A multi-scale symmetry analysis of uninterrupted trends returns of daily financial indices

C.M. Rodríguez-Martínez[a], H.F. Coronel-Brizio[a,b], A.R. Hernández-Montoya[a,b,*]

[a]*Centro de Investigación en Inteligencia Artificial. Universidad Veracruzana. Sebastián Camacho 5, Xalapa Veracruz 91000, México. Tel/Fax: 52-228-8172957/8172855.*

[b]*Facultad de Física. Universidad Veracruzana, Apdo. Postal 475. Xalapa, Veracruz. México.*

Abstract

We present a symmetry analysis of the distribution of variations of different financial indices, by means of a statistical procedure developed by the authors based on a symmetry statistic by Einmahl and Mckeague. We applied this statistical methodology to financial uninterrupted daily trends returns and to other derived observable. In our opinion, to study distributional symmetry, trends returns offer more advantages than the commonly used daily financial returns; the two most important being: 1) Trends returns involve sampling over different time scales and 2) By construction, this variable time series contains practically the same number of non-negative and negative entry values. We also show that these time multi-scale returns display distributional bi-modality. Daily financial indices analyzed in this work, are the Mexican IPC, the American DJIA, DAX from Germany and the Japanese Market index Nikkei, covering a time period from 11-08-1991 to 06-30-2017. We show that, at the time scale resolution and significance considered in this paper, it is almost always feasible to find an interval of possible symmetry points containing one most plausible symmetry point denoted by C. Finally, we study the temporal evolution of C showing that this point is seldom zero and responds with sensitivity to extreme market events.

Keywords: Econophysics, Symmetry Test, Returns Distribution, Gain/Loss Asymmetry, Symmetry point, Symmetry interval
PACS: 05.40, 02.50.-r, 02.50.Ng, 89.65.Gh, 89.90.+n

*Corresponding author
  Email address: alhernandez@uv.mx (A.R. Hernández-Montoya)
  URL: www.uv.mx/alhernandez (A.R. Hernández-Montoya)

1. Introduction

Properties of financial markets and their behavior are of great interest to academics, policy makers, traders and investors. These markets are complex systems and all of them present universal emerging properties of statistical nature, nowadays commonly called by the name of stylized facts [1, 2].

In particular, the empirical gain/loss asymmetry, i.e. "draw-downs in stock prices or stock index values are larger than the upward variations", is one of those stylized facts, and it is of great and current interest under both, practical/applied and academic points of view (i.e. symmetry based automatic trading systems, skewness induced risk, validity of Efficient Market Hypothesis, etc). Indeed, studying the symmetry of the distribution of financial returns has been researched for many years and it has always been an important and tricky business: results of symmetry measurements depend on the period of time of data records, the time scale analyzed, returns definition, leverage effect, etc. [3–8].

Initially, most of the market symmetry or gain/loss asymmetry studies relied on studying the third standardized moment of the price or index variations or other similar measurement. More recently, symmetry of financial variations or related problems has been approached by very ingenious methodologies, as for example, the analysis of the returns distribution of stocks ensembles during crash and rally days [9]; the study of large fluctuation dynamics under time reversal (TR) symmetry (large fluctuations dynamics at daily scale are not TR symmetric, but at the scale of high frequency data they are)[10]; study of the investment horizons distribution [11, 12]; empirical analysis of the clustering on the asymmetry properties in financial time series [13]; symmetry break mechanisms [14]; symmetry in trading volume [15]; analysis related to time-scale effects on gain/loss asymmetry in stock indices [16]; the use of the non-extensive formalism of Physics [17], or focusing in searching possible symmetry points of returns [18] and many more interesting empirical and agents modeled studies [19, 20].

Even though it has been researched for many years, the study of the symmetry of the unconditional distribution of financial returns remains, in our opinion a not fully understood subject. For instance, [21, 22] have analyzed returns of a large sample of diverse financial indices without finding important symmetry deviations or fully rejecting the symmetry hypothesis. On the other hand, many studies under different conditions and points of view have reported the emergence of asymmetries in the financial returns distribution [9–20].

By the analysis presented in this paper we expect to shed new light in what it may be grasped when we speak of symmetry of the distribution of financial

variations. Usually, in finance, a symmetric distribution of financial variations is understood as a distribution with a mirrored symmetry around the y vertical axis in $x = 0$, i.e. the mean values of the distribution coincides with the origin. In this paper, we will see that although it is possible to find a symmetry point of the variations distribution, most of the time this point it is not situated in the origin or symmetrically around it.

1.1. Organization of the present paper

This work is organized as follows: in the next section we briefly review the statistical methodology used for the symmetry analysis: we show the results of the numerical calculation of the asymptotic distribution of the T_n statistic as obtained in [18]. These values are important because we use them to find the interval where it is more plausible to observe symmetry points of market variations, as is explained in subsection 2.1. In section 3, and following [23], we present the time multi-scale observables analyzed in this work. We explain their construction, report some interesting statistical distributional properties that they display, such as bi-modality, and we include a brief digression on the reasons these observables were chosen to study the symmetry of market variations. In section 4, we apply our methodology for assessing symmetry to four data samples of different markets from the period of time including 11-08-1991 to 06-30-2017. Also, in this section we show how the most plausible symmetry point for a given confidence level from market variations evolves in time and behaves around extreme market movements. Finally, conclusions and a summary are discussed in section 5.

2. The T_n Statistic and methodology to test symmetry

Our symmetry test is based in the distribution-free test statistic T_n for symmetry testing proposed by Einmahl and Mckeague [24] and it follows an empirical likelihood approach. Suppose a sample $X_1, \ldots, X_n$ consisting of independent and identically distributed random variables with common absolutely continuous cumulative distribution function F. The null hypothesis is that there exists a point of symmetry around zero, i.e. , $H_0 : F(0-x) = 1 - F(x-0)$, for all $x > 0$ belonging to the sample.

