The European Journal of Finance

Routledge
Taylor & Francis Group

On the statistics of scaling exponents and the Multiscaling Value at Risk

Journal:	*The European Journal of Finance*
Manuscript ID	REJF-2020-0043.R1
Manuscript Type:	Special Issue
Keywords:	Multiscaling time series, Robust scaling, Value at Risk, Model Construction and Estimation < Econometric Modeling, Hypothesis Testing < Econometric and Statistical Methods: General, Financial Forecasting < General Financial Markets, Long-memory

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

StatsHurst_final_a_v2.bbl StatsHurst_final_a_v2.aux StatsHurst_final_a_v2.tex

SCHOLARONE™
Manuscripts

On the statistics of scaling exponents and the Multiscaling Value at Risk

ARTICLE HISTORY

Compiled January 6, 2021

ABSTRACT

Scaling and multiscaling financial time series have been widely studied in the literature. The research on this topic is vast and still flourishing. One way to analyze the scaling properties of time series is through the estimation of their scaling exponents, which are recognized as being valuable measures to discriminate between random, persistent, and anti-persistent behaviors in these time series. In the literature, several methods have been proposed to study the multiscaling property. In this paper, we use the generalized Hurst exponent (GHE) tool and we propose a novel statistical procedure based on GHE which we name Relative Normalized and Standardized Generalized Hurst Exponent (RNSGHE). This methodology is used to robustly estimate and test the multiscaling property and, together with a combination of t-tests and F-tests, serves to discriminate between real and spurious scaling estimates. Furthermore, we introduce a new tool to estimate the optimal aggregation time used in our methodology which we name Autocororrelation Segmented Regression. We numerically validate this procedure on simulated time series by using the Multifractal Random Walk (MRW) and we then apply it to real financial data. We present results for times series with and without anomalies and we compute the bias that such anomalies introduce in the measurement of the scaling exponents. We also show how the use of proper scaling and multiscaling can ameliorate the estimation of risk measures such as Value at Risk (VaR). Finally, we propose a methodology based on Monte Carlo simulation, which we name Multiscaling Value at Risk (MSVaR), that takes into account the statistical properties of multiscaling time series. We mainly show that by using this statistical procedure in combination with the robustly estimated multiscaling exponents, the one year forecasted MSVaR mimics the VaR on the annual data for the majority of the stocks analyzed.

1.   Introduction

Nowadays, scaling and multiscaling are widely accepted as empirical stylized facts in financial time series. Since they provide important information to risk and asset managers, they need to be properly addressed and analyzed. The (multi)scaling property of time series is particularly important in risk management and has been recently employed as a warning tool for financial events (Antoniades et al. 2020). In particular, models that implicitly or explicitly assume independence of asset returns should be tested against long-term dependence alternatives. In fact, if the assumption of independence of price increments is not met, risk measures might be severely biased, especially if the long-range dependence is acting with a different degree across the time series statistical moments. In particular, multiscaling has been adopted as a formalism in two different branches of quantitative Finance, i.e. econophysics and mathematical Finance. The former devoted most of the attention to price and returns series in order to understand the source of multifractality from an empirical and theoretical point of view (Mandelbrot 1963, 1967; Mantegna and Stanley 1999; Gençay et al. 2001; Mantegna and Stanley 1995; Di Matteo 2007; Calvet and Fisher 2002; Lux 2004; Lux and Marchesi 1999; Di Matteo, Aste, and Dacorogna 2005; Buonocore et al. 2019) and has recently identified a new stylized fact which relates (non-linearly) the strength of multiscaling and the dependence between stocks (Buonocore et al. 2019). The latter instead, builds on the work of (Gatheral, Jaisson, and Rosenbaum 2018) on rough volatility and has been used to construct stochastic models with anti-persistent volatility dynamics (Gatheral, Jaisson, and Rosenbaum 2018; Takaishi 2020; Fukasawa, Takabatake, and Westphal 2019; Livieri et al. 2018). Even if the research question comes from different perspectives, it is important to recognize the relevance that its study has in Finance.

2

Multiscaling has been understood to originate from one or more phenomenon related to trading dynamics.[1] In particular, it can be attributed to the fat tails, the autocorrelation of the absolute value of log-returns, liquidity dynamics, or (non-linear) correlation between high and low returns generated by the different time horizons at which traders operate and the consequent volumes traded. It can also be caused by the endogeneity of markets for which a given order generates many other orders. This occurs especially in markets where algorithmic trading is frequently adopted. There are different methodologies used to compute scaling exponents from time series (Jiang et al. 2019). Among all, let us recall the Multifractal Detrended Fluctuation Analysis (MFDFA) proposed in (Kantelhardt et al. 2002), the Wavelet Transform Modulus Maxima (WTMM) introduced by (Muzy, Bacry, and Arneodo 1991, 1993) and the Structure function approach also known as the Generalized Hurst exponent (GHE) method (Van Atta and Chen 1970; Kolmogorov 1962; Di Matteo, Aste, and Dacorogna 2003; Di Matteo 2007). In a recent paper, (Barunik and Kristoufek 2010) tested different methodologies against some data specification and empirically showed that the GHE approach outperforms the other models. For this reason, throughout this work, we will use the GHE approach. Notwithstanding the importance of the correct estimation of the Hurst exponent, the analysis has been rarely addressed from a statistical point of view.

In this paper, we propose a step-by-step procedure that provides a robust estimation and that tests the multiscaling property in a statistically significant way. Application to simulated data and empirical data allows us also to demonstrate the impact of bias on these estimations. We show how the use of proper scaling and multiscaling can ameliorate the estimation of risk measures such as Value at Risk(VaR). We also propose a methodology based on Monte Carlo simulation, which we name Multiscaling VaR, which takes into account the statistical properties of multiscaling time series by using a multiscaling consistent data generating process.

The paper is structured as follows. Sections 2 and 3 provide a brief description of multiscaling in Finance and of the statistical procedure proposed to consistently estimate and test the scaling spectrum. Section 4 shows the results of this methodology

---

[1] We refer as trading dynamics the results of the set of actions undertaken by investors in buying and selling financial instruments.

3

applied to synthetic data while Section 5 reports the results of an empirical application to real financial time series. Section 6 is devoted to a practical application of scaling and multiscaling property to VaR while Section 7 concludes.

2. Multiscaling in Finance

In this section, we explain the importance of the multifractal (multiscaling) formalism in financial markets. Let us first fix the notation by defining the prices time series as P_t and the log-prices $p_t = \ln(P_t)$. From this, the log-returns over a time aggregation τ are $r_\tau(t) = p_{(t+\tau)} - p_t$, where τ is expressed in days. Financial models are usually based on the assumption that log-prices follow a Brownian Motion and the for this model, the rescaled second moment of the log-returns over time aggregation τ follows

$$\sigma_\tau = \mathbb{E}[|r_\tau(t)|^2]^{\frac{1}{2}} \sim \sigma \tau^{\frac{1}{2}}, \tag{1}$$

where σ_τ is the standard deviation at aggregation horizon τ while σ is the standard deviation at daily aggregation. This equation is usually referred to as the *square root of time rule* and it is widely applied in quantitative Finance (Danielsson and Zigrand 2006; Wang, Yeh, and Cheng 2011). Examples are the Black and Scholes model in which the volatility evolves as $\sigma \tau^{\frac{1}{2}}$, or the VaR which under Basel regulatory framework can be computed for higher time aggregation, e.g. the τ days VaR can be computed as the daily VaR multiplied by $\tau^{\frac{1}{2}}$. In the analysis of the Nile river, Hurst found that the scaling behaviour described by a Brownian Motion was not in line with the empirical data (Hurst 1956). Scaling and multiscaling analyses have been later introduced in Finance (Mandelbrot 1963, 1967, 2013; Jiang et al. 2019). To detect multiscaling, it is necessary to study the non-linearity of the scaling exponents of the q-order moments of the absolute value of log-returns (Mandelbrot, Fisher, and Calvet 1997; Calvet, Fisher, and Mandelbrot 1997; Di Matteo 2007). In particular, for a process (p_t) with stationary increments, the GHE methodology considers a function

4

of increments (Di Matteo 2007) of the form

$$\Xi(\tau, q) = \mathbb{E}\left[|r_\tau(t)|^q\right] \sim K_q \tau^{qH_q}, \tag{2}$$

where $q = \{q_1, q_2, \ldots, q_M\}$ is the set of evaluated moments, $\tau = \{\tau_1, \tau_2, \ldots, \tau_N\}$ is the set of lags used to compute the log-returns, N and M are the maximum numbers of moments and lag specification, i.e. $q_1 = q_{min}$, $q_M = q_{max}$, $\tau_1 = \tau_{min}$ and $\tau_N = \tau_{max}$, K_q is the q-moment for $\tau = 1$, and H_q is the so called generalized Hurst exponent which is a function of q. Finally, the function qH_q is concave (Mandelbrot, Fisher, and Calvet 1997; Calvet, Fisher, and Mandelbrot 1997) and codifies the scaling exponents of the process. A process is uniscaling when the function H_q does not depend on q, i.e. $H_q = H$ (Di Matteo 2007), while it is multiscaling otherwise. If $H \neq 0.5$, the process does not behave as a standard Brownian Motion (Wiener process) and neglecting this feature, would significantly bias the estimation of the true risk. In particular, if $H < 0.5$ ($H > 0.5$) the process is said to be anti-persistent (persistent) while if $H = 0.5$ the process can be of two types, i.e. it can have independent increments or it can be a short-term dependent process (Lillo and Farmer 2004; Kristoufek 2010). Given Equation 2, a possible way to define a multiscaling proxy is by quantifying the degree of non-linearity of the function qH_q. The standard procedure used in order to extract qH_q consists in running a linear regression in log-log scale of Equation 2, which reads as

$$\ln(\Xi(\tau, q)) = qH_q \ln(\tau) + \ln(K_q), \tag{3}$$

where τ is defined in the range $\tau = [\tau_{min}, \tau_{max}]$ and $q = [q_{min}, q_{max}]$ (Di Matteo 2007). A multiscaling proxy can be obtained by fitting the measured scaling exponent with a second degree polynomial (Buonocore, Aste, and Di Matteo 2016; Buonocore et al.

5

2019) of the form[2]

$$qH_q = Aq + Bq^2, \tag{4}$$

where A and B are two constants. In this mathematical setting, as for different multifractal models in Finance (Bacry, Delour, and Muzy 2001b; Calvet and Fisher 2002, 2004; Jiang et al. 2019), we implicitly assume a quadratic function of qH_q. The measured B, $\widehat{B}$, represents the curvature of qH_q. If $\widehat{B} = 0$, the process is uniscaling, while if $\widehat{B} \neq 0$, the process is multiscaling (Buonocore, Aste, and Di Matteo 2016; Buonocore et al. 2019). In order to widely apply the multiscaling formalism in Finance, it is of vital importance the ability to correctly estimate the value of qH_q and consequently, of A and B.

3. Methodology

As highlighted in the previous section, estimating the Hurst exponent from empirical data is a challenging task and these challenges can be categorized in two different classes:

- Those due to the statistical procedure adopted;
- Those linked to the financial data themselves.

Within the first class, we identify two main issues related to the following two points:

- The statistical model used to compute the scaling exponents;
- The input variables used in the statistical procedure.

