Volatility Models. *J. Bus. Financ. Account.*, 34, 1002–1024.

[36] Ihlen, E.A.F. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Front. Physiol.*, 3, 141; MATLAB code available at: http://www.mathworks.com/matlabcentral/fileexchange/38262-multifractal-detrended-fluctuation-analyses.

[37] Ignatieva, K., Rodrigues, P., & Seeger, N. (2009). Stochastic Volatility and Jumps: Exponentially Affine Yes or No? An Empirical Analysis of S&P500 Dynamics. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1344226.

[38] Jacquier, E., Polson, N.G., & Rossi, P.E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *J. Econom.,* 122: 185–212.

[39] Jarrow, R.A., & Larsson, M. (2012). The meaning of market efficiency. *Math. Financ.,* 22, 1–30.

[40] Jensen, M.J. (1999). Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. *J. Forecast.,* 18, 17–32.

[41] Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H.E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A,* 316, 87–114.

[42] Kendal, W.S., & Jørgensen, B. (2011). Tweedie convergence: A mathematical basis for Taylor's power law, 1/f noise, and multifractality. *Phys. Rev. E,* 84(6), 066120.

[43] Kinney, J.B. & Atwal, G.S. (2014). Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci*., 111, 3354–3359.

[44] Kristoufek, L., & Vosvrda, M. (2013). Measuring capital market efficiency: Global and local correlations structure. *Phys. A Stat. Mech. Its Appl.,* 392, 184–193.

[45] Kristoufek, L., & Vosvrda, M. (2014). Measuring capital market efficiency: long-term memory, fractal dimension and approximate entropy. *Eur. Phys. J. B,* 87(7), 162.

[46] Koutrouvelis, I.A. (1980). Regression-Type Estimation of the Parameters of Stable Laws. *J. Am. Stat. Assoc.*, 75(372), 918-928.

[47] Koutrouvelis, I.A. (1981). An iterative procedure for the estimation of the parameters of stable laws. *Commun. Stat.Simulat.*, 10 (1), 17-28.

[48] Li, Y., Peng, F., & Xu, H. (2012). Bayesian Testing for Asset Volatility Persistence on Multivariate Stochastic Volatility Models. *J. Math. Financ.*, 2(1), 83-89.

[49] Lillo, F., & Farmer, D.J. (2004). The Long Memory of the Efficient Market. *Stud. Nonlinear Dyn. Econom.*, 8, 1–33.

[50] Liu, X.B., & Li, Y. (2014). Bayesian testing volatility persistence in stochastic volatility models with jumps. *Quant. Financ.*, 14(8), 1415-1426.

[51] Liu, Y., Liang, X.S., & Weisberg, R.H. (2007). Rectification of the bias in the wavelet power spectrum. *J. Atmos. Ocean. Technol*., 24, 2093–2102.

[52] Mandelbrot, B.B., & Van Ness, J.W. (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Rev.* 10, 422–437.

[53] Matos, J.A.O., Gama, S.M.A., Ruskin, H.J., Al Sharkasi, A., & Crane, M. (2008). Time and scale Hurst exponent analysis for financial markets. *Phys. A Stat. Mech. Its Appl.*, 387, 3910–3915.

[54] McCulloch, J.H. (1986). Simple consistent estimators of stable distribution parameters. *Commun. Stat. - Simul. Comput.*, 15, 1109–1136.

[55] Messow, P., & Krämer, W. (2013). Spurious persistence in stochastic volatility. *Econ. Lett.*, 121, 221–223.

[56] Montanari, A., Taqqu, M.S., & Teverovsky, V. (1999). Estimating long-range dependence in the presence of periodicity: An empirical study. *Math. Comput. Model.*, 29, 217–228.

[57] Ng, E.K.W., & Chan, J.C.L. (2012). Geophysical applications of partial wavelet coherence and multiple wavelet coherence. *J. Atmos. Ocean. Technol.*, 29, 1845–1853.

[58] Paninski, L. (2003). Estimation of Entropy and Mutual Information. *Neural Comput.*, 15, 1191–1253.

[59] Patton, A., & Sheppard, K. (2013). Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. Economic Research Initiatives at Duke (ERID) Working Paper, 168 (November 16). Available at http://ssrn.com/abstract=1943825.

[60] Poterba, J.M., & Summers, L.H. (1986). The Persistence of Volatility and Stock Market Fluctuations. *Am. Econ. Rev.*, 76, 1142–1151.

[61] Rea, W., Oxley, L., Reale, M., & Brown, J. (2013). Not all estimators are born equal: The empirical properties of some estimators of long memory. *Math. Comput. Simul.*, 93, 29–42.

[62] Reisen, V.A. (1994). Estimation of the fractional difference parameter in the ARIMA(p,d,q) model using the smoothed periodogram. *J. Time Ser. Anal.*, 15, 335–350.