We use the notation $F_n(-x) := F_n(0 - x)$ and $F_n(x-) := F_n(x - 0)$, where the empirical distribution function F_n is defined as

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$$

and the function I, named the indicator function, takes a value of 1 if its argument is true and 0 if its argument is false. The test statistic by Einmahl and Mckeague is:

$$T_n = -2 \int_0^\infty \log H(x) dG_n(x) = -\frac{2}{n} \sum_{i=1}^n \log H\left(|X_i|\right).$$ (1)

where G_n is the empirical distribution function of the $|X_i|$, and H satisfies:

$$\log H(x) = n F_n\left(-x\right) \log \frac{F_n(-x)+1-F_n(x-)}{2F_n(-x)} + n\left[1 - F_n\left(x-\right)\right] \log \frac{F_n(-x)+1-F_n(x-)}{2[1-F_n(x-)]}$$

More details on the above shown, can be found in [24].

In the previous cited paper by the authors [18], the asymptotic percentage points of the statistic T_n were calculated numerically and they are shown in table 1.

Cumulative Probability	Percentage point
0.50	0.659
0.75	1.258
0.85	1.768
0.90	2.200
0.95	2.983
0.975	3.798
0.990	4.909
0.995	5.768
0.999	7.803

Table 1: Asymptotic percentage points of T_n calculated numerically from [18].

These limiting distribution points were found by proving that T_n converges to:

$$T_n \xrightarrow{D} \int_0^1 \frac{W(t)^2}{t} dt$$

where W denotes a standard Wiener process.

Asymptotic points were obtained using the series representation [25]

$$T_n \xrightarrow{D} \sum_{i=1}^\infty \lambda_i \nu_i,$$

4

where $\nu_1, \nu_2, \ldots$ are independent chi-squared random variables with one degree of freedom, and $\lambda_1, \lambda_2, \ldots$ are the eigenvalues of the integral equation:

$$\int\limits_0^1 \sigma(s,t)f_i ds = \lambda_i f_i(t) \tag{2}$$

where $\sigma(s,t)$ denotes the covariance function of the process $\frac{W(t)}{\sqrt{t}}$. Finally, the asymptotic percentage points of the distribution of T_n were found solving equation 2 numerically.

It is worth mentioning that in table A1 of [26], Einmahl and Mckeague had only provided critical values of T_n based on 100,000 simulations for selected sample sizes up to n=150. Their simulation results showed a fast convergence to the asymptotic percentage points.

Values of table 1 will be used in our statistical procedure to assessing the symmetry of the financial indices variations as summarized in section 3. For further details on this methodology see again [18].

2.1. Statistical Methodology

Let us now summarize the statistical methodology based on T_n to assess symmetry of a given set of observations from an unknown probability distribution. For more details, consult [18].

As usual, given a daily financial time series of prices or indice values $P_1, \ldots P_{N+1}$ we define its corresponding time series of daily logarithmic returns, called in this paper only Returns, as $R_i := \log(P_{i+1}) - \log(P_i)$, $i = 1, \ldots N$.

Let us define for Returns time series their "*shifted returns around c*" as $S_{R_t}(c) = R_t - c$, with $t = 1, \ldots, N$, where c is a real number, not necessarily a symmetry point. In particular, for the present explanation, R_t may represent either daily logarithmic returns or the multi-scale returns, defined in next section.

We will use the time series of shifted returns to obtain the corresponding values of their test-statistic T_n calculated from $S_{R_1}(c), \ldots, S_{R_N}(c)$, and denoted $T_n(c)$ for a particular value of c.

Now, let us suppose the data to test have at least a symmetry point. Then a *plausible* value of the symmetry point (for a given significance level α), will be any real number c, such that

$$T_n(c) < T(\alpha), \tag{3}$$

where $T(\alpha)$ denotes the $\alpha-$level upper point of the distribution of T_n selected from table 1.

The interval of symmetry of our returns distribution, with a significance level α, will be the interval (C_{min}, C_{max}), where C_{min} and C_{max} are, respectively the infimum and supremum of the set of all plausible symmetry points i.e. those points satisfying equation 3.

Finally, the most plausible symmetry point for the chosen significance level α, is the one that minimizes $T_n(c)$ in the interval (C_{min}, C_{max}) and it will denoted C.

To obtain (C_{min}, C_{max}) it is necessary to determine the intersection of the curve of $T_n(c)$ versus c and the horizontal line of significance level calculated from table 1 in an interval that contains the set of all possible values of c which would not lead to the rejection of the null hypothesis of symmetry for the probability distribution of the random variable R_t at the given significance level α. Of course, C is the point in this interval that minimizes $T_n(c)$.

The previously described methodology will be illustrated for all our data samples in section 3.

2.2. Skewness vs T_n

We close this section with a few words explaining our motivation to choose a methodology based on the T_n statistic over more traditional tests based on skewness:

- It is well known that a small skewness of a distribution is a necessary condition for symmetry, but it is not a sufficient condition, i.e., a small value of skewness does not guarantee the symmetry of the corresponding distribution.

- A statistical test of symmetry or a statistic based on skewness has to overcome the above mentioned problem and for that reason should be more complex than our test making necessary to set up additional conditions to the tested distribution.

- Our assessment of symmetry is independent of the distribution of data. No additional constrains are required for its application and T_n is based in the empirical distribution function only.

- Finally, and as previously mentioned at the beginning of this work, since many years, most of the statistical studies of symmetry of financial data

were elaborated by analyzing skewness of data distribution. In our opinion, new statistic perspectives to face this important and interesting problem are needed.

3. Data Sample and construction of our observables

We construct the observables of market variations from the following financial indices daily time series:

1. American DJIA (Dow Jones Industrial Average).
2. American Nasdaq (National Association of Securities Dealers Automated Quotation).
3. Mexican IPC (Indice de Precios y Cotizaciones in Spanish of Index of Prices and Quotations in English).
4. Japanese Nikkei 225.

All data samples were downloaded from `https://finance.yahoo.com/` website, the period of time considered in our analysis is from 11-08-1991 to 06-30-2017.

3.1. Observables construction

In order to study the symmetry of financial markets variations, we construct two time multi-scale observables of financial returns calculated by analogy with daily returns and using the logarithmic differences of the extreme values of daily uninterrupted trends [1][27]. This has the advantage of not having to define a fixed and arbitrary scale for returns, since that time scale is naturally defined by the different uninterrupted trends durations. For a graphic illustration using DJIA data, showing what we mean by uninterrupted trends, see figure 1.