Within the second class, issues arise mainly from the following question:

- If the data contain an anomaly, how is this impacting the estimation of the scaling and multiscaling exponents?

In this work, we address the above challenges. First, we focus on the statistical

---

[2]Technical details of the choice of this functional form can be found in (Buonocore, Aste, and Di Matteo 2016; Buonocore et al. 2019).

6

procedure and the implication on financial time series with and without anomalies. Then, we discuss practical implications related to Finance, with special attention to Value at Risk.

3.1. *Statistical procedure*

Multiscaling properties of financial time series have been understood to come from one or more phenomena related to trading dynamics. From the point of view of the financial microstructure, scaling can be attributed to liquidity dynamics, endogeneity of markets, or any other dynamic existing in the market. In particular, the superimposition of distinct strategies and investment horizons generates long-range dependence with different degrees of strength when evaluated at different order moments and this is precisely the definition of multiscaling. In this section we propose a methodology to estimate the Hurst exponent qH_q and the multiscaling depth (curvature) coefficient B in a robust manner. As specified in Equation 3, the estimation of scaling laws is generally performed through a linear regression in log-log scales. The statistical problem which might arise in this context is that the regression is performed minimizing the squared log-errors instead of the true errors. This procedure might, in case of strong deviation from the assumed statistical model for the errors, severely impact the results. The solution to this problem consists of applying a nonlinear regression to the original (i.e. not transformed) data, comparing the fit of the two specifications to the original data and using the one which performs better. Another issue related to the statistical model is the uncertainty associated to the intercept for the q regressions. In particular, we can exactly compute the value of K_q rather than estimating it, thus eliminating possible errors and bias. We can define the standardized $\Xi(\tau, q)$ as

$$\widetilde{\Xi}(\tau, q) = \frac{\Xi(\tau, q)}{K_q}, \tag{5}$$

based on which, Equation 2 can be rewritten as

$$\widetilde{\Xi}(\tau, q) \sim \tau^{qH_q}. \tag{6}$$

7

Equation 6 eliminates the possible bias introduced by the estimation of K_q via regression. To easily exploit and model the multiscaling behavior, we define the q-order normalized moment as

$$\dddot{\Xi}(\tau, q) = \widetilde{\Xi}(\tau, q)^{\frac{1}{q}} \tag{7}$$

which transforms Equation 6 in

$$\dddot{\Xi}(\tau, q) \sim \tau^{H_q}. \tag{8}$$

Within this new formulation, the analysis is much easier since now, all the q regressions have a 0 intercept and the multiscaling is present only if the regression coefficients H_q differ for distinct values of q. In fact, for uniscaling time series all regression lines are overlapping while, for multiscaling time series they diverge. Given the formalism introduced by Equation 8, it is easy to check whether a process is multiscaling or not. In addition, we can now rewrite Equation 4 for the normalized and standardized structure function of Equation 8 as

$$H_q = A + Bq. \tag{9}$$

Even if mathematically equivalent to Equation 4, this equation has a statistical advantage. Eliminating the multiplication by q from both sides of the Equation, we reduce the possibility of spurious results in case q is a dominant factor in the multiplication. Indeed, the interpretation is equivalent, i.e. A is the linear scaling index while B is the multiscaling proxy. Finally, let us define the relative structure function between two consecutive moments, namely q_i and q_j ($q_j > q_i$), as follows

$$\dddot{\Xi}(\tau, q_i, q_j) = \frac{\dddot{\Xi}(\tau, q_j)}{\dddot{\Xi}(\tau, q_i)} \sim \frac{\tau^{H_{q_j}}}{\tau^{H_{q_i}}} = \tau^{H_{q_j} - H_{q_i}} = \tau^{H(q_i, q_j)}, \tag{10}$$

8

Where $H(q_i, q_j) = H_{q_j} - H_{q_i}$. This formalization has a similar structure as the Extended Self Similarity (ESS) methodology (Jiang et al. 2019) since the scaling exponents are computed taking as reference another moment function while it diverges from it as the ESS has a reciprocal effect while Equation 10 has an incremental effect relative to the reference moment. This approach helps in the statistical analysis since we can now test if a process is statistically multiscaling using a significance test on the estimated $H(q_i, q_j)$. In fact, for uniscaling time series we have that $H_q = H$, which implies that the difference between different order moments is always 0.[3] On the contrary, for multiscaling time series it should be different from 0 for all q. This reduces to a t-test on the regression coefficients estimated using Equation 10. Besides the multi-regression approach, it is possible to perform a multivariate regression by rewriting Equation 10 as

$$
\begin{bmatrix}
\ddot{\Xi}(\tau, 0, q_1) \\
\ddot{\Xi}(\tau, q_1, q_2) \\
\vdots \\
\ddot{\Xi}(\tau, q_{M-1}, q_M)
\end{bmatrix}
= \tau \begin{bmatrix} H_{(0,q_1)}, H_{(q_1,q_2)}, \cdots, H_{(q_{M-1},q_M)} \end{bmatrix},
\tag{11}
$$

where M is the maximum number of moments used. This is a multivariate nonlinear regression that can be easily solved via a nonlinear optimization algorithm. Such a methodological approach implies a possible relationship between the q-moments used in the regression. Depending on the model assumptions, one can use Equation 11 or perform M separate regressions for each exponent. In the first case, it is then possible to use an F-test to test if all the coefficients except for the first one ($H(0, q_1)$) are jointly equal to 0 against the alternative that some coefficients are different from 0. This is a less restrictive multiscaling test compared to the multiple t-tests. We call *strongly multiscaling processes* those processes which reject both the null hypothesis for all the t-tests and the null of the F-test. Conversely, we call *weakly multiscaling processes* those processes for which the null hypothesis of all the t-tests is rejected but

---

[3]Exluding the special case $H(0, q_1)$.

9

not the null of the F-test.[4] This is quite intuitive since if a process is multiscaling, all the relative increments are statistically significant. However, if the process reconstructed with a single exponent is statistically equivalent to the one reconstructed with the full multiscaling spectrum, this means that such multiscaling behavior is weak. As already mentioned, it is recommended to estimate the model both in the log-log scale and in the original coordinate system, and to base the choice of the model on a goodness of fit measure.

3.1.1. The choice of q and τ

The choice of q and τ is an important step in the statistical evaluation of the multiscaling exponents. They must be selected using specific statistical criteria. In fact, using the wrong values of q and τ can severely bias the evaluation. Regarding q, many research papers about multiscaling systems propose the use of a vast spectrum of qs. This approach has two fallacies. The first one lies in the fact that multiscaling processes are such even for small values of q. Secondly and most importantly, given a distribution of returns with tail exponent α, for $q \geq \alpha$, $\mathbb{E}[r^q]$ diverges (Jiang et al. 2019). For empirical data, this effect is characterized by a distortion of the moment function which can be misinterpreted as multiscaling even without the presence of temporal dependence. Hence, to have a robust measure of multiscaling, it is necessary to have $q < \alpha$. Any multiscaling behaviour found by neglecting or ignoring this fact, is severely biased and possibly false. The method used to set q can derive from two different approaches, i.e. established research results or direct tail exponent computation. Since it has been empirically shown that financial time series have fat tails with tail exponents ranging from ~ 1.5 (Scalas and Kim 2006; Eom, Kaizoji, and Scalas 2019; Weron 2001) to ~ 3 (Jiang et al. 2019), a conservative approach would be to use $q \leq 1$. Alternatively, it is possible to estimate α on the empirical distribution through a tail estimator (e.g. (Clauset, Shalizi, and Newman 2009; Virkar, Clauset et al. 2014)) and use it as threshold for q. In this paper, we follow the conservative approach and use $q \leq 1$. In fact, if the multiscaling phenomenon is present, it can be extrapolated from this range of

---

[4]If the null hypothesis for one or more t-tests is not rejected but the F-test rejects the null hypothesis, the process is a non-stable multiscaling process.

moments.

Regarding the time aggregation τ, a general rule would be to use the minimum possible value of τ, denoted as τ^*, such that the autocorrelation information of the series is preserved. The autocorrelation ρ of the return series at lag τ is defined as:

$$\rho_\tau(r(t)) = \frac{\mathbb{E}[(r(t+\tau) - \mu)(r(t) - \mu)]}{\sigma_\tau^2} \tag{12}$$

where μ and σ_τ^2 are respectively the mean and variance of $r(t)$. It is a well known stylized fact that returns are expected to be uncorrelated at daily frequency while the absolute and squared returns exhibit long-range persistence (Cont 2001; Chakraborti et al. 2011). Among the different procedures used to estimate τ^*, it is worth mentioning:

- Segmented regression on the structure function (Yue et al. 2017);
- Autocorrelation significance test (Buonocore, Aste, and Di Matteo 2017).

The first procedure computes the structure function for each q-moment and succesfully fits a segmented regression in log-log coordinates between τ and $\Xi(\tau, q)$, and finds two slopes: one for the scaling component and one for the non-scaling component (Yue et al. 2017). The second approach instead, chooses the value of τ prior to computing the structure function, setting the value of τ^* as the minimum value of τ for which the autocorrelation is not statistically significant (Buonocore, Aste, and Di Matteo 2017). In this paper, we propose a new approach which takes the advantages of both methods. We name this Autocorrelation Segmented Regression. The rationale behind this approach is to perform a segmented regression on the autocorrelation (or the autocovariance) function computed on the absolute returns and take τ^* as the value which minimizes the sum of squared residuals for the high autocorrelation state and the random noise state, i.e. plateau.[5] This approach has the advantage of setting the value of τ in advance, avoiding ad-hock solutions and reducing computations. Nevertheless, the method is less sensitive to a unique non-significant lag. In fact, in noisy data it can happen that for a lag the autocorrelation is not significant while it is significant for a considerable number of subsequent lags. The equation for the proposed Autocorrelation

---

[5](Valsamis, Husband, and Chan 2019) reviews different segmented regression specifications.

11

Segmented Regression (ACSR) takes the form

$$\rho_\tau(r(t)) = \begin{cases} \alpha + \beta\tau, & \text{if } \tau < \tau^* \\ \alpha + \beta\tau^*, & \text{if } \tau \geq \tau^* \end{cases} \tag{13}$$

where α is the intercept of the regression and can be fixed to be equal to ρ_τ with $\tau = 1$, β is a slope parameter for the autocorrelation function, τ is the lag at which the autocorrelation is computed, and τ^* is the value of aggregation which maximizes the autocorrelation information.[6] We use $\ln(\tau)$ instead of τ for a better detection of τ^*. Figure 1 shows how this method works. We generated a process with known $\tau_{max} = 250$ and run the ACSR to empirically estimate $\ln(\tau^*)$. As shown in Figure 1, we get a value of $\ln(\tau^*) = 5.51$ which corresponds to $\tau^* = 247$.
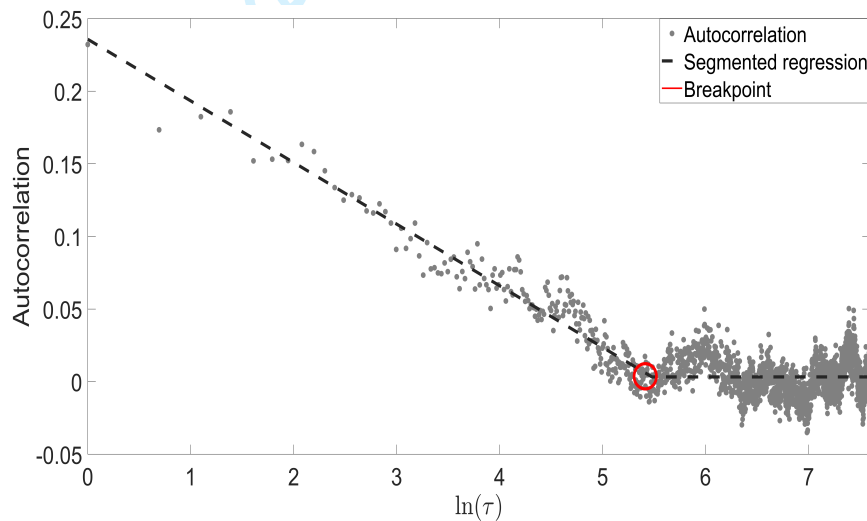
Figure 1. ACSR methodology computed on the autocorrelation function of the absolute values of the log-returns time series. The red circle shows the breakpoint where the regression line has a break.