[63] Reisen, V.A., Abraham, B., & Lopes, S. (2001). Estimation of Parameters in ARFIMA Processes: A Simulation Study. *Commun. Stat. Simulat*., 30(4), 787-803.

[64] Reisen, V.A., Abraham, B., & Toscano, E.M.M. (2001). Parametric and semiparametric estimations of stationary univariate ARFIMA model. *Braz. J. Probab. Stat.*, 14, 185–206.

[65] Robinson, P.M. (1995). Gaussian semiparametric estimation of long run dependence. *Ann. Stat.*, 23(5), 1630-1661.

[66] Said, S.E., & Dickey, D.A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71, 599–607.

[67] Samoradnitsky, G., & Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*. New York: Chapman and Hall.

[68] Schwert, G.W. (1989). Tests For Unit Roots: A Monte Carlo Investigation. *J. Bus. Econ. Stat*., 2, 147–159.

[69] Taqqu, M.S., Teverovsky, V.T, & Willinger, W. (1995). Estimators for long-range dependence: an empirical study. *Fractals*, 03, 785–798.

[70] Todorov, V., & Tauchen, G. (2011). Volatility Jumps. *J. Bus. Econ. Stat*., 29, 356–371.

[71] Veillette, M. (2012). STBL: Alpha stable distributions for MATLAB, MATLAB code at: http://www.mathworks.com/matlabcentral/fileexchange/37514-stbl--alpha-stable-distributions-for-matlab.

[72] Vo, M. (2011). Oil and stock market volatility: A multivariate stochastic volatility

perspective, *Energy Econ.*, 33, 956–965.

[73] Yong, L., & Zhang, J. (2014). Bayesian testing for jumps in stochastic volatility models with correlated jumps. *Quant. Financ.*, 14(10), 1693-1700.

[74] Zeileis, A., Leisch, F., Hornik, K., Kleiber, C., Hansen, B., & Merkle, E.C. (2015). Testing, Monitoring, and Dating Structural Changes, R package 'strucchange' version 1.5-1. Available at: http://cran.r-project.org/web/packages/strucchange/strucchange.pdf.

[75] Zeileis, A., Kleiber, C., Walter, K., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Comput. Stat. Data Anal.*, 44, 109–123.

[76] Zeileis, A., Shah A., & Patnaik, I. (2010). Testing, monitoring, and dating structural changes in exchange rate regimes. *Comput. Stat. Data Anal.*, 54, 1696–1706.

[77] Zivot, E., & Andrews, D.W.K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *J. Bus. Econ. Stat.*, 10(3), 251–270.

[78] Zolotarev, V.M. (1986). One-dimensional stable distribution. *Transl. Math. Monographs*, vol. 65, Amer. Math. Soc., Providence.

[79] Whitcher, B., & Jensen, M.J. (2000). Wavelet estimation of a local long memory parameter. *Explor. Geophys.*, 31, 094-103.

[80] World Bank (2015). World Development Indicators, World Data Bank. Available at: http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators, accessed on: 27.06.2015.

[81] Wuertz, D. (2013). ARMA Time Series Modelling. Environment for teaching „Financial Engineering and Computational Finance", R package 'fArma' available at: http://cran.r-project.org/web/packages/fArma/fArma.pdf.

[82] Wuertz D. (2014). fUnitRoots: Trends and Unit Roots. Environment for teaching "Financial Engineering and Computational Finance", R Package, version 3010.78, available at: http://cran.r-project.org/web/packages/fUnitRoots/index.html.

Tables and figures

Table 1
Market capitalization of listed companies for some Central and Easter European countries (% of GDP; averages 1997-2010).

Bulgaria	13.433
Czech Republic	22.030
Hungary	24.460
Poland	25.076
Romania	12.017
Slovak Republic	5.594
Slovenia	21.072

Source of data: Word Bank (2015)

Table 2
Main statistics for BET index

	Level	Log-returns
Mean	4038.996	0.021
Median	4562.320	0.021
Maximum	10813.590	4.588
Minimum	281.090	-5.697
Std. Dev.	2762.631	0.744
Skewness	0.151	-0.336
Kurtosis	1.886	10.123
Jarque-Bera	241.507	9283.154
Probability	0.000	0.000
Zivot-Andrews Unit Root Test (Null Hypothesis: BET index has a unit root with a structural break in both the intercept and trend)	-4.366 (p=0.000)	-49.307 (p=0.001)
Tests of normal distributions		
Lilliefors (D)	0.126 (p=0.000)	0.089 (p=0.000)
Cramer-von Mises (W2)	15.797 (p=0.000)	14.335 (p=0.000)
Watson (U2)	15.764 (p=0.000)	14.307 (p=0.000)
Anderson-Darling (A2)	105.728 (p=0.000)	82.217 (p=0.000)

Note: the probabilities for the Zivot-Andrews test are calculated from a standard t-distribution and do not take into account the breakpoint selection process.