More formally, given a price time series $P_1, \ldots, P_n$, we define an "uninterrupted trend" of duration k, as a succession of $k+1$ consecutive values of the given time series where each value is greater than the preceding one. In this case we say that these consecutive $k+1$ increasing values form an uninterrupted uptrend. On the other hand, if each one of the $k+1$ consecutive values is smaller or equal than the preceding one we say we have an uninterrupted downtrend of length k.

Now, let us suppose that for a fixed integer $m > 0$, an uninterrupted trend (uptrend or downtrend), starts at P_m with a duration of k days: $P_m, \ldots, P_{m+k}$.

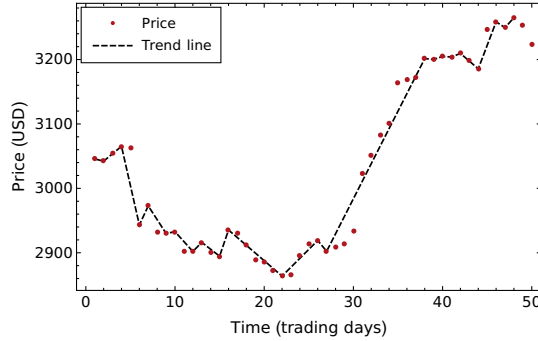[1]In probability theory, this kind of consecutive events are called "runs".

Figure 1: Uninterrupted daily trends for DJIA. Time period from Nov/08/1991 to Jan/21/1992.

In this case, we define their corresponding trend return as:

$$S_m^k := \log(P_{m+k}) - \log(P_m) \qquad (4)$$

Where m indexes the different trends and k indicates the duration in days of the m-th trend.

An other useful and closely related observable we analyze in this paper is defined as:

$$TS_m^k := \frac{\log(P_{m+k}) - \log(P_m)}{k} \qquad (5)$$

i.e. the terms of the time series of TReturns divided by their corresponding duration. It gives us an estimation of how fast trends[2] are. Let us name TS_m^k as TVReturns, meaning Trends Velocity [3] returns.

Figure 2 is divided in four panels and they will be referred as figures 2a to 2d respectively.

For illustration purposes figure 2a displays only for DJIA data, the evolution of Returns, TReturns and TVReturns for the period of time analysed in the present work. The another figures 2b to 2d show Returns, TReturns and TVReturns smoothed probability density functions respectively. A summary of their descriptive statistics can be found in table 2, showing number of entries, mean, standard deviation, skewness and kurtosis values.

As expected from the idea of aggregating same signed variations, TReturns reach the bigger amplitude fluctuations, then Returns and finally TVReturns.

---

[2]From now on, for simplicity, we will refer to uninterrupted trends only as trends

[3]Independently of knowing that TVReturns is a discrete observable, we use this term because every TVReturn is obtained dividing its corresponding TReturn by time duration in number of days

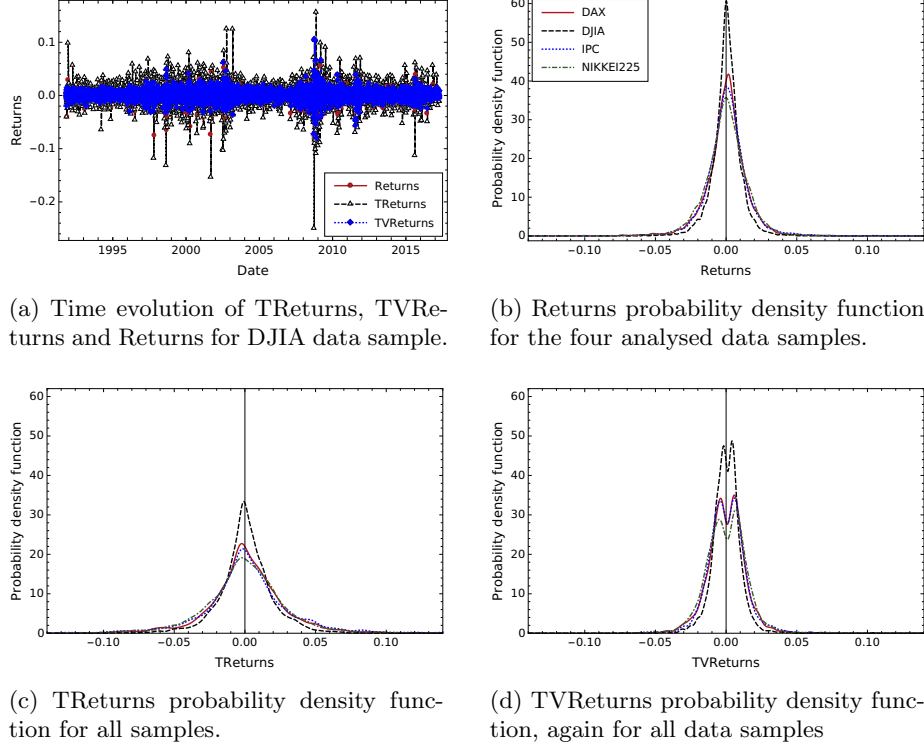Compare their RMS values in fifth column of table 2.

(a) Time evolution of TReturns, TVReturns and Returns for DJIA data sample.

(b) Returns probability density function for the four analysed data samples.

(c) TReturns probability density function for all samples.

(d) TVReturns probability density function, again for all data samples

Figure 2: a) Returns, TReturns and TVReturns evolution for DJIA sample. Figures (b),(c) and (d) Smoothed probability density function of TReturns and the constructed Multi-scale returns for all data samples.Descriptive statistics of these distributions is show in table 2

It is interesting to observe in the figure 2d that TVReturns for the four samples have a bi-modal distribution. Even more, bi-modality is also observed in the distributions of TReturns obtained from trends with duration longer than one day as is shown in all left sub-figures of figure 3, where for clarity, we have plotted TReturns and TVReturns distributions only with duration from one to four days.