3.1.2. Multiscaling estimation and testing procedure

Before turning the attention to the simulation experiment, let us recall the full procedure required to robustly extract the scaling exponents:

(1) Compute τ^* with the Autocorrelation Segmented Regression method;

---

[6]It is important to notice that the segmented regression in the structure function and the ACSR method yields similar results.

(2) Compute $q = \alpha$ or rely on the empirical evidence available in the literature;

(3) Perform the linear and nonlinear regressions with the above parameters (Equation 11);

(4) Asses the goodness of fit of the two models and select the one that overperforms;

(5) Compute the multiscaling curvature using Equation 9 and test for statistical significance.

Concerning point (5), in this paper we propose a full procedure in order to run what we call the multiscaling test. The testing procedure is divided into four steps. In the first step, we test if each scaling increment $H(q_i, q_j)$ is statistically significant through a t-test. The second step is devoted to the F-test. In particular, we perform the F-test using the predicted relative moments from both the regression with the full estimated scaling spectrum, and the regression where only the first scaling is different from 0, i.e. $\widehat{\bar{\Xi}}(\tau, q)$ and $\widehat{\bar{\Xi}}(\tau, \bar{q})$, where $\bar{q} = \{q_1, 0, \ldots, 0\}$. If the null hypothesis is rejected, that the full spectrum is necessary to recover all the relative moments.[7] The third step of the procedure consists of a random walk (RW) hypothesis. Assuming the multiscaling parameter $B = 0$, we perform the regression of Equation 9 with only the constant A and test if $\widehat{A} = 0.5$ with a t-test. In fact, for financial returns $H_2 = 0.5$ which implies $H_2 = 0.5 = A + 2B$. Hence, when $B = 0$ for monoscaling time series, we expect $A = 0.5$. In case the null hypothesis is rejected, this means that the RW scaling is incorrect and the use of the *square-root of time rule* severity creates a bias in the risk measures. The last step involves a confirmatory test of the results deriving from the first and second steps of the test procedure. In particular, we perform the full regression of Equation 9 and test for $\widehat{A} = 0.5$ and $\widehat{B} = 0$ using a t-test. If this test gives a conflicting result with respect to the first and second steps, we cannot assert anything on the process with precision and a deeper analysis is required by controlling for different input specifications.

---

[7] For the non-linear regression, in order to use the F-test we have to use $\ln(\widehat{\bar{\Xi}}(\tau, q))$ and $\ln(\widehat{\bar{\Xi}}(\tau, \bar{q}))$.

4.   Simulation experiment

In the simulation experiment, we focus on one of the most used models to generate multifractal time series: the Multifractal Random Walk (MRW) proposed by (Bacry, Delour, and Muzy 2001b,a). This model is capable to generate multifractal time series with a known multiscaling spectrum. In addition, this model is able to generate time series which are consistent with the financial stylized facts. In the discrete version of the MRW, the process $r_\tau(t) = p(t+\tau) - p(t)$ is defined as (Bacry, Delour, and Muzy 2001b):

$$r_\tau(t) = p_{(t+\tau)} - p_t = \sum_{k=\frac{t}{\Delta t}+1}^{\frac{t+\tau}{\Delta t}} \epsilon_{\Delta t}(k) e^{\omega_{\Delta t}(k)}, \tag{14}$$

with

$$\epsilon_{\Delta t} \sim N(0, \sigma^2 \Delta t), \quad \omega \sim N(-\lambda^2 \ln(L/\Delta t), \lambda^2 \ln(L/\Delta t))$$

where λ is called intermittency parameter and determines the strength of the multifractality, L is the autocorrelation length, σ^2 is the variance of the process, and Δt is the discretization step. The distinctive feature of the MRW is that, even if the $\epsilon_{\Delta t}(k)$ are independent, the $\omega_{\Delta t}(k)$ are not, having autocovariance

$$Cov(\omega_{\Delta t}(k_1), \omega_{\Delta t}(k_2)) = \lambda^2 \ln \rho_{\Delta t}(k_1 - k_2),$$

whit

$$\rho_{\Delta t}(k_1 - k_2) = \begin{cases} \dfrac{L}{(|k_1 - k_2| + 1)\Delta t} & |k_1 - k_2| < L/\Delta t, \\ 1 & \text{otherwise.} \end{cases}$$

14

In the continuous limit, the scaling exponents of this model are

$$\zeta(q) = qH(q) = (\lambda^2 + \frac{1}{2})q - \frac{\lambda^2}{2}q^2. \tag{15}$$

The power of this model is that it encompasses all the major stylized facts using only three parameters (λ, L, σ). In fact, this model is able to reproduce fat tails, volatility clustering and multiscaling spectrum. For the purpose of simulation, we generated 100 paths each of dimension $T = 10000$ and we set the model parameters to $L = 250$, $\sigma = 1$ and $\lambda = 0.05, 0.1, 0.3$ in accordance to empirical findings (Bacry, Kozhemyak, and Muzy 2013, 2008; Løvsletten and Rypdal 2012). As explained in previous section, $q_{max} = 1$. In particular, we use $q \in [0.02, 1]$ with steps 0.02 which converts to 50 evaluated moments. To select τ_{max}, we use the τ^* estimated by the ACSR. Table 1 shows results for the different specifications of λ. As shown in the table, the procedure is quite accurate and the 95% confidence intervals (C.I.) always contain the value of $L = 250$ which is the truncation parameter.

	$\bar{\tau}^*$	95% C.I.
$\lambda = 0.05$	297.2	(239.8,354.7)
$\lambda = 0.10$	267.8	(243.2,292.4)
$\lambda = 0.30$	259.9	(231.6,288.3)

Table 1. Results of the ACSR for the estimation of τ_{max}. $\bar{\tau}^*$ is the mean over all the paths and the 95% C.I. are computed over 200000 bootstrapped samples.

Once the parameters are estimated, we compute the multiscaling exponents and evaluate their statistical significance. Since Equation 15 gives the true multiscaling spectrum, we can easily test the performance of the GHE approach and compare it with the new proposed methodology of this paper. We use the normalized and standardized structure function (NSSF) proposed in Equation 8 and the relative normalized and standardized structure function (RNSSF) proposed in Equation 11, which we name Normalized and Standardized Generalized Hurst Exponent (NSGHE) and the Relative Normalized and Standardized Generalized Hurst Exponent (RNSGHE), respectively. The latter methodology will be used to test the multiscaling spectrum. Tables 2 and 3 present the root mean squared errors (RMSE) of the different method-

15

ologies computed over the 100 realizations for both A and B parameters of Equation 4.[8] As it is possible to notice, the RNSGHE generally outperforms with respect to the other specifications. It is important to highlight that by removing the slope ambiguity, results have considerably improved. In fact, the standard GHE approach has the highest RMSE among all the specifications. This result was expected since the new methodology helps to remove uncertainty and thus, to ameliorate the estimation performance.

	GHE	RNSGHE$_L$	NSGHE$_{NL}$	RNSGHE$_{NL}$
$\lambda = 0.05$	0.0338	**0.0148**	0.0163	0.0163
$\lambda = 0.10$	0.0289	**0.0147**	0.0154	0.0153
$\lambda = 0.30$	0.0427	0.0352	0.0343	**0.0339**

Table 2. RMSE for the parameter A for the different methodologies. Subscript L refers to the linear regression while NL to the non-linear regression.

	GHE	RNSGHE$_L$	NSGHE$_{NL}$	RNSGHE$_{NL}$
$\lambda = 0.05$	0.0112	**0.0039**	0.0047	**0.0039**
$\lambda = 0.10$	0.0098	**0.0043**	0.0046	**0.0043**
$\lambda = 0.30$	0.0167	0.0144	**0.0137**	0.0144

Table 3. RMSE for the parameter B for the different methodologies. Subscript L refers to the linear regression while NL to the non-linear regression.

Tables 2 and 3 show a better performance of the RNSGHE$_L$ and RNSGHE$_{NL}$ compared to the other models, in terms of RMSE. For this reason, we use these models in the paper. Now, we show the nature of the process by performing the multiscaling test. Figure 2 shows the p-values of all the 50 coefficients related to the q moments equations for a realization of the MRW model, assuming the same parameters specification as above and $\lambda = 0.05$. What we can observe is that by choosing a confidence level of 5%, the null hypothesis of scaling increments equal to 0 for all the evaluated moments is rejected. However, if we set a more stringent confidence level, for example 1%, the null hypothesis is not rejected for some coefficients, resulting in a uniscaling process.

For $\lambda = 0.1, 0.3$ all the p-values are almost 0. Since we generated processes with a non-negligible amount of multiscaling, this was expected. Once the t-test has been

---

[8]For the linear regression case, the NSGHE and RNSGHE are equivalent models, so we report only the result for the RNSGHE.

16
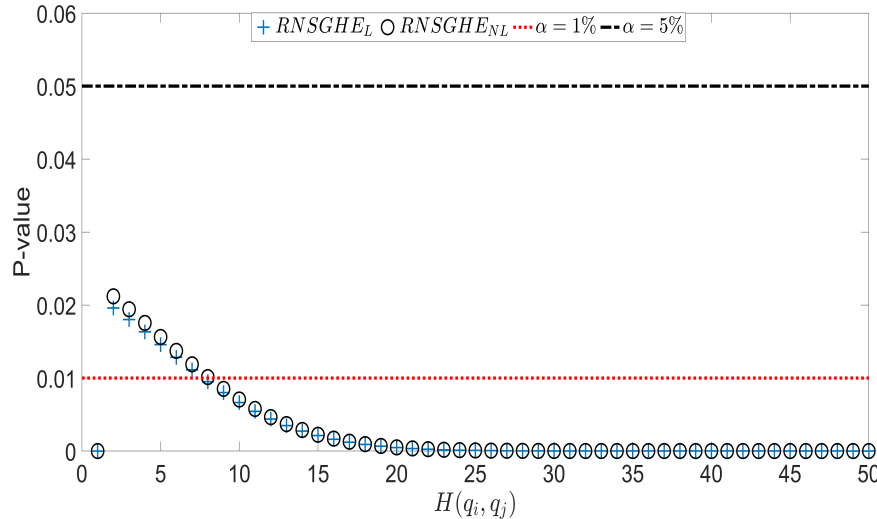
Figure 2. P-values of all the 50 coefficients related to q moments equations for a multifractal random walk with $T = 10000$, $\lambda = 0.05$, $L = 250$ and $\sigma = 1$.

carried out, we perform an F-test on the overall scaling spectrum (Equation 11). Results are reported in Table 4. It is possible to infer that the process generated with $\lambda = 0.05$ does not reject the null for which all the scaling increments are equal to 0 while, for $\lambda = 0.10, 0.30$ the null is rejected and the full scaling spectrum is necessary to reconstruct the relative moments. By combining this result with the outcome of the t-test, we can conclude that the process generated with $\lambda = 0.05$ is weakly multiscaling at 5% confidence level but it is not multiscaling at 1% confidence level. The other two specifications are strongly multiscaling at any reasonable confidence level.