Table 3
Posterior means, standard deviations and quantiles of model parameters

Parameter	Posterior mean	Posterior standard deviation	5% quantile	95% quantile
μ	0.030	0.008	0.017	0.042
μ_h	-1.345	0.126	-1.553	-1.139

ϕ_h	0.962	0.007	0.949	0.974
$\sigma_h^{\ 2}$	0.087	0.016	0.061	0.115
ψ	0.133	0.016	0.108	0.159
ν	27.271	10.665	13.040	46.598

Table 4

Main statistics for time-varying standard deviation $\exp(h_t/2)$

	Level	First order differences
Mean	0.591	0.000
Median	0.516	0.000
Maximum	2.593	0.194
Minimum	0.161	-0.227
Std. Dev.	0.317	0.031
Skewness	1.742	0.359
Kurtosis	7.421	9.710
Jarque-Bera	5745.347	8256.870
Probability	0.000	0.000
Zivot-Andrews Unit Root Test (Null Hypothesis: Volatility has a unit root with a structural break in both the intercept and trend)	-9.198 (p=0.000)	-24.916 (p=0.026)
Tests of normal distributions		
Lilliefors (D)	0.118 (p=0.000)	0.099 (p=0.000)
Cramer-von Mises (W2)	23.880 (p=0.000)	16.882 (p=0.000)
Watson (U2)	18.475 (p=0.000)	16.768 (p=0.000)
Anderson-Darling (A2)	142.357 (p=0.000)	94.345 (p=0.000)
Parameters of an alfa-stable distribution:		
Koutrouvelis (1980,1981):		
α	1.586	
β	1	
γ	0.155	
δ	0.665	
McCulloch (1986):		
α	1.622	
β	1	
γ	0.167	
δ	0.591	

Note: The probabilities for the Zivot-Andrews test are calculated from a standard t-distribution and do not take into account the breakpoint selection process.

Table 5

Various unit root tests for S&P 500 series (levels of returns)

Test	Value	Conclusion
Augmented Dickey–Fuller	-9.959	Reject the unit root null
Elliott–Rothenberg–Stock (DF-GLS)	-6.785	Reject the unit root null
Elliott–Rothenberg–Stock (feasible point optimal test)	0.226	Reject the unit root null
KPSS (mu-test)	1.362	Accept the stationarity null
KPSS (tau-test)	0.680	Accept the stationarity null
Phillips–Perron (Z-alpha)	-78.340	Reject the unit root null
Phillips–Perron (Z-tau)	-6.260	Reject the unit root null
Schmidt–Phillips (tau-test)	-11.957	Reject the unit root null
Schmidt–Phillips (rho-test)	-285.405	Reject the unit root null

Table 6

Hurst parameter estimates for time-varying standard deviation $\exp\left(h_t / 2\right)$

Aggregated Variance Method	0.863
Differenced Aggregated Variance Method	0.859
Aggregated Absolute Value/Moment Method	0.916
Higuchi or Fractal Dimension Method	0.986
The R/S Method	0.944
Hurst parameter based on Whittle Estimator for Fractional Gaussian Noise / Fractional ARIMA	0.990
Generalized Hurst Exponent (DiMatteo, 2007)($q=1$)	0.803
Generalized Hurst Exponent (DiMatteo, 2007)($q=2$)	0.792
Generalized Hurst Exponent (DiMatteo, 2007)($q=3$)	0.775

Table 7

The fractional ("memory") parameter *d* in an ARFIMA(p,d,q) model for time-varying standard deviation $\exp\left(h_t / 2\right)$

	The bandwidth used in the regression equation			
	0.30	0.40	0.50	0.60
Geweke and Porter-Hudak (1983) estimator	0.633 (0.256)	0.476 (0.146)	0.399 (0.089)	0.568 (0.055)
Reisen (1994); Reisen et al. (2001) estimator	0.665 (0.096)	0.425 (0.055)	0.408 (0.033)	0.591 (0.021)

Notes: Asymptotic standard deviation in (). For Reisen estimator, the parameter β (the exponent of the bandwidth used in the lag Parzen window) is set equal with 0.90.

Fig. 1. Log-returns of BET index.