The bi-modality of TReturns from trends longer than one day and TVReturns for all trends durations is explained as follows:

TReturns for a trend duration equal to one day, are a sub-sample of the uni-modal, daily returns distribution of any of our data samples; the TReturns distribution becomes bi-modal for trends with duration bigger than one day as shown in all left sub-figures of figure 3, because as trend duration increases, the magnitude of TReturns tends to increase and the probability of observing small positive and negative fluctuations around zero becomes smaller, giving place to

Market	Observable	Entries	μ	σ	Skewness	Kurtosis
DJIA	Returns	6447	$(2.9 \pm 1.3) \times 10^{-4}$	0.011	-0.176	11.56
DJIA	TReturns	3340	$(5.6 \pm 3.5) \times 10^{-4}$	0.021	-0.693	14.95
DJIA	TVReturns	3340	$(1.6 \pm 1.7) \times 10^{-4}$	0.010	0.378	13.96
IPC	Returns	6409	$(5.5 \pm 1.8) \times 10^{-4}$	0.015	0.022	9.355
IPC	TReturns	2942	$(1.2 \pm 0.6) \times 10^{-3}$	0.034	-0.241	9.678
IPC	TVReturns	2942	$(4.6 \pm 2.4) \times 10^{-4}$	0.013	0.0152	5.845
DAX	Returns	6465	$(3.1 \pm 1.7) \times 10^{-4}$	0.014	-0.106	7.300
DAX	TReturns	3274	$(6.3 \pm 5.1) \times 10^{-4}$	0.029	-0.729	10.83
DAX	TVReturns	3274	$(1.2 \pm 2.2) \times 10^{-4}$	0.013	-0.027	5.431
Nikkei	Returns	6282	$(-0.3 \pm 1.9) \times 10^{-4}$	0.015	-0.203	8.094
Nikkei	TReturns	3272	$(-0.5 \pm 5.2) \times 10^{-4}$	0.030	-0.569	11.95
Nikkei	TVReturns	3272	$(-0.6 \pm 2.4) \times 10^{-4}$	0.014	-0.257	6.813

Table 2: Descriptive statistics of Returns, TReturns and TVReturns for all data samples.

the bi-modal shape of TReturns distributions for trends duration bigger than one day. See left sub-figures of figure 3.

For the case of TVReturns, as shown in figure 2d, the bi-modal effect is clearly observed at the overall of TVReturns distributions. The reason of this is because the definition of TVReturns implies the division or TReturns by their corresponding duration in days, this operation narrows the obtained TVReturns distribution with trend duration bigger than one day, with respect to the distribution of TReturns, giving more entries closer to zero, without canceling the bi-modal behavior of TReturns and increasing the amplitude of peaks for trends durations bigger than one day. Finally, all right sub-figures on figure 3 show that the subsample of TVReturns for trends of length equal to one day, i.e. uni-modal daily returns, unlike the TReturns case, has not enough statistics to mask the bi-modal distribution of TVReturns for trends with duration bigger than one day, as is evidenced again in figure 2d.

It is important to point out that in the case of the bi-modal distributions of figure 2d, our symmetry test results in the non rejection of the assumption of symmetry, due to the fact that, even though the TVReturns series shows two modes with slightly different heights, the difference was not large enough for rejecting the null hypothesis, according to the distribution of the test statistic T_n. Additional studies have to be performed with much larger data samples to examine in detail this behavior. To have an idea of the difference in peaks amplitude of TVReturns samples, table 3 indicates their higher to lower peaks ratio.

Name	Height ratio
DJIA	1.0246
IPC	1.0233
DAX	1.0235
Nikkei	1.0878

Table 3: Ratio between higher to lower peaks for TVReturns samples. See figure 2d.

3.1.1. Justification of our selection of observables

Since in this paper we are interested in the symmetry of market indices variations, in our opinion, our selection of observables TReturns and TVReturns is more interesting and realistic for this goal than using only standard daily returns for the following reasons:

1. Sign symmetry: By construction the time series of TReturns and TVReturns alternate in sign, having the same number of negative and non-negative terms, or differing in the number of common signed entries only by one unit if the time series have an odd number of terms. Entries of the TReturns series are obtained by aggregating consecutive same signed variations, spreading its distribution in relation to the Returns distribution, as may be appreciated in figure 2 and confirmed from RMS values from Table 3. We expect this aggregation process to make the observable more sensitive to symmetry fluctuations by unmasking and emphasizing the effect on symmetry of large fluctuations composed by smaller ones. In fact small consecutive same signed variations in real life can cause large financial gains or losses. The TVReturns, obtained by the same process of aggregating consecutive common signed variations, become smaller and spread less than the TReturns and Returns due to the effect of dividing the latter by their corresponding durations. See again figure 2 and the corresponding RMS values from table 3.

2. Time multi-scale observables: The time series of TReturns and TVReturns involve in their construction different and non arbitrary time scales determined by trends duration. This makes our analysis more general, allowing us to explore a more realistic situation and giving us the opportunity to construct new indicators and observables based on this, for example market volatility from TReturns or TVReturns.