	RNSGHE$_L$	**RNSGHE$_{NL}$**
$\lambda = 0.05$	0.99999	0.99999
$\lambda = 0.10$	0.00003	0.00003
$\lambda = 0.30$	0.00000	0.00000

Table 4. P-values of the F-test for the null that only $H(0, q_1)$ is different from 0. Subscript L refers to linear regression while NL to non-linear regression.

Since these processes have been generated such that they have a specific multiscaling spectrum, the last two tests of the multiscaling test procedure have trivial results.

5.   Empirical application

In this section, we perform the empirical application of the proposed statistical methodology on financial data described in Section 5.1, and produce a statistical analysis of their multiscaling properties in Section 5.2. In Section 5.3, we also study how anomalies impact these estimation.

5.1.   *Data*

The dataset used for the analyses is composed of stocks listed in the *Dow Jones* (DJ). In particular, close prices of stocks are recorded on a daily basis from $03/05/1999$ to $20/11/2019$, i.e. 5363 trading days. We use 27 over the 30 listed stocks since they are the ones for which the entire time series is available. For the purpose of our analysis, we use log-prices and log-returns. Table 5 reports the summary statistics of the data.

As shown in this table, all the empirical stylized facts can be observed. Indeed, log-returns are centered at 0, the Skewness is (in most of the cases) different from 0, while the high value of the Kurtosis clearly depicts fat tails of the log-returns distributions.

5.2.   *Multiscaling test*

In this section, we report results of the multiscaling test.[9] We report the results of all the steps of the testing procedure described at the end of Section 3.1.2. Results are presented using the RNSGHE_L because as explained in the previous section, it has the best performance in the correct estimation of the scaling spectrum.[10] Results are summarized in Table 6. The second column of the table presents the τ^* calculated using the ACSR methodology, which is presented in Section 3.1.1. For its estimation, we fix a maximum value for the choice of τ equal to $\frac{T}{5} = 1072$ in order not to bias the scaling estimation with too few values. We notice that several stocks reach the boundary value, suggesting a very high rate of persistency in the time series. The third column of this table reports the response to the weak multiscaling (weak M-S) process

---

[9]All the tests are performed using a confidence level of 5% unless differently stated.

[10]It is important to highlight that if the other methods proposed in this paper are adopted, results remain qualitatively unchanged.

18

	Summary statistics							
Stock	Mean	Median	Min	Max	Std	Skewness	Kurtosis	Obs
AAPL	0.00	0.00	-0.73	0.13	0.03	-4.08	115.62	5362
AXP	0.00	0.00	-0.19	0.19	0.02	-0.05	13.72	5362
BA	0.00	0.00	-0.19	0.14	0.02	-0.27	8.72	5362
CAT	0.00	0.00	-0.16	0.14	0.02	-0.16	7.59	5362
CSCO	0.00	0.00	-0.17	0.22	0.02	0.14	12.31	5362
CVX	0.00	0.00	-0.13	0.19	0.02	0.07	13.68	5362
DIS	0.00	0.00	-0.20	0.15	0.02	-0.10	14.05	5362
GS	0.00	0.00	-0.21	0.28	0.02	0.61	18.47	5362
HD	0.00	0.00	-0.34	0.13	0.02	-0.98	26.78	5362
IBM	0.00	0.00	-0.21	0.12	0.02	-0.43	15.84	5362
INTC	0.00	0.00	-0.25	0.18	0.02	-0.44	11.89	5362
JN'	0.00	0.00	-0.17	0.12	0.01	-0.61	18.33	5362
JPM	0.00	0.00	-0.23	0.22	0.02	0.25	17.55	5362
KO	0.00	0.00	-0.11	0.13	0.01	0.05	12.27	5362
MCD	0.00	0.00	-0.14	0.09	0.01	-0.09	9.97	5362
MMM	0.00	0.00	-0.14	0.10	0.01	-0.20	9.69	5362
MRK	0.00	0.00	-0.31	0.12	0.02	-1.36	31.86	5362
MSFT	0.00	0.00	-0.17	0.18	0.02	-0.09	12.65	5362
NKE	0.00	0.00	-0.22	0.18	0.02	-0.18	15.88	5362
PFE	0.00	0.00	-0.12	0.10	0.02	-0.27	8.59	5362
PG	0.00	0.00	-0.38	0.10	0.01	-4.21	120.73	5362
TRV	0.00	0.00	-0.20	0.23	0.02	0.31	19.69	5362
UNH	0.00	0.00	-0.22	0.30	0.02	0.05	23.74	5362
UTX	0.00	0.00	-0.33	0.13	0.02	-1.44	36.30	5362
VZ	0.00	0.00	-0.13	0.14	0.01	0.13	9.89	5362
WMT	0.00	0.00	-0.11	0.11	0.01	0.21	10.11	5362
XOM	0.00	0.00	-0.15	0.16	0.01	0.05	13.43	5362

Table 5. Summary statistics of the log-returns of the analyzed stocks.

hypothesis, i.e. hypothesis that a single scaling exponent is enough to approximate the full scaling spectrum but individual scaling increments are statistically significant. As we can observe, none of the analyzed stocks are weakly multiscaling. In fact, as reported in the fourth column, all the stocks pass both the tests and result in strong multiscaling processes. The fifth column of the table reports the result of the RW hypothesis. To perform this analysis we run the regression $H_q = A + \varepsilon$ and test if the estimated A, $\widehat{A}$, is equal to 0.5.[11] We note that only for two stocks the null hypothesis is not rejected, namely Cisco and Pfizer. However, this is a first order approximation of the process and do not check if the process is multiscaling. The sixth column summarizes results

---

[11]This is the standard procedure to estimate the Hurst exponent for uniscaling processes, i.e. $\widehat{A} = \widehat{H}$.

19

of the confirmation test, which is equivalent to test $\widehat{A} = 0.5$ and $\widehat{B} = 0$ in the full regression model of Equation 9. Finally, the last three columns of the table report the estimated Hurst exponent $\widehat{H}$ computed for the RW test, the linear scaling index $\widehat{A}$, and the multiscaling proxy $\widehat{B}$. These results point out that multiscaling is a stylized fact and can be statistically tested by rewriting the structure function in a convenient way.

### 5.3. *Effect of anomalies in the multiscaling estimation*

Multiscaling time series are generated from trading dynamics. One of the fundamental aspects of systems exhibiting multiscaling properties is the strong endogeneity of the sample paths, an aspect which is considered to be originated by financial trading dynamics. For this reason, transient exogenous shocks only distort the analysis and consequently, the estimation procedure. Hence, the statistical procedure used to analyze multiscaling systems are highly sensitive to exogenous shocks. In this context, we refer to an exogenous shock as an unexpected and transient behavior of the stock price, not explainable by the market conditions or by the price path. In addition, anomalies in the time series can occur due to errors or algorithmic trading crashes. Anomalies in financial time series can be grouped in 3 main categories: spikes, jumps and contamination errors. Figure 5.3 shows these possible anomalies. The top left panel is dedicated to the original log-price time series for Verizon. This time series is quite volatile and in fact, the log-returns have a Kurtosis index equal to 9. However, although the distribution of log-returns is fat-tailed, there are not clear anomalies. The top right panel of Figure 5.3 depicts the same log-price time series to which a strong fat-tailed series (Kurtosis larger than 1000) is added. This is the case of contamination error. This is generally due to machine errors in the data transmission process. The bottom left panel reflects the Verizon log-price time series with a random spike added. The spike can arise from multiple sources, among all algorithmic trading errors or contamination errors due to data manipulations. The last panel in the bottom right corner represents the log-price series with an added jump. Jumps per se can arise from endogenous or exogenous

20

Stock	τ^*	Weak M-S	Strong M-S	RW t-test	M-S F-test	$\widehat{H}$	$\widehat{A}$	$\widehat{B}$
AAPL	963	NO	YES	NO	YES	0.5567 (0.5534, 0.5600)	0.5774 (0.5770, 0.5779)	-0.0407 (-0.0414, 0.0399)
AXP	1072	NO	YES	NO	YES	0.5203 (0.5149, 0.5256)	0.5535 (0.5525, 0.5545)	-0.0653 (-0.0670, -0.0636)
BA	1072	NO	YES	NO	YES	0.5391 (0.5353, 0.5429)	0.5630 (0.5623, 0.5637)	-0.0468 (-0.0480, -0.0456)
CAT	527	NO	YES	NO	YES	0.5422 (0.5393, 0.5451)	0.5603 (0.5597, 0.5608)	-0.0355 (-0.0364, -0.0345)
CSCO	1072	NO	YES	YES	YES	0.5011 (0.4966, 0.5056)	0.5292 (0.5280, 0.5304)	-0.0551 (-0.0571, -0.0530)
CVX	313	NO	YES	NO	YES	0.4722 (0.4699, 0.4746)	0.4872 (0.4868, 0.4876)	-0.0293 (-0.0300, -0.0286)
DIS	1072	NO	YES	NO	YES	0.5211 (0.5180, 0.5243)	0.5410 (0.5406, 0.5413)	-0.0389 (-0.0395, -0.0382)
GS	1070	NO	YES	NO	YES	0.4861 (0.4821, 0.4901)	0.5110 (0.5102, 0.5118)	-0.0489 (-0.0503, -0.0475)
HD	1072	NO	YES	NO	YES	0.5370 (0.5332, 0.5409)	0.5614 (0.5608, 0.5619)	-0.0477 (-0.0486, -0.0467)
IBM	1072	NO	YES	NO	YES	0.4842 (0.4808, 0.4876)	0.5053 (0.5049, 0.5057)	-0.0413 (-0.0420, -0.0407)
INTC	1072	NO	YES	NO	YES	0.4568 (0.4531, 0.4605)	0.4797 (0.4786, 0.4809)	-0.0450 (-0.0470, -0.0430)
JN	488	NO	YES	NO	YES	0.4664 (0.4629, 0.4698)	0.4880 (0.4874, 0.4885)	-0.0423 (-0.0432, -0.0415)
JPM	1072	NO	YES	NO	YES	0.4851 (0.4808, 0.4895)	0.5124 (0.5118, 0.5131)	-0.0535 (-0.0546, -0.0523)
KO	1072	NO	YES	NO	YES	0.4937 (0.4905, 0.4968)	0.5134 (0.5129, 0.5139)	-0.0387 (-0.0395, -0.0379)
MCD	480	NO	YES	NO	YES	0.5116 0.5092, 0.5140)	0.5266 (0.5261, 0.5270)	-0.0293 (-0.0301, -0.0286)
MMM	1032	NO	YES	NO	YES	0.4564 (0.4534, 0.4593)	0.4747 (0.4743, 0.4751)	-0.0360 (-0.0366, -0.0353)
MRK	867	NO	YES	NO	YES	0.5231 (0.5189, 0.5274)	0.5498 (0.5484, 0.5512)	-0.0522 (-0.0546, -0.0499)
MSFT	1072	NO	YES	NO	YES	0.4737 (0.4698, 0.4776)	0.4981 (0.4975, 0.4988)	-0.0479 (-0.0490, -0.0468)
NKE	1072	NO	YES	NO	YES	0.4763 (0.4739, 0.4787)	0.4913 (0.4912, 0.4915)	-0.0295 (-0.0298, -0.0292)
PFE	1072	NO	YES	YES	YES	0.4999 (0.4952, 0.5046)	0.5285 (0.5263, 0.5308)	-0.0562 (-0.0600, -0.0523)
PG	969	NO	YES	NO	YES	0.4617 (0.4582, 0.4652)	0.4834 (0.4829, 0.4839)	-0.0426 (-0.0435, -0.0417)
TRV	604	NO	YES	NO	YES	0.4673 (0.4643, 0.4703)	0.4861 (0.4857, 0.4865)	-0.0368 (-0.0375, -0.0362)
UNH	580	NO	YES	NO	YES	0.5163 (0.5134, 0.5192)	0.5346 (0.5342, 0.5351)	-0.0360 (-0.0367, -0.0352)
UTX	1072	NO	YES	NO	YES	0.4799 (0.4759, 0.4840)	0.5053 (0.5048, 0.5059)	-0.0497 (-0.0507, -0.0488)
VZ	1072	NO	YES	NO	YES	0.4712 (0.4683, 0.4742)	0.4897 (0.4889, 0.4905)	-0.0362 (-0.0376, -0.0348)
WMT	1072	NO	YES	NO	YES	0.4314 (0.4277, 0.4350)	0.4540 (0.4534, 0.4546)	-0.0444 (-0.0454, -0.0434)
XOM	344	NO	YES	NO	YES	0.4642 (0.4608, 0.4675)	0.4852 (0.4846, 0.4858)	-0.0413 (-0.0422, -0.0403)