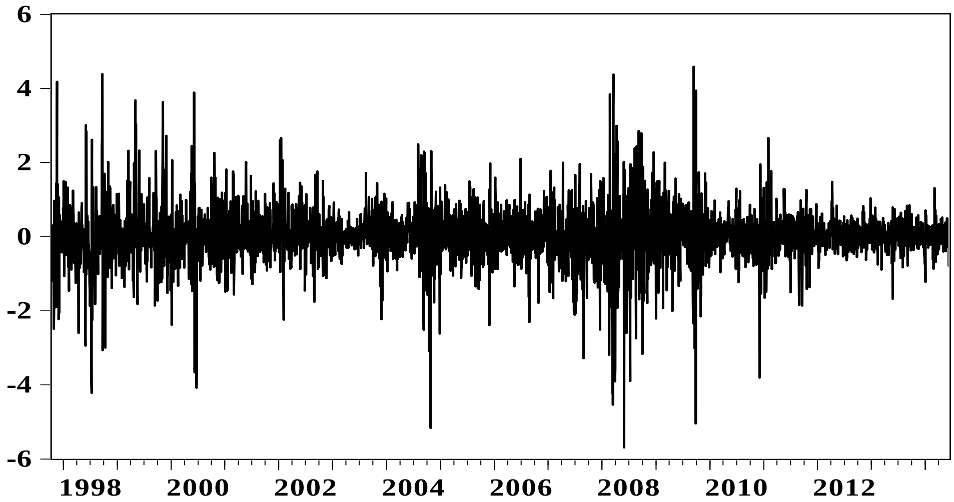


Fig. 2. Estimates of $p(\psi|y)$ (left panel) and $p(\nu|y)$ (right panel)

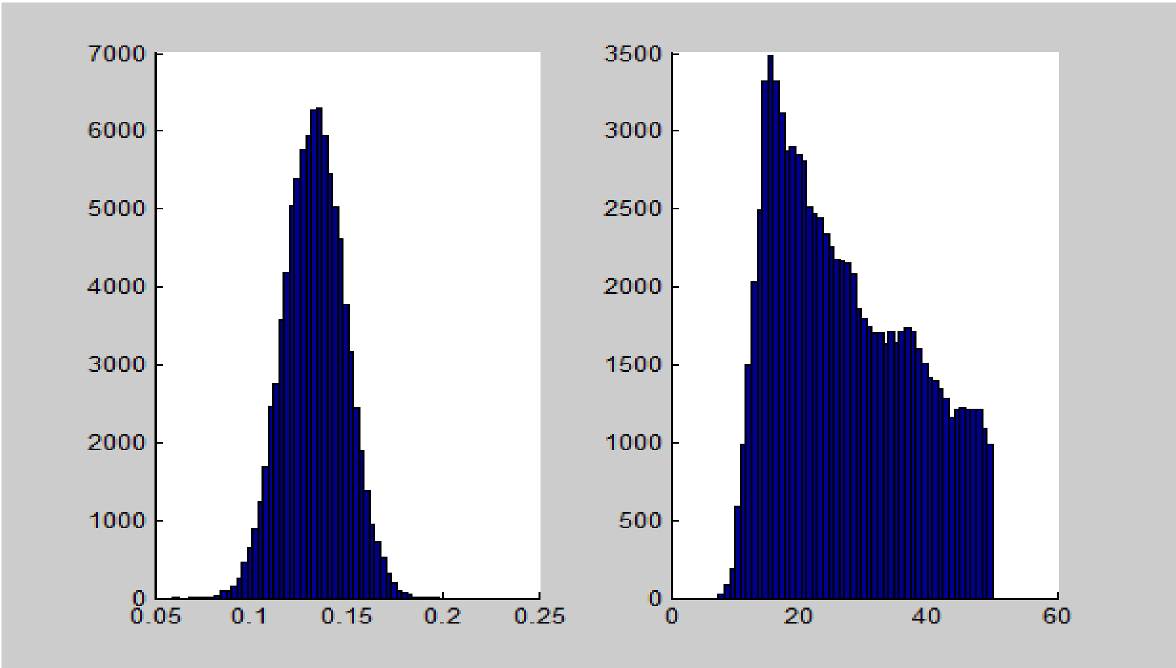


Fig. 3. Posterior means (solid line) and 90% credible intervals (dash lines) of the time-varying standard deviation $\exp\left(h_t / 2\right)$