4. Data analyses

Applying the statistical methodology summarized in subsection 2.1, the interval of symmetry (C_{min}, C_{max}), i.e. the *plausible* set of values of the unknown
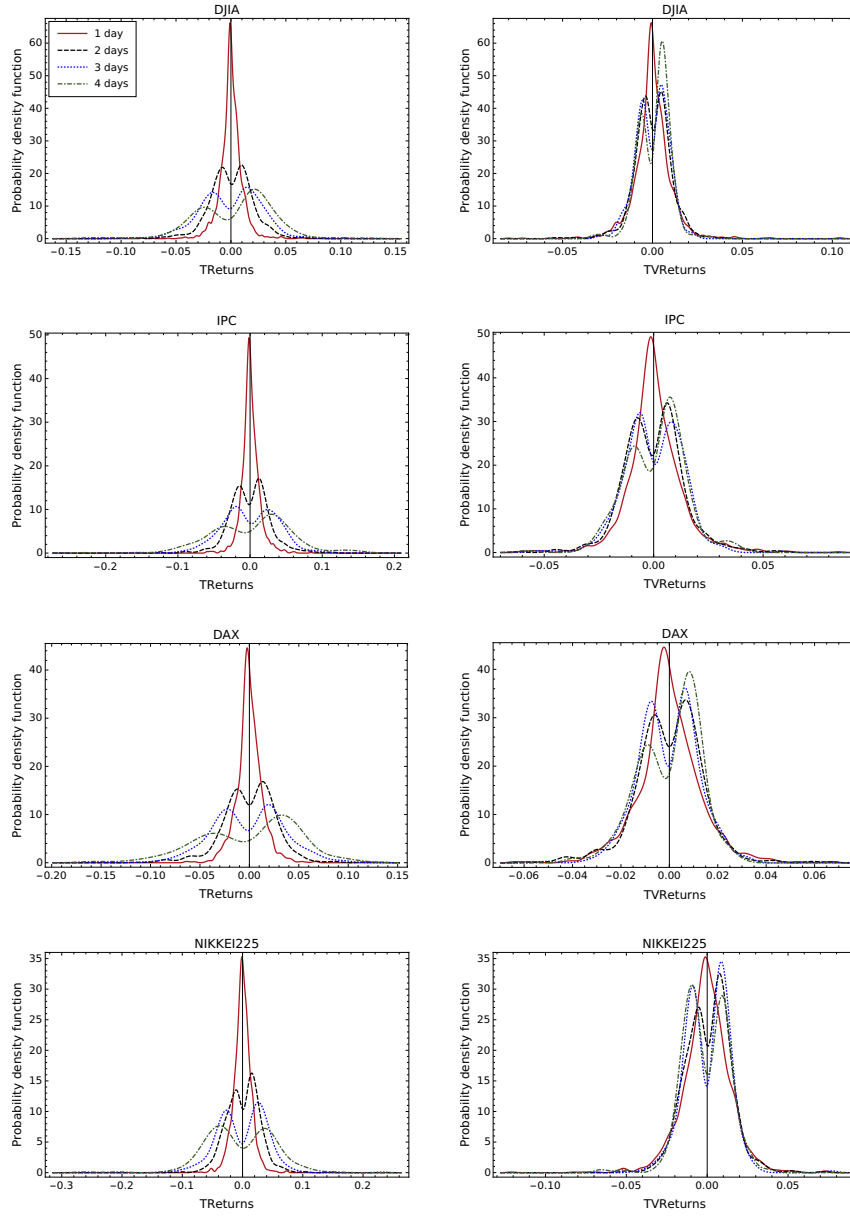
Figure 3: Bi-Modal behavior of TReturns (left column) and TVReturns (right column). We show only variations from trends not longer than 4 days.

symmetry points c, for different values of significance level α can be seen in the plots of $T_n(c)$ versus c for our multi-scale returns samples in figures 4 to 6.
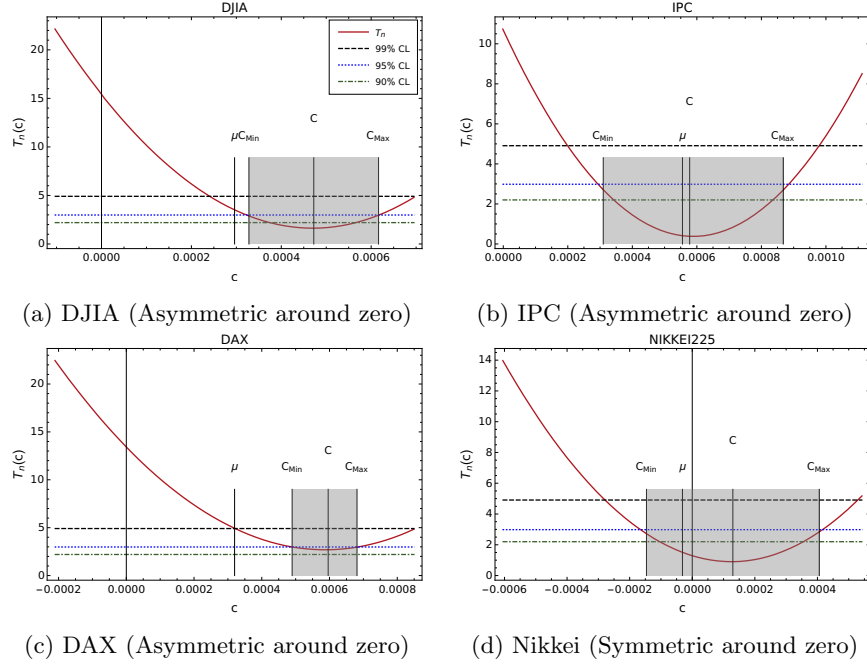
(a) DJIA (Asymmetric around zero)   (b) IPC (Asymmetric around zero)

(c) DAX (Asymmetric around zero)   (d) Nikkei (Symmetric around zero)

Figure 4: Returns statistic $T_n(c)$ versus selected values of the symmetry point c. Horizontal straight lines correspond to the 99, 95 and 90 upper percentage points, as indicated. Returns mean value μ, the origin, and the most plausible symmetry point C are indicated by vertical lines. Gray rectangles determine the interval (C_{min}, C_{max}) for $\alpha = 0.05$.

As mentioned in subsection 2.1, the symmetry interval (C_{min}, C_{max}) is determined by the intersection of the curve of $T_n(c)$ and the horizontal line of significance level calculated in table 1 in an interval that contains the set of all possible values of c which would not lead to the rejection of the null hypothesis of symmetry for the probability distribution of the analysed random variable. This interval is shown in gray for our different multi-scale returns data, in figures 4, 5 and 6, where the horizontal lines $T_n = 4.909$, $T_n = 2.983$ and $T_n = 2.200$, correspond to the asymptotic 0.99, 0.95 and 0.90 percentiles of the distribution of the T_n statistic corresponding to $\alpha = 0.01$, 0.05 and 0.1 respectively, displayed in table 1. For a significance level $\alpha = 0.10$ (or lower), in each one of our data samples, it was always possible to find an interval of plausible values for the unknown point of symmetry which would not lead us to the rejection of the assumption of symmetry. The table 4 shows all this symmetry intervals for $\alpha = 0.05$, where each point belonging to them can be statistically considered as a plausible point around which the distribution of the market variations (Returns, TReturns and TVReturns) is symmetric.