Table 6. Results of the multiscaling estimation and testing procedure.

shocks. However, if they derive from an endogenous driving force, they persist in the jump direction. Conversely, if they come from an exogenous source instead, they tend

21

to be transient. In a relatively recent paper (Sornette, Malevergne, and Muzy. 2003), the authors explain that huge financial crashes can be originated from endogenous shocks which have a huge persistence behavior. These kinds of shocks are inherent in the price process so they are not transient anomalies.
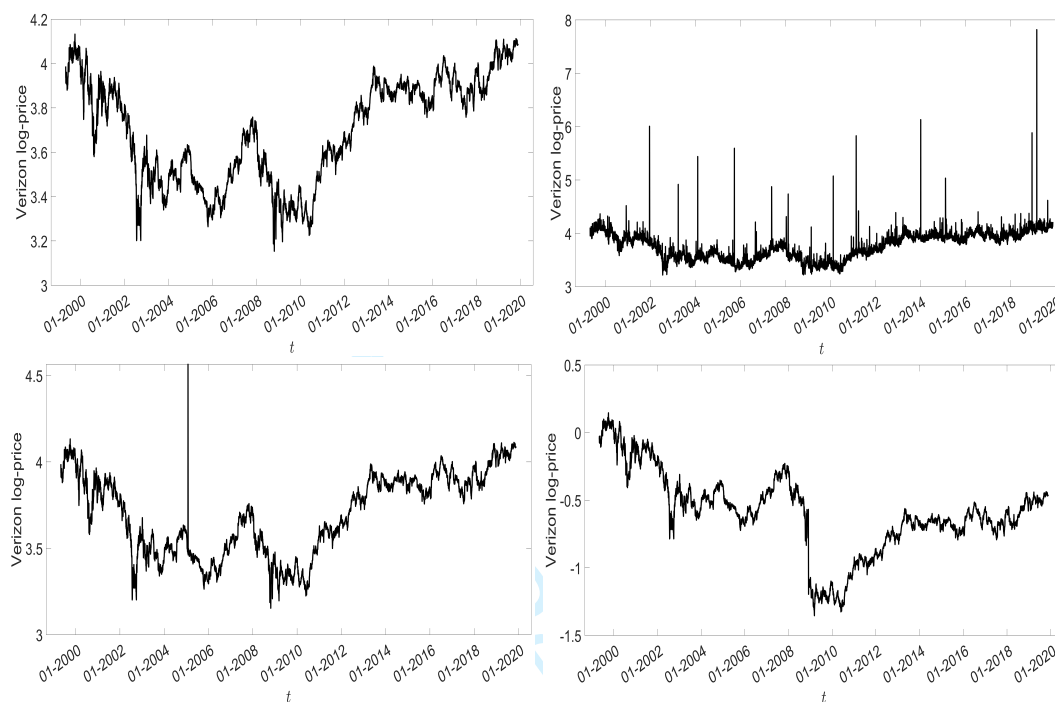
Figure 3. Verizon log-prices time series and some anomalies of financial time series.

In a mostly technical paper, (Katsev and L'Heureux 2003) show both theoretically and experimentally that such data anomalies can strongly bias results, especially for short datasets. In particular, the paper shows that under certain circumstances, these irregularities can generate spurious scaling. For these reasons, it is suggested to analyze the time series and eliminate such anomalies before proceeding with the scaling estimation. In order to do so, we propose a methodology based on financial stylized facts. More precisely, we use volatility clustering and long-range dependence of asset returns (Cont 2001; Chakraborti et al. 2011). In this empirical context, the quantities

22

that we name Cumulative Variance (CV) and Cumulative Auto-Covariance (CAV):[12]

$$CV(t) = \sum_{i=1}^{t} r_i^2 \tag{16}$$

and

$$CAV(t) = \sum_{i=1}^{t} |r_{i+1}||r_i| \tag{17}$$

should be very similar, except when an exogenous (unexpected) anomaly exists. The volatility clustering drives the similarity in the short period since $|r_{i+1}|$ and $|r_i|$ are expected to be very similar (same cluster), while the long-range dependence drives the similarity of the two measures over the long-run. Figure 4 represents the two quantities for the Verizon stock. These two quantities are approximately equal, confirming that even with high volatility and many tail events, the time series does not contain exogenous shocks.
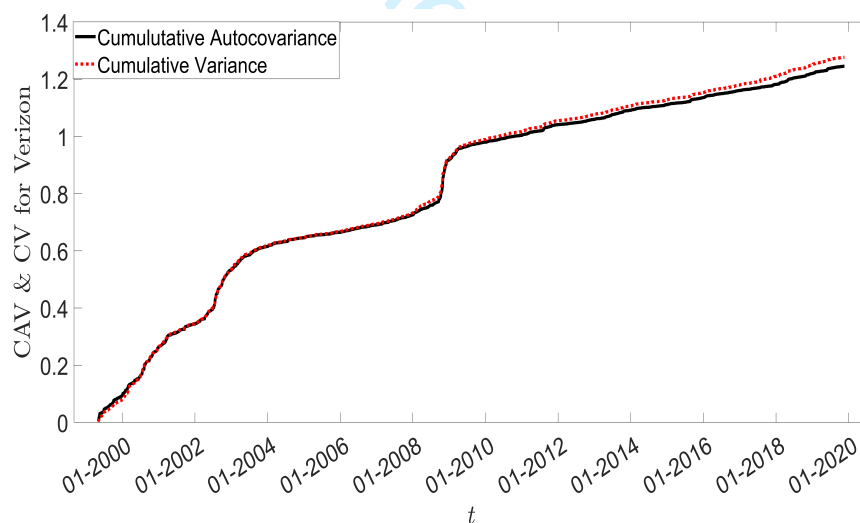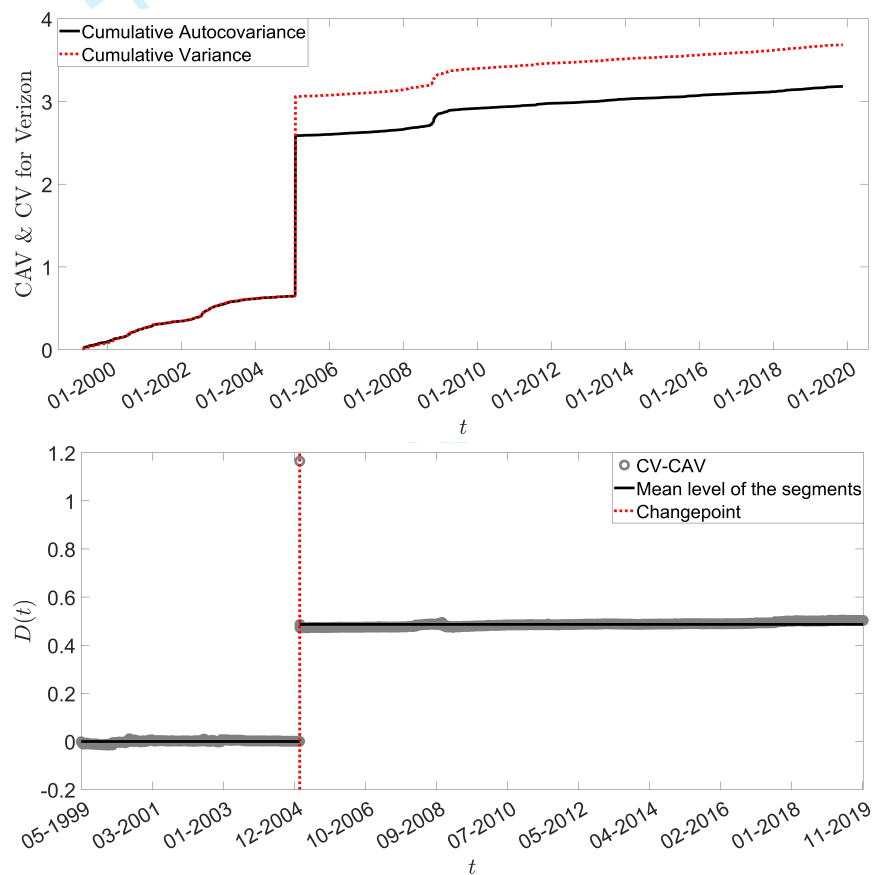

Figure 4. Cumulative Variance (CV) and Cumulative Auto-Covariance (CAV) for the Verizon time series.

The difference between the two cumulative series is given by $D(t) = CV(t) - CAV(t)$. By running a change-point detection in the intercept and slope of $D(t)$, it is possible

---

[12]In the financial High-frequency literature, these quantities are strongly related to the Realized Variance and Bipower Variation.