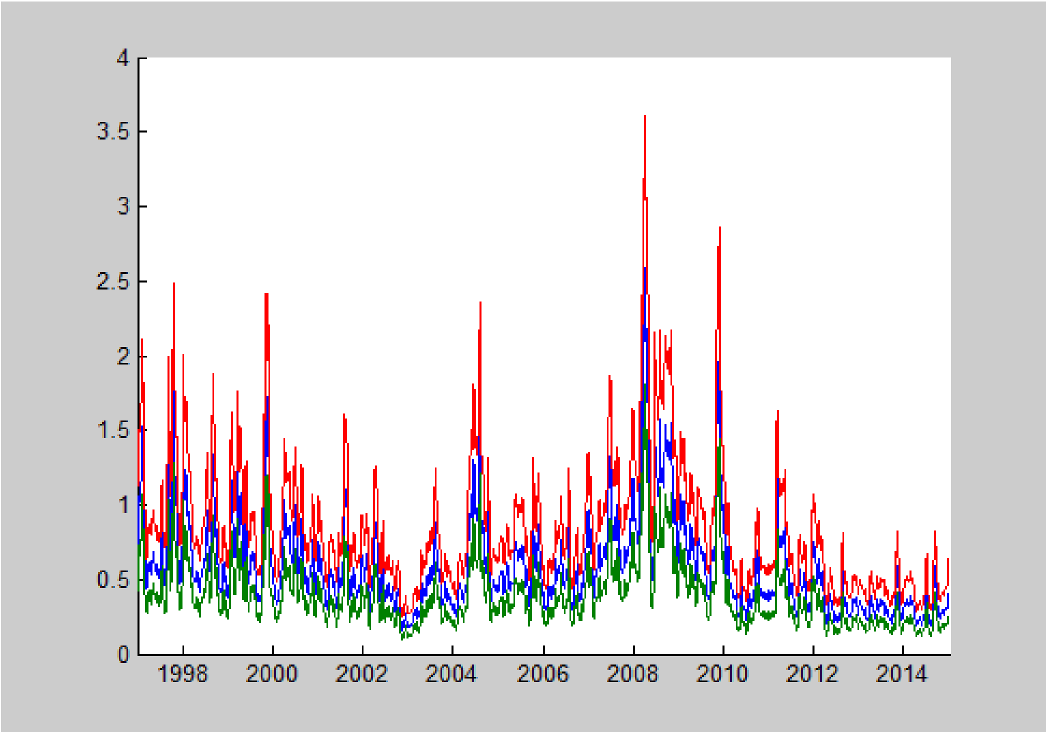
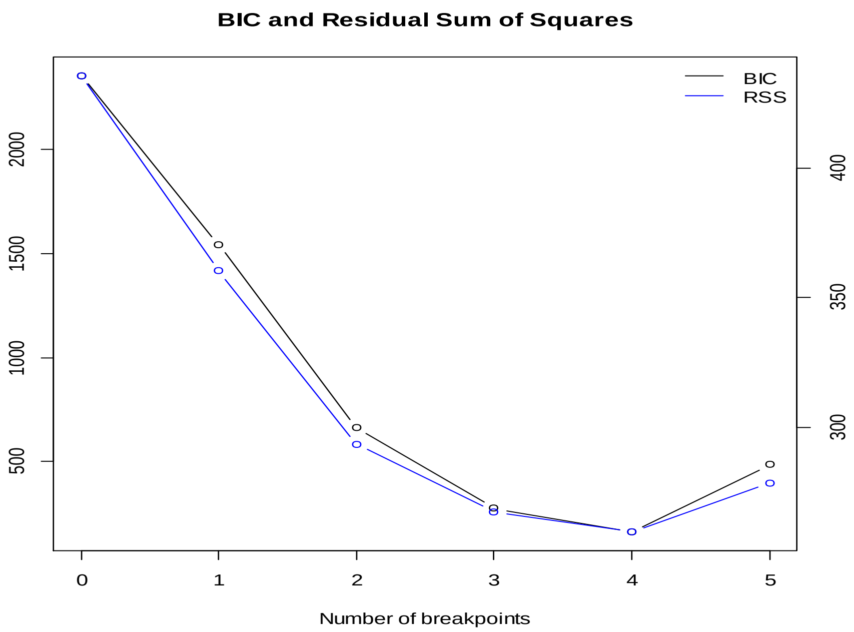


Fig. 4. The Bai and Perron test for the optimal number of structural breaks in volatility



Note: The minimal segment size in which the regression coefficients are constant: 653 observations (15% of the data).

Fig.5. Tweedie distribution of the time-varying standard deviation $\exp\left(h_t/2\right)$

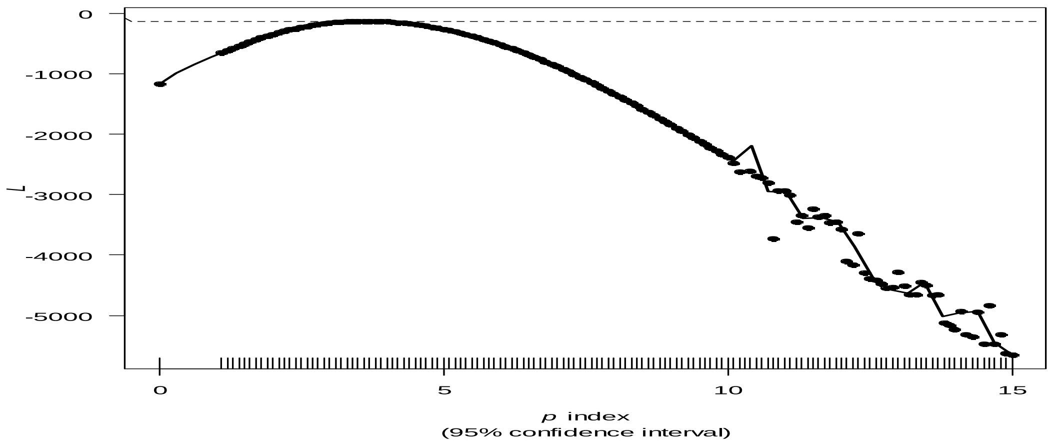


Fig. 6. Multifractal Detrended Fluctuation Analysis of the time-varying standard deviation $\exp\left(h_t/2\right)$ - q-order Hurst exponent