Table 4 shows a summary of the information presented visually in figures 4, 5
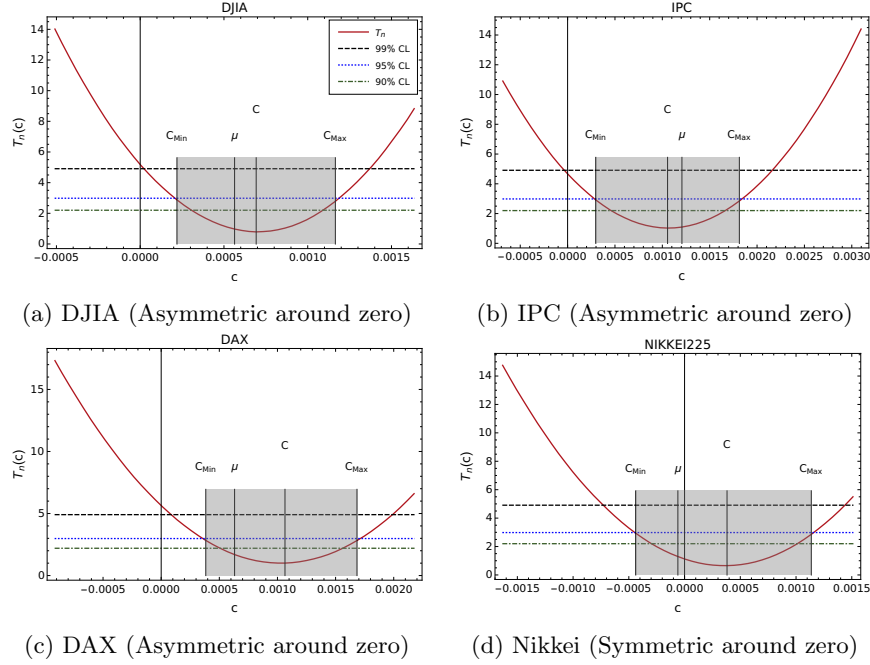
Figure 5: TReturns Statistic $T_n(c)$ versus selected values of the symmetry point c. Again, horizontal straight lines indicate the 99, 95 and 90 upper percentage points, TReturns mean μ, the origin and C are signaled by vertical lines. Again the symmetry interval is marked in gray for $\alpha = 0.05$.

and 6, including observables mean values, symmetry intervals, confidence level and whether or not the daily and multi-scale returns distributions past the symmetry test.

Again, as mentioned in [18] it is important to have in mind that our statistical approach is different from that of maximizing a test-statistic, as it is done generally, see for example in [28–30]. Instead, our methodology statistically sustains the idea that whenever there exist one or many plausible values for the points of symmetry, these plausible symmetry points may be found, and if $T_n(c)$ has a minimum value, denoted in this paper by C, this point will be the most plausible symmetry point for a given significance level α.

4.1. Time evolution of T_n and the most plausible symmetry point C

Until now we have only assessed the overall symmetry of our four data samples for three observables. We would like to have a more complete image of symmetry changes on time, i.e. we would like to see how T_n and the most plausible symmetry point C found with our method, evolve over time, at least at a coarse

(a) DJIA (Asymmetric around zero)   (b) IPC (Symmetric around zero)

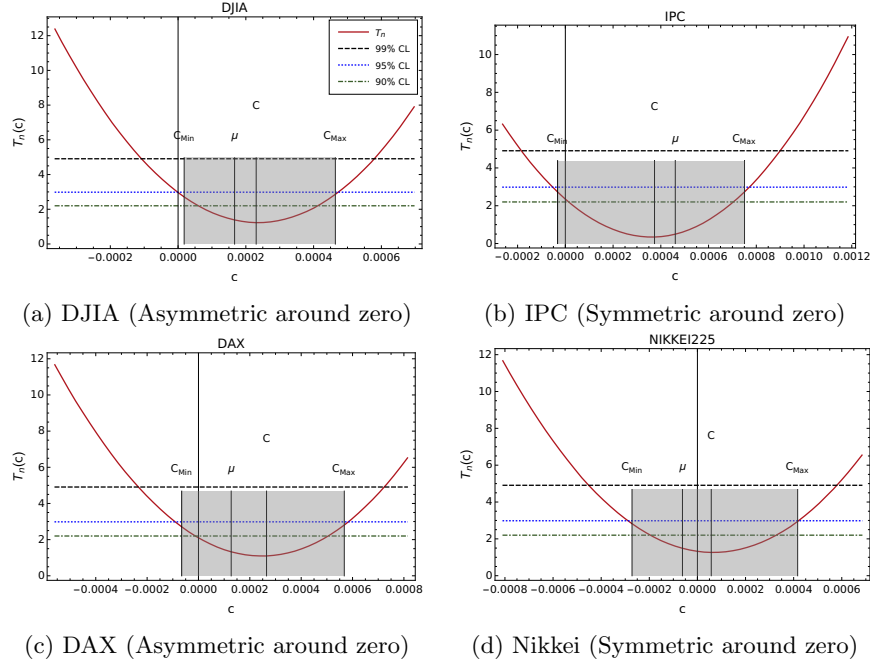(c) DAX (Asymmetric around zero)   (d) Nikkei (Symmetric around zero)

Figure 6: TVReturns statistic $T_n(c)$ versus selected values of the symmetry point c. As before, horizontal straight lines correspond to the 99, 95 and 90 upper percentage points. Mean value μ, C and the origin are again shown by vertical lines. As two previous plots (C_{min}, C_{max}) for $\alpha = 0.05$ is emphasized in gray.

grained level as determined by the chosen observables. To achieve this we use a movable time window for all our samples and observables and perform the statistical test successively. We chose a time window of 252 trading days since it is the average number of trading days in a year for the selected markets and its resolution is not enough noisy to allow us appreciate some market days with extreme movements as it will be shown in subsection 4.2. Results on the dynamics of of T_n are shown in figure 7, where these plots show that at the chosen time window resolution level, the plausibility of symmetry around $c = 0$ is variable, being rejected at different periods of time for the different significance levels. We will show explicitly the dynamics of C and its confidence interval next.