23

to detect the anomalies in the price series and replace the corresponding values on the original series according to a specific rule, e.g. the mean of previous and subsequent data points. Top panels of Figures 5 and 6 reflect the case in which a spike and a jump have been added to the log-price time series of Verizon, respectively. As shown in the figures, the two measures start to diverge significantly exactly in correspondence of the anomaly point. The bottom panels instead, represent the change point detection performed on the quantity $D(t)$. The panel shows that the procedure correctly identifies the position of the anomaly.

Figure 5. Cumulative Variance ($CV(t)$) and Cumulative Auto-Covariance ($CAV(t)$) for the Verizon time series (top panel) and the measure $D(t)$ (bottom panel) in case of the added spike.

Given the fact that such events (spikes or jumps) are rare and have unconventional magnitude, their removal can only benefit the analysis. Let us estimate the multiscaling exponent for the Verizon stock when the anomalies reported in Figure 5.3 (bottom panels) are not removed by the time series. Table 7 reports the results. The estimated

24
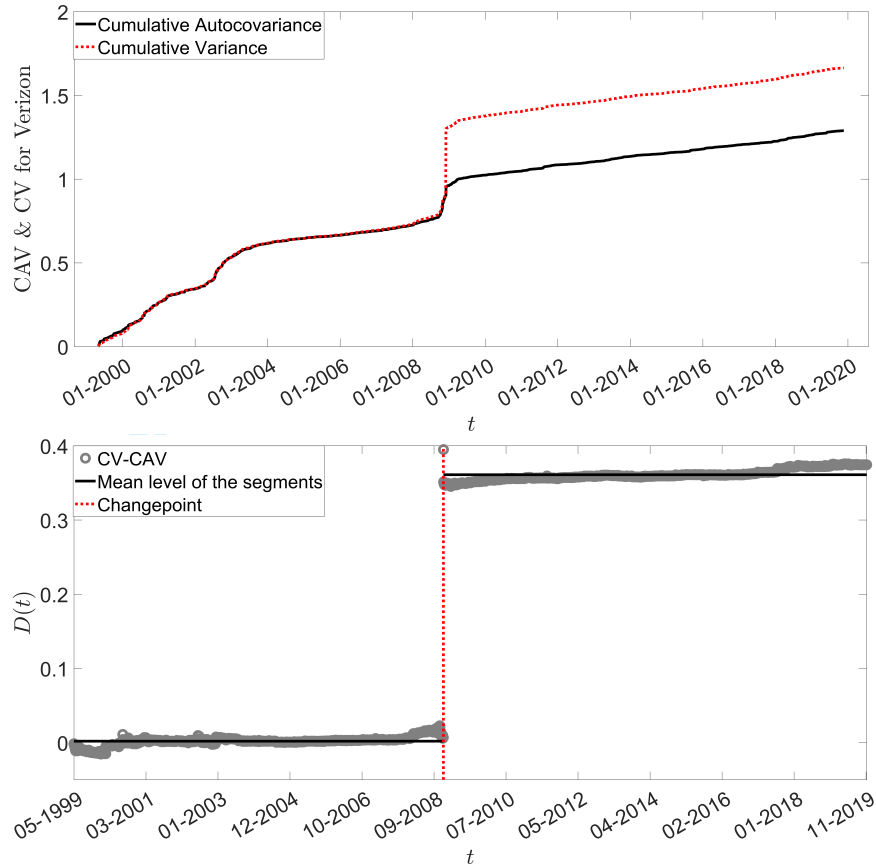
Figure 6. Cumulative Variance ($CV(t)$) and Cumulative Auto-Covariance ($CAV(t)$) for the Verizon time series (top panel) and the measure $D(t)$ (bottom panel) in case of the added jump.

values change considerably, especially in the scenario where a spike is added.

	$\widehat{H}$	$\widehat{A}$	$\widehat{B}$
Spike	0.5290	0.6013	-0.1417
Jump	0.4942	0.5220	-0.0543

Table 7. Results of the multiscaling estimation on the times series reported in Figure 5.3 (bottom panels) with anomalies not removed.

For completeness, we also perform a t-test with the null hypothesis of no difference between the estimates with anomalies and the estimates reported in Table 6. The null hypothesis for all the coefficients is strongly rejected at any confidence level.

To show how these anomalies can generate spurious multiscaling, we generate 100 fractional Brownian motions (uniscaling process) of length 1000 with Hurst exponent $H = 0.47$ (the one estimated for Verizon). To these simulated time series we add a

spike and a jump and estimate $\widehat{H}$, $\widehat{A}$ and $\widehat{B}$ for both the series with and without the anomalies. Results are reported in Table 8. As we can see, when the anomalies are not present in the time series, the average values for $\widehat{H}$, $\widehat{A}$ and $\widehat{B}$ are in line with the true values and not statistically different from them. In the scenario with the added spike, we can see that $\widehat{H}$, $\widehat{A}$ and $\widehat{B}$ are severely biased. In particular, due to the spike, the times series look multiscaling, while it is not. Finally, for the case of the added exogenous jump, we have that the scaling exponent curves and the parameter B is not equal to 0. Also the other two estimated coefficients are upward biased and statistically different from the theoretical ones. These results clearly show that the

	$\widehat{H}$	$\widehat{A}$	$\widehat{B}$
No anomalies	0.4663	0.4654	-0.0018
	(0.4597, 0.4710)	(0.4596, 0.4728)	(-0.0048, 0.0012)
Spike	0.4403	0.4893	-0.0961
	(0.4330, 0.4481)	(0.4799, 0.4998)	(-0.1027, -0.0899)
Jump	0.4845	0.4806	0.0076
	(0.4785, 0.4905)	(0.4737, 0.4874)	(0.0045, 0.011)

Table 8. Average of the 100 estimated $\widehat{H}$, $\widehat{A}$ and $\widehat{B}$ for a fractional Brownian motion with $H = 0.47$. 95% C.I. computed over 200000 bootstrapped samples are reported in parenthesis.

scaling exponents are sensitive to such anomalies and estimates can be biased if such anomalies are not carefully analyzed.

5.4. *Jump test on the empirical data*

We performed the jump detection analysis on the 27 stocks used for the analysis and found that 6 of them exhibited jumps. Table 9 summarizes the results. We report the date(s) in which the jump is identified by the change point detection algorithm,[13] as well as the estimated $\hat{H}$, $\hat{A}$ and $\hat{B}$ when the jump(s) have been removed.[14]

From Table 9, we can notice the following: when removal of the jump increases the estimate of $\widehat{H}$ and $\widehat{A}$, multiscaling is reduced, while the opposite is true apart for UTX stock. Furthermore, we can see that, apart for $\widehat{H}$ for the MRK stock, all the other estimates fall outside of the confidence intervals of the same parameters' estimates

---

[13] For the change point detection we used $\frac{1}{2} \parallel D(t) - \mathbb{E}[D(t)] \parallel_2$ as penalty.

[14] We imputed the value corresponding to the jump with the average between previous and subsequent data points.

	Jump date	$\widehat{H}$	$\widehat{A}$	$\widehat{B}$
AAPL	29-Sept-2000	0.5315	0.5527	-0.0416
HD	12-Oct-2000	0.5203	0.5535	-0.0653
MRK	30-Sept-2004	0.5391	0.5630	-0.0468
PG	07-Mar-2000	0.5422	0.5603	-0.0355
UNH	13-Oct-2008	0.5011	0.5292	-0.0551
UTX	17-Sept-2001	0.4722	0.4872	-0.0293

Table 9. Results of the jump detection analysis and estimated exponent parameters

reported in Table 6. This means that the estimated coefficients reported in Table 6 are statistically different (at 5% significance level) from the ones estimated without the jumps. Results make clear that in case of abrupt jumps or spikes, the estimates of the scaling exponent parameters can be severely impacted.

6.   Practical application of scaling and multiscaling to VaR

In this section, we show that by using a simple VaR configuration without a scaling or multiscaling consideration might bias the VaR estimation at higher aggregation scales. We use daily stocks data from the Dow Jones index to estimate the multiscaling spectrum and to carry out the multiscaling test described in Section 5. After the procedure is concluded, we estimate the two most common VaR models, i.e. the Historical and Gaussian VaR at 1 day. Successively, we use these estimates to compute the yearly VaR using the *square root of time rule*.[15] We then compare them with the fractional VaR with proper scaling and highlight eventual biases. To conclude, we propose a methodology to compute a multiscaling consistent VaR.

6.1.   *Value at Risk*

VaR is an easy and intuitive way to quantify risk for assets and portfolios. Let $VaR(\tau, 1 - \alpha)$ be the Value at Risk at frequency τ for a confidence level equal to

---

[15]Theoretically, this is only valid with the Gaussian formulation. Nevertheless, it is commonly used also for Historical VaR time aggregation.

27

$1 - \alpha$ which satisfies

$$P(r_\tau(t) < VaR(\tau, 1 - \alpha)) = \alpha \qquad (18)$$

where $r_\tau(t)$ are the log-returns at frequency τ. Several methodologies are used to compute the VaR. Among all, we recall the Historical VaR (HVaR) and the Gaussian VaR(GVaR).[16] The former is a non-parametric approach that uses historical data to compute the VaR, while the latter assumes a Gaussian distribution of stock returns and applies the Gaussian formula for the percentiles computation to extract the VaR at a given confidence level. The issue faced in applied Finance is that the *square root of time rule* works only under the assumption of *iid* Gaussian returns. However, this technique is widely adopted regardless of its assumptions.

In our analysis, VaR is computed using the two aforementioned approaches at $\tau = 1$ day, and 95% confidence level ($Var(1, 95\%)$). Annual VaR ($Var(250, 95\%)$) is calculated with the scaling exponent equal to 0.5, i.e. $\widehat{Var}(250, 95\%) = Var(1, 95\%) \times 250^{0.5}$ and to the estimated H, i.e. $\widehat{Var}_H(250, 95\%) = Var(1, 95\%) \times 250^{\widehat{H}}$.[17] We further estimate the true $Var(250, 95\%)$ using annual returns ($\tau = 250$ days) and compare them. Results are shown in Figures 7 and 8. These figures show that when we compare the VaR calculated using the H scaling time rule and the VaR with *the square root of time rule*, the deviation from the true VaR is lower when the former approach is used. In fact, for the VaR with H scaling time, the bias with respect to the true VaR is considerably lower for most of the stocks in both the HVaR and GVaR settings. This is due to the fact that over the long-run, even a small divergence from the assumption of scaling exponent equal to 0.5 can have a substantial impact.

To conclude the analysis with a quantitative assessment of the performance of the different methodologies, we also report the relative error (RE), i.e.

$$RE = \frac{|Var(250, 95\%) - Var(1, 95\%) \times 250^K|}{|Var(250, 95\%)|},$$

---

[16] Besides the common knowledge of non-Gaussianity of stocks' log-returns, the Gaussian VaR is widely adopted both in academia and industry.