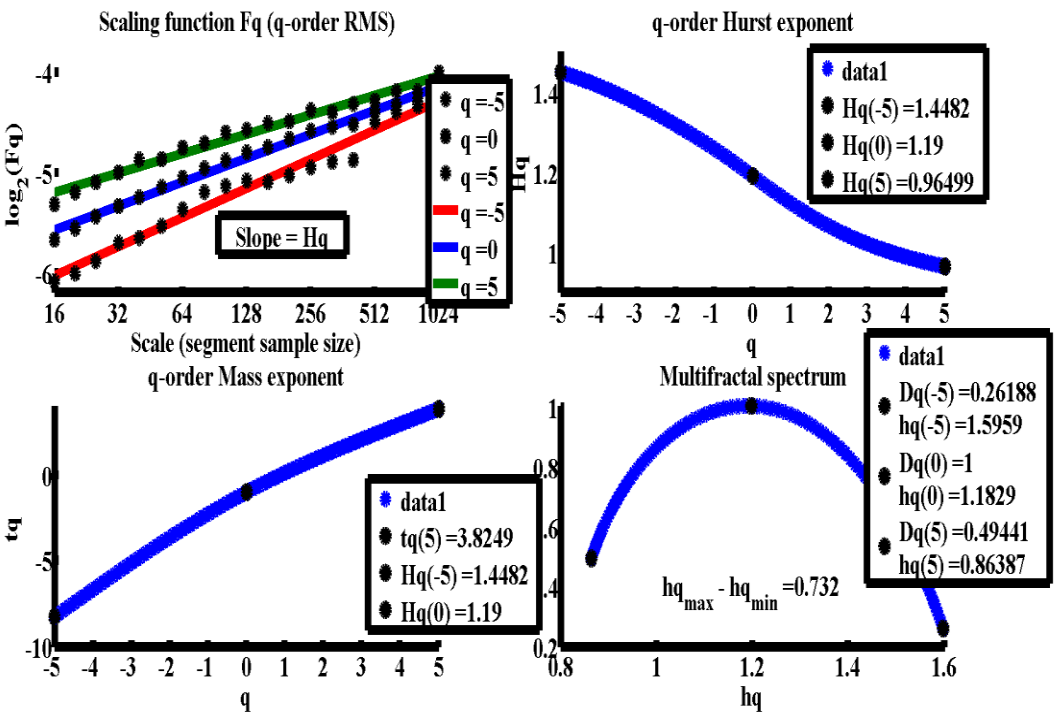


Fig. 7. Multifractal Detrended Fluctuation Analysis of the time-varying standard deviation $\exp(h_t/2)$ - local Hurst exponent