The corresponding coarse grained images of the time evolution of C for all data samples and observables were obtained. Figure 8 displays the desired results. In Appendix A we show these plots separately including its confidence interval evolution.

Market	Observable	μ	Symmetry interval	C	Zero symmetric
DJIA	Returns	$2.9{\times}10^{-4}$	$(3.2{\times}10^{-4}, 6.1{\times}10^{-4})$	$4.7{\times}10^{-4}$	No
DJIA	TReturns	$5.6{\times}10^{-4}$	$(2.0{\times}10^{-4}, 1.1{\times}10^{-3})$	$6.9{\times}10^{-4}$	No
DJIA	TVReturns	$1.6{\times}10^{-4}$	$(-1.5{\times}10^{-7}, 4.7{\times}10^{-4})$	$2.3{\times}10^{-4}$	Yes
IPC	Returns	$5.5{\times}10^{-4}$	$(2.9{\times}10^{-4}, 8.8{\times}1.0^{-4})$	$5.9{\times}10^{-4}$	No
IPC	TReturns	$1.2{\times}10^{-3}$	$(2.8{\times}10^{-4}, 1.8{\times}10^{-3})$	$1.0{\times}10^{-3}$	No
IPC	TVReturns	$4.6{\times}10^{-4}$	$(-5.3{\times}10^{-5}, 7.6{\times}10^{-4})$	$3.6{\times}10^{-4}$	Yes
DAX	Returns	$3.1{\times}10^{-4}$	$(4.8{\times}10^{-4}, 6.8{\times}10^{-4})$	$5.8{\times}10^{-4}$	No
DAX	TReturns	$6.3{\times}10^{-4}$	$(3.6{\times}10^{-4}, 1.7{\times}10^{-3})$	$1.0{\times}10^{-3}$	No
DAX	TVReturns	$1.2{\times}10^{-4}$	$(-8.8{\times}10^{-5}, 5.8{\times}10^{-4})$	$2.4{\times}10^{-4}$	Yes
Nikkei	Returns	$-3.1{\times}10^{-5}$	$(-1.6{\times}10^{-4}, 4.1{\times}10^{-4})$	$1.2{\times}10^{-4}$	Yes
Nikkei	TReturns	$-5.9{\times}10^{-5}$	$(-4.3{\times}10^{-4}, 1.1{\times}10^{-3})$	$3.6{\times}10^{-4}$	Yes
Nikkei	TVReturns	$-6.2{\times}10^{-5}$	$(-2.8{\times}10^{-4}, 4.1{\times}10^{-4})$	6.1×10^{-5}	Yes

Table 4: Mean value μ, symmetry interval, most probable symmetry point and result of our symmetry test around zero for all data samples analyzed and our three observables. Measurements were performed with a significance level $\alpha = 0.05$.

4.2. Behavior of the Symmetry point C around extreme market movements dates

By their construction, since the observables TReturns and TVReturns are coarse grained, in order to clarify the sensitivity of C on market big movements, we show only plots for daily returns from figure 8a, this time separately in figure 9, where in the graphs of evolution of C on time, we appreciate clearly the days listed in table 5 when market experienced extreme events, as well as the dates and duration of important crisis, such as the dotcom bubble and the subprime bear market. Because it is not the goal of this research, we do not compare the reliability of this methodology with other financial indicators to study extreme events or market cycles, but it is in our opinion that it can be applied to this kind of financial analyses. In fact traders use many indicators simultaneously to follow market behavior and construct trading systems. Implications for trading applications of our methodology are clear.

4.2.1. Note on the variation ranges of C and μ

It is pertinent to include here a brief comment on the small range of variation of our variable C displayed in figures 4 to 6, 8 and 9. Table 4 shows the C interval range of all our observables for all our data samples.

Figures A.10 to A.12 included in Appendix A show the evolution of C and it confidence interval for all our observables and all analyzed data samples. Small values of C must be expected, since that bigger the C value, the further the

(a) Returns

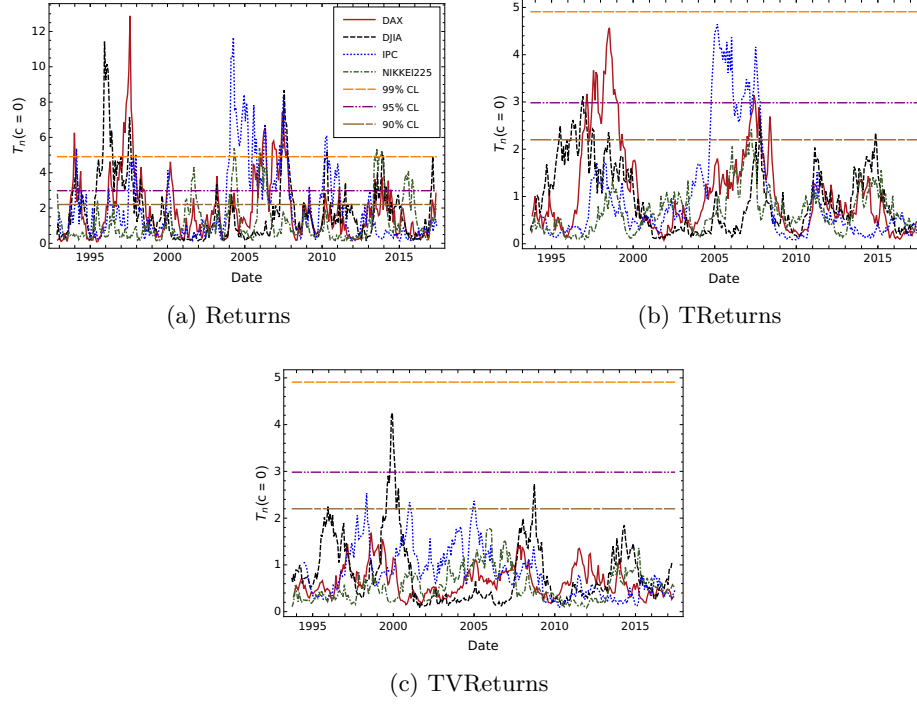(b) TReturns

(c) TVReturns

Figure 7: Plots of statistic $T_n(c = 0)$ versus date for our four different markets series data. Also in this plot, horizontal straight lines that correspond to the 99, 95 and 90 upper percentage points, are indicated.