[17] We will discuss the multiscaling case later in the paper.
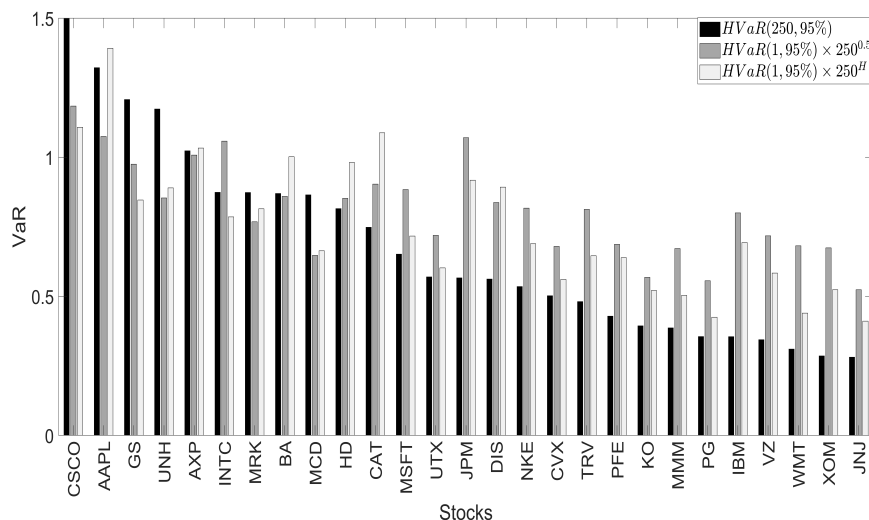
Figure 7. HVaR using annual data and using daily data rescaled by the factor $250^{0.5}$ and by $250^{\widehat{H}}$. Stocks are sorted by the magnitude of the annual Historical VaR.
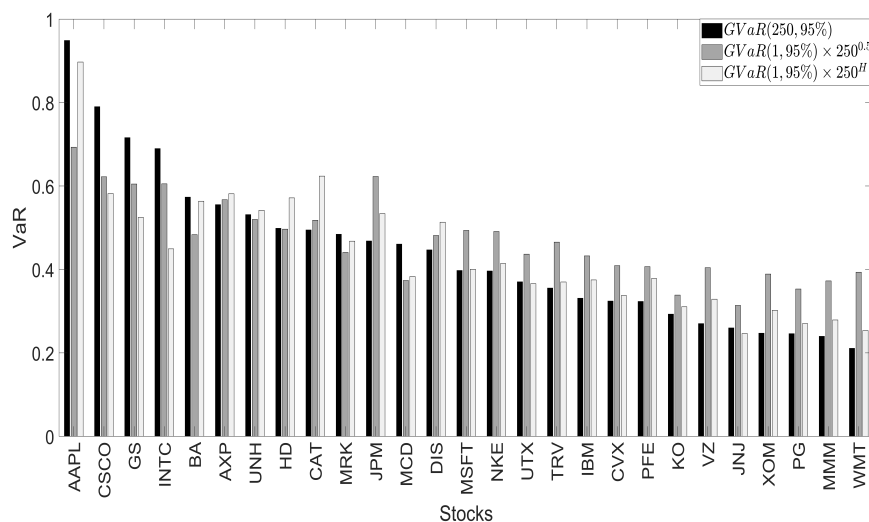
Figure 8. GVaR using annual data and using daily data rescaled by the factor $250^{0.5}$ and by $250^{\widehat{H}}$. Stocks are sorted by the magnitude of the annual Gaussian VaR.

where K is equal to 0.5 for the VaR computed with the square root of time rule or equal to the estimated H, $\widehat{H}$, as reported in Table 6. This helps to identify the magnitude of the deviation from the true VaR and to compare the two scaling approaches. Figure 9 shows the results. As the figure shows, using the correct scaling results in a smaller relative error. This confirms that the choice of a proper scaling exponent should not be neglected by the financial community, considering that its estimation and testing are relatively simple.

29
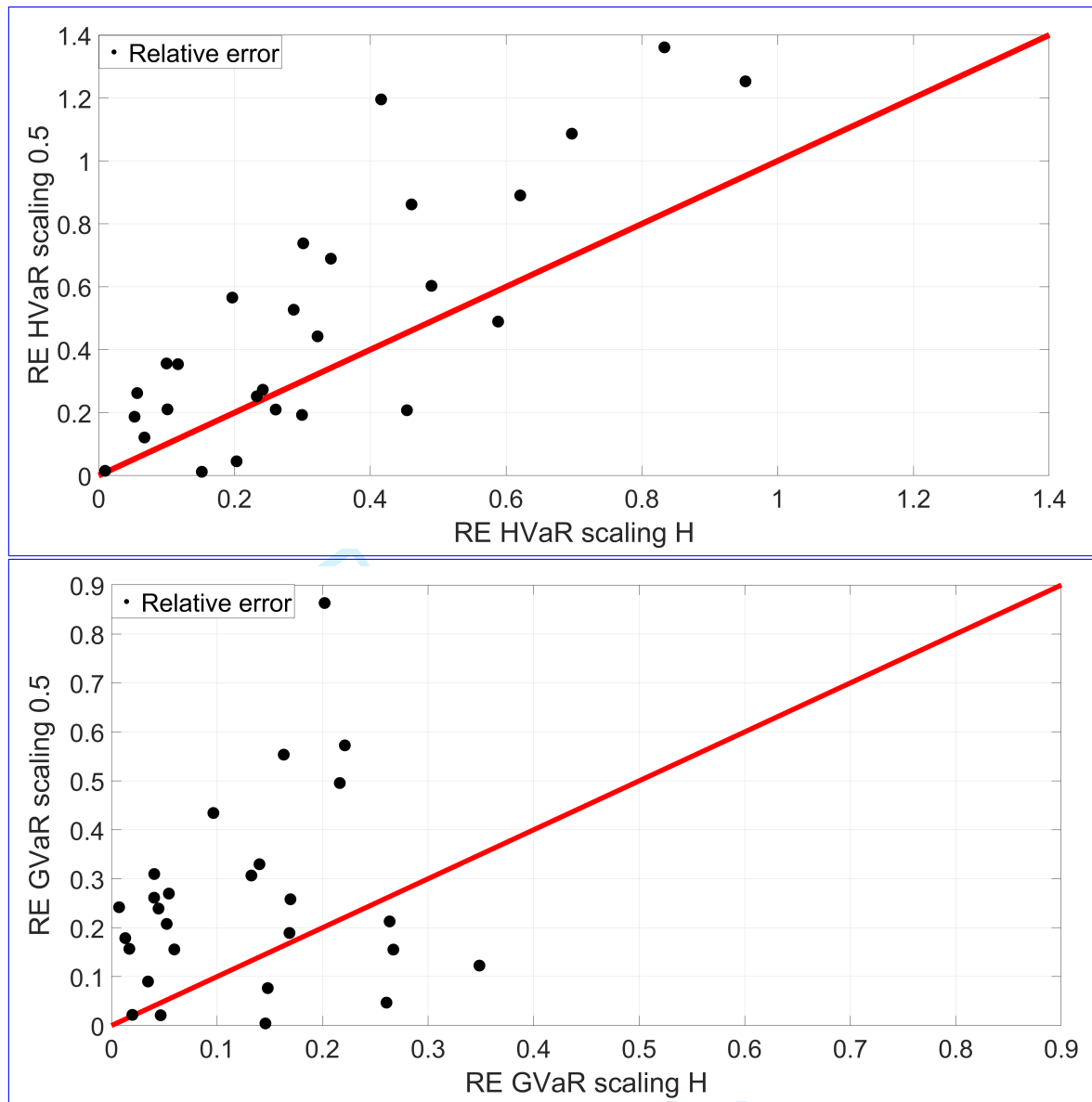
Figure 9. Relative error between the true VaR calculated using annual data, HVaR and GVaR computed using daily data scaled by the factor 0.5 and by the estimated factor H. The red line is the 45 degrees reference line.

6.1.1. Multiscaling consistent VaR

In the previous subsection, we showed that using the correct scaling contributes to reduce the computation error for VaR at smaller frequencies. However, as explained in Section 5.2, all time series analyzed are strongly multiscaling. To deal with such situations, we discuss a possible solution. While VaR is related to the log-returns, multiscaling is a property of the moments of the log-returns. For this reason, there is

30

not a straightforward formula to compute VaR which takes into account multiscaling. An exception to this is the Multifractal VaR proposed in (Lee, Song, and Chang 2016), where the author introduces a VaR consistent with the multifractality of financial time series using the Multifractal Model of Asset Returns (MMAR). In a previous paper (Batten, Kinateder, and Wagner 2014), a similar analysis is performed but it relies on the MMAR Monte Carlo simulations and it computes the VaR on the simulated time series. The Monte Carlo approach has the advantage of letting the researcher use the model that best depicts the data. In fact, one can calibrate the MRW or the MMAR and generate a large number of sample paths which can be used to compute VaR. In the case of moderate multiscaling, the difference can be low but for multiscaling processes with a $|B| > 0.05$, neglecting such a feature can strongly distort the VaR. In this work, we use the MRW to simulate 250 trading days (i.e. 1 year) of log-returns and compute the VaR of the simulated paths. For this purpose, three parameters need to be estimated: the variance σ^2, the autocorrelation scale parameter L and the intermittency parameter λ. The variance can be estimated from the log-returns time series as $\sigma^2 = Var(r_\tau(t))$ with $\tau = 1$, the parameter L is set to be equal to τ^*, while the intermittency parameter λ can be extracted by equating the estimated coefficients of Equation 4 to the parameters in Equation 15 and getting two (possibly different) estimates of λ, i.e. λ_A and λ_B.[18] For each stock, we estimate the three parameters, σ^2, L and λ. Hence, we generate 100000 independent paths of daily returns for a year (i.e. 250 days). Finally, we quantify the VaR, which we name Multiscaling VaR (MSVaR), by computing the 95% percentile on these simulations. Results are depicted in Figure 10.

It is possible to appreciate that the MSVaR computed on the simulated paths has a comparable size to the Historical VaR computed on annual data. It is also important to note that the values predicted using λ_A and λ_B are very similar, suggesting that the stocks log-returns can be adequately approximated by the MRW model. Nevertheless, we remark on the importance of the full multiscaling estimation and testing procedure

---

[18]If the data generating process is not a fully MRW, the estimation of λ by using A or B can differ substantially. In our case, for most of the stocks analyzed the intermittency parameter computed with the two estimated coefficients, i.e. λ_A and λ_B, are very similar.

31
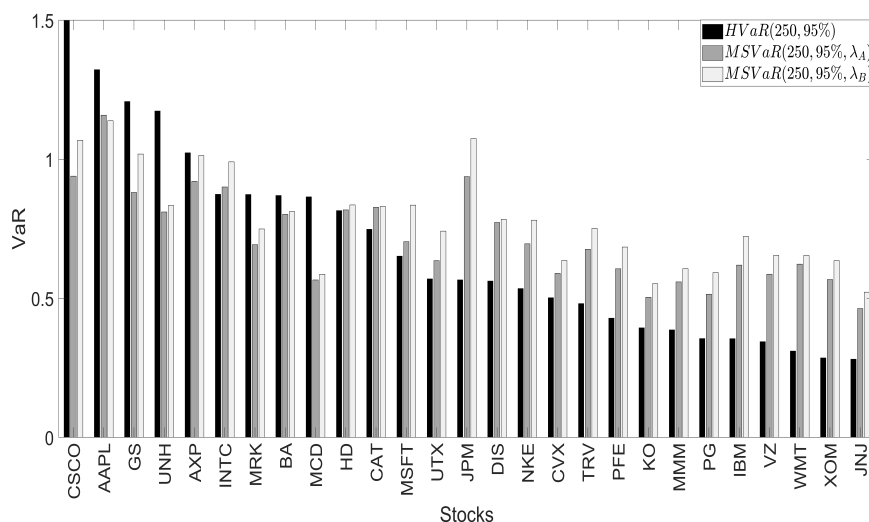
Figure 10. Annual Historical VaR (HVaR) and Multiscaling VaR (MSVaR). Confidence level 95%. Stocks are sorted according to the magnitude of the Historical VaR.

which lead to the MSVaR. In fact, if the previous analysis is bypassed the estimated risk metrics can be severely biased and inconsistent.