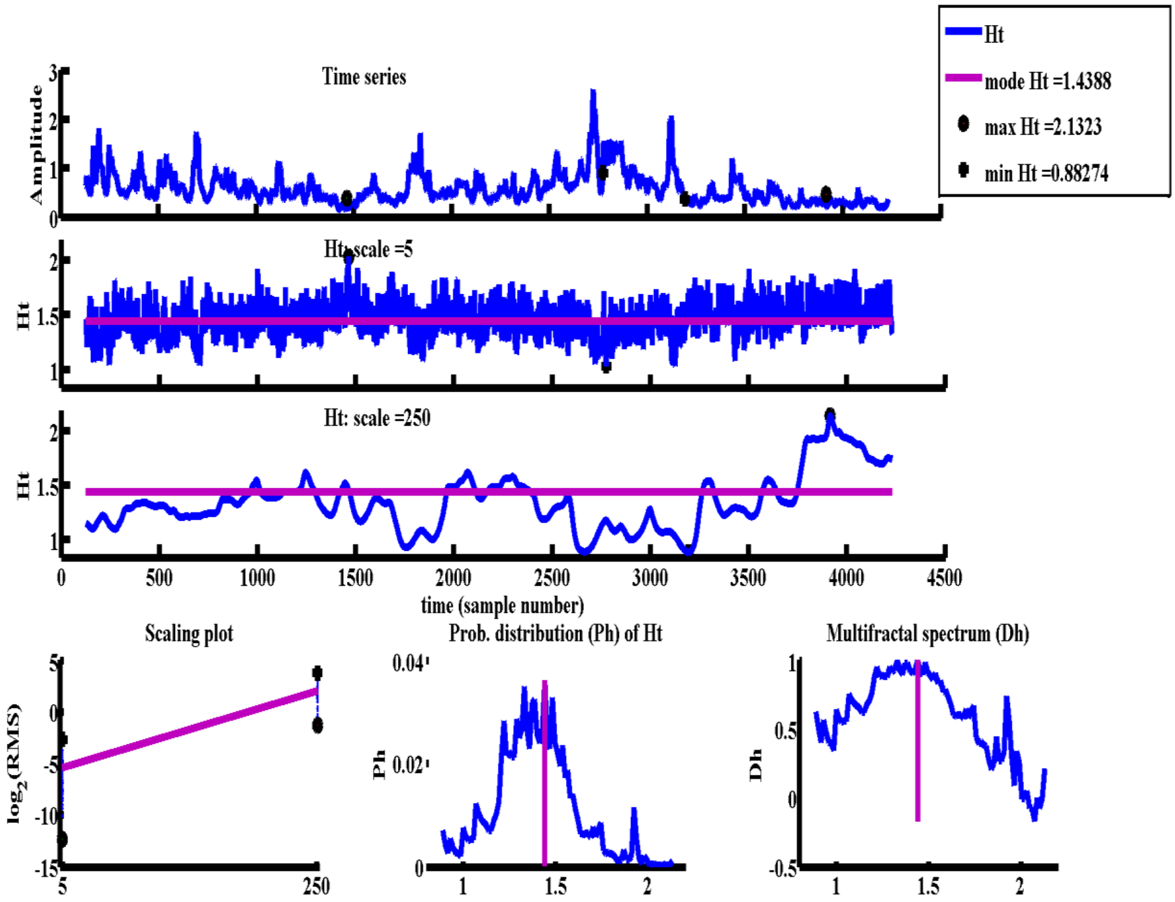
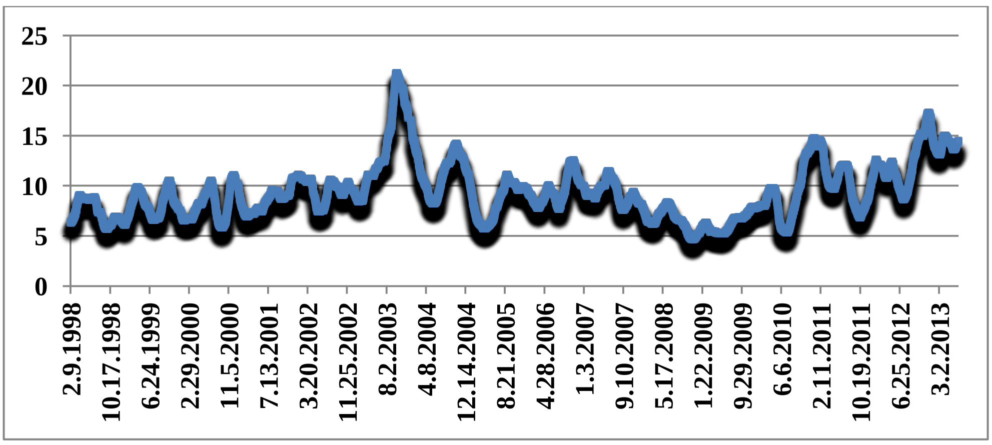
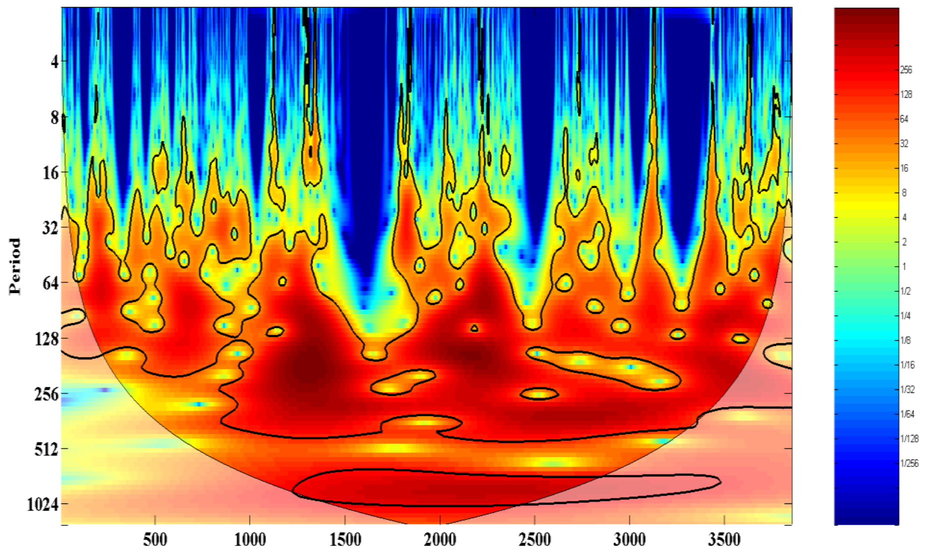