7. Conclusions

In this paper, we propose a step-by-step procedure to robustly estimate and test multiscaling in financial time series. By rewriting the structure function in a convenient way we perform multiple tests on the scaling spectrum and assess the statistical significance of multiscaling, discriminating between weak and strong multiscaling. We have shown the effect of anomalies in financial time series and studied the impact on the estimated scaling exponents. Moreover, we have shown how the use of proper scaling can help to reduce the error in the VaR forecasting at a smaller frequency with respect to the commonly used *square root of time rule*. Finally, we have proposed a Multiscaling consistent VaR using a Monte Carlo MRW simulation calibrated to the data and on which the VaR is then computed. Results are encouraging and confirm the goodness of the proposed methodological approach. Multiscaling is a stylized fact which can make a difference in the assessment of risk measures and in building quantitative models. It can be easily extrapolated from data and should not be overlooked by risk managers and authorities.

32

Acknowledgement

References

Antoniades, Ioannis P, Giuseppe Brandi, LG Magafas, and T Di Matteo. 2020. "The use of scaling properties to detect relevant changes in financial time series: a new visual warning tool." *Physica A: Statistical Mechanics and its Applications* 125561.

Bacry, Emmanuel, Jean Delour, and Jean-François Muzy. 2001a. "Modelling financial time series using multifractal random walks." *Physica A: Statistical Mechanics and its Applications* 299 (1): 84–92.

Bacry, Emmanuel, Jean Delour, and Jean-François Muzy. 2001b. "Multifractal random walk." *Physical Review E* 64 (2): 026103.

Bacry, Emmanuel, Alexey Kozhemyak, and Jean-François Muzy. 2008. "Continuous cascade models for asset returns." *Journal of Economic Dynamics and Control* 32 (1): 156–199.

Bacry, Emmanuel, Alexey Kozhemyak, and Jean-François Muzy. 2013. "Log-normal continuous cascade model of asset returns: aggregation properties and estimation." *Quantitative Finance* 13 (5): 795–818.

Barunik, Jozef, and Ladislav Kristoufek. 2010. "On Hurst exponent estimation under heavy-tailed distributions." *Physica A: Statistical Mechanics and its Applications* 389 (18): 3844–3855.

Batten, Jonathan A, Harald Kinateder, and Niklas Wagner. 2014. "Multifractality and value-at-risk forecasting of exchange rates." *Physica A: Statistical Mechanics and its Applications* 401: 71–81.

Buonocore, Riccardo J., Tomaso Aste, and Tiziana Di Matteo. 2017. "Asymptotic scaling properties and estimation of the generalized Hurst exponents in financial data." *Phys. Rev. E* 95: 042311.

Buonocore, Riccardo Junior, Tomaso Aste, and Tiziana Di Matteo. 2016. "Measuring multi-

33

scaling in financial time-series." *Chaos, Solitons & Fractals* 88: 38–47.

Buonocore, RJ, G Brandi, RN Mantegna, and T Di Matteo. 2019. "On the interplay between multiscaling and stock dependence." *Quantitative Finance* 20 (1): 133–145.

Calvet, Laurent E., and Adlai J. Fisher. 2002. "Multifractality in asset returns: Theory and evidence." *Review of Economics and Statistics* 84 (3): 381–406.

Calvet, Laurent E., and Adlai J. Fisher. 2004. "How to forecast long-run volatility: Regime switching and the estimation of multifractal processes." *Journal of Financial Econometrics* 2 (1): 49–83.

Calvet, Laurent E., Adlai J. Fisher, and Benoit B. Mandelbrot. 1997. "Large deviations and the distribution of price changes." *Cowles Foundation Discussion Paper No. 1165* .

Chakraborti, Anirban, Ioane Muni Toke, Marco Patriarca, and Frédéric Abergel. 2011. "Econophysics review: I. Empirical facts." *Quantitative Finance* 11 (7): 991–1012.

Clauset, Aaron, Cosma Rohilla Shalizi, and Mark EJ Newman. 2009. "Power-law distributions in empirical data." *SIAM Review* 51 (4): 661–703.

Cont, Rama. 2001. "Empirical properties of asset returns: stylized facts and statistical issues." *Quantitative Finance* 1 (2): 223–236.

Danielsson, Jon, and Jean-Pierre Zigrand. 2006. "On time-scaling of risk and the square-root-of-time rule." *Journal of Banking & Finance* 30 (10): 2701–2713.

Di Matteo, Tiziana. 2007. "Multi-scaling in finance." *Quantitative finance* 7 (1): 21–36.

Di Matteo, Tiziana, Tomaso Aste, and Michel M. Dacorogna. 2003. "Scaling behaviors in differently developed markets." *Physica A: Statistical Mechanics and its Applications* 324 (1): 183–188.

Di Matteo, Tiziana, Tomaso Aste, and Michel M. Dacorogna. 2005. "Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development." *Journal of Banking & Finance* 29 (4): 827–851.

Eom, Cheoljun, Taisei Kaizoji, and Enrico Scalas. 2019. "Fat tails in financial return distributions revisited: Evidence from the Korean stock market." *Physica A: Statistical Mechanics and its Applications* 526: 121055.

Fukasawa, Masaaki, Tetsuya Takabatake, and Rebecca Westphal. 2019. "Is Volatility Rough?" *arXiv preprint arXiv:1905.04852* .

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum. 2018. "Volatility is rough." *Quantitative Finance* 18 (6): 933–949.

Gençay, Ramazan, Michel Dacorogna, Ulrich A Muller, Olivier Pictet, and Richard Olsen.

34

2001. *An introduction to high-frequency finance*. Elsevier.

Hurst, Harold Edwin. 1956. "Methods of using long-term storage in reservoirs." *Proceedings of the Institution of Civil Engineers* 5 (5): 519–543.

Jiang, Zhi-Qiang, Wen-Jie Xie, Wei-Xing Zhou, and Didier Sornette. 2019. "Multifractal analysis of financial markets: a review." *Reports on Progress in Physics* 82 (12): 125901.

Kantelhardt, Jan W., Stephan A. Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, and H. Eugene Stanley. 2002. "Multifractal detrended fluctuation analysis of nonstationary time series." *Physica A: Statistical Mechanics and its Applications* 316 (1): 87–114.

Katsev, Sergei, and Ivan L'Heureux. 2003. "Are Hurst exponents estimated from short or irregular time series meaningful?" *Computers & Geosciences* 29 (9): 1085–1089.

Kolmogorov, Andrey Nikolaevich. 1962. "A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number." *Journal of Fluid Mechanics* 13 (1): 82–85.

Kristoufek, Ladislav. 2010. *Long-range dependence in returns and volatility of Central European Stock Indices*. Technical Report. IES Working Paper.

Lee, Hojin, Jae Wook Song, and Woojin Chang. 2016. "Multifractal value at risk model." *Physica A: Statistical Mechanics and its Applications* 451: 113–122.

Lillo, Fabrizio, and J Doyne Farmer. 2004. "The long memory of the efficient market." *Studies in nonlinear dynamics & econometrics* 8 (3).

Livieri, Giulia, Saad Mouti, Andrea Pallavicini, and Mathieu Rosenbaum. 2018. "Rough volatility: evidence from option prices." *IISE transactions* 50 (9): 767–776.

Løvsletten, Ola, and Martin Rypdal. 2012. "Approximated maximum likelihood estimation in multifractal random walks." *Physical Review E* 85 (4): 046705.

Lux, Thomas. 2004. "Detecting multi-fractal properties in asset returns: The failure of the scaling estimator." *International Journal of Modern Physics C* 15 (04): 481–491.

Lux, Thomas, and Michele Marchesi. 1999. "Scaling and criticality in a stochastic multi-agent model of a financial market." *Nature* 397 (6719): 498.

Mandelbrot, Benoit. 1967. "The variation of some other speculative prices." *The Journal of Business* 40 (4): 393–413.

Mandelbrot, Benoit B. 1963. "The variation of certain speculative prices." *The Journal of Business* 36 (4): 394–419.

Mandelbrot, Benoit B. 2013. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Selecta Volume E*. Springer Science & Business Media.

Mandelbrot, Benoit B., Adlai Fisher, and Laurent E. Calvet. 1997. *A Multifractal Model of Asset Returns.* Cowles Foundation Discussion Papers 1164. Cowles Foundation for Research in Economics, Yale University.

Mantegna, Rosario N., and H. Eugene Stanley. 1995. "Scaling behaviour in the dynamics of an economic index." *Nature* 376 (6535): 46–49.

Mantegna, Rosario N., and H. Eugene Stanley. 1999. *Introduction to Econophysics: Correlations and Complexity in Finance.* Cambridge University Press.

Muzy, Jean-François, Emmanuel Bacry, and Alain Arneodo. 1991. "Wavelets and multifractal formalism for singular signals: Application to turbulence data." *Physical Review Letters* 67 (25): 3515–3518.

Muzy, Jean-François, Emmanuel Bacry, and Alain Arneodo. 1993. "Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method." *Physical Review E* 47 (2): 875–884.

Scalas, Enrico, and Kyungsik Kim. 2006. "The art of fitting financial time series with Levy stable distributions." *arXiv preprint physics/0608224* .

Sornette, D., Y. Malevergne, and J.-F. Muzy. 2003. "What causes crashes?" *Risk, 16:67–71* .

Takaishi, Tetsuya. 2020. "Rough volatility of Bitcoin." *Finance Research Letters* 32: 101379.

Valsamis, Epaminondas Markos, Henry Husband, and Gareth Ka-Wai Chan. 2019. "Segmented Linear Regression Modelling of Time-Series of Binary Variables in Healthcare." *Computational and Mathematical Methods in Medicine* 2019 (3478598): 1–7.

Van Atta, CW, and WY Chen. 1970. "Structure functions of turbulence in the atmospheric boundary layer over the ocean." *Journal of Fluid Mechanics* 44 (1): 145–159.

Virkar, Yogesh, Aaron Clauset, et al. 2014. "Power-law distributions in binned empirical data." *The Annals of Applied Statistics* 8 (1): 89–119.

Wang, Jying-Nan, Jin-Huei Yeh, and Nick Ying-Pin Cheng. 2011. "How accurate is the square-root-of-time rule in scaling tail risk: A global study." *Journal of Banking & Finance* 35 (5): 1158–1169.

Weron, Rafał. 2001. "Levy-stable distributions revisited: tail index> 2 does not exclude the Levy-stable regime." *International Journal of Modern Physics C* 12 (02): 209–223.

Yue, Peng, Hai-Chuan Xu, Wei Chen, Xiong Xiong, and Wei-Xing Zhou. 2017. "Linear and non-linear correlations in the order aggressiveness of Chinese stocks." *Fractals* 25 (5): 1750041.

36