Fig.8. Nonparametric estimation of Shannon's index of diversity by Chao and Shen method for the time-varying standard deviation $\exp(h_t/2)$

(a) Estimator

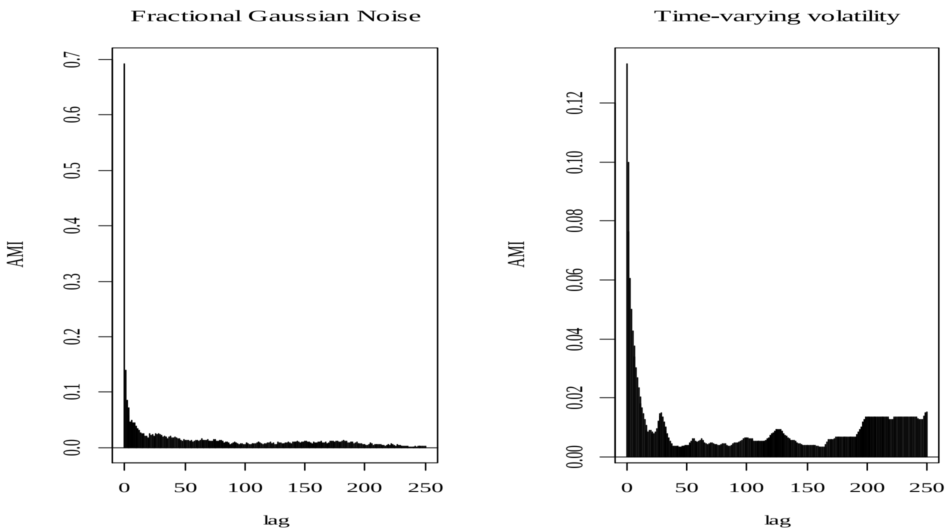


(b) Rectified wavelet power spectrum of estimator



Notes: Power spectrum rectified according to Liu et al. [51] – The thick black contour designates the 5% significance level against red noise which is estimated from Monte Carlo simulations using phase-randomized surrogate series. The cone of influence, which indicates the region affected by edge effects, is also shown with a light black line.

Fig. 9. The average mutual information index (AMI) for time-varying standard deviation $\exp\left(h_t / 2\right)$



Note: Number of bins is equal with 2.

Fig. 10. Average squared differences from 10000 fGn simulations of mutual information

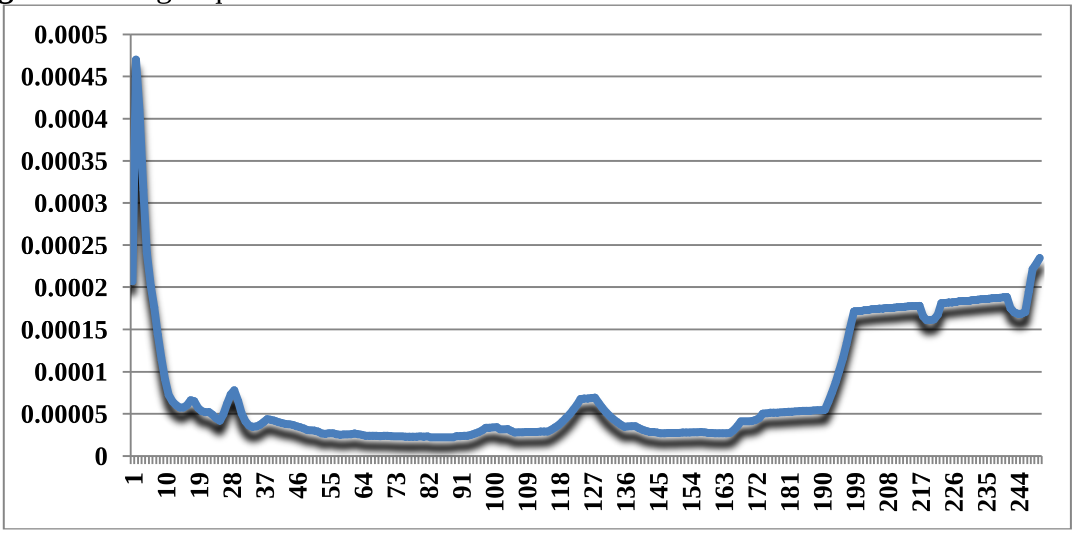


Fig. 11. The minimal mutual information on overlapping windows