Name	date
a) Japanese asset price bubble	1-1-1990
b) Tequila Effect	12-20-1994
c) Dotcom bubble	03-10-2000
d) Subprime crisis	08-09-2007
e) Brexit	06-23-2018

Table 5: Recent crisis and critical market days chronologically ordered and pointed out in figures 9.

market would be from efficiency. Since the market seems to be efficient most of the time, C usually must take very small values. A quantitative study of the above mentioned could be addressed in another paper. Same comment applies also to the variation range of the mean value μ of the observables analyzed here.
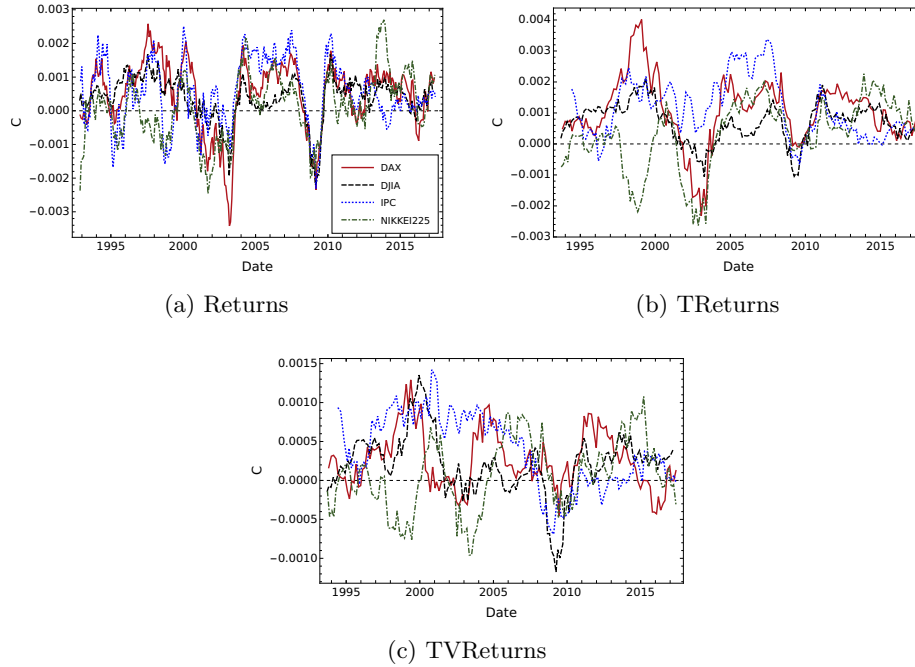
17

(a) Returns

(b) TReturns

(c) TVReturns

Figure 8: Plots of the most plausible point of symmetry C versus time for the three different observables and four data samples. A 252 trading days time window was used. To appreciate independently each plot, see Appendix A.

5. Conclusions

In this article we study the symmetry of distributions of two coarse-grained, multi-scale observables computed from daily uninterrupted trends, named TReturns and TVReturns of four financial daily data samples: DJIA, DAX, Nikkei and IPC indices. Per se, these distributions display interesting properties such as bi-modality, smoother variations, a time multi-scale nature, etc., as discussed in section 3, deserving a deeper and additional analysis to be presented in a following paper. For comparison and completeness purposes, the same analyses are performed on the usual daily returns, named here only Returns.

Studying the symmetry of the distribution of variations of stocks prices, financial index or other assets is a topic of great interest in finance. It gives us information of market fairness and how easily wealth can be created or destroyed. Also, it is strongly linked to the Efficient Market Hypothesis (EMH), which implies that for efficient markets, variations of assets prices or indices values must be symmetric. It is precisely due to this reason that people from the financial community almost automatically consider that these kind of distributions must be symmetrical around the origin. In this work we show that this is not the general case: the
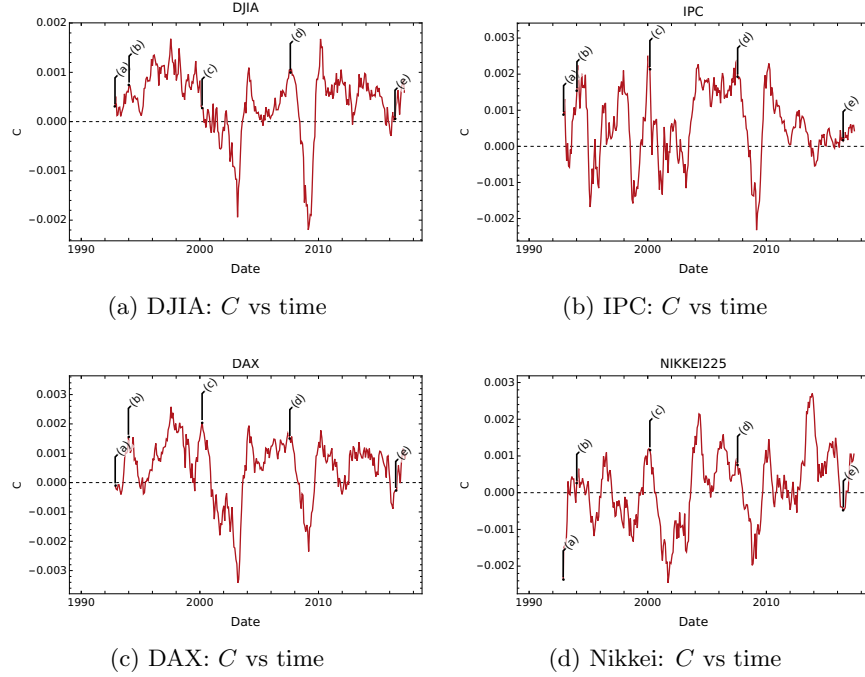
18

(a) DJIA: C vs time

(b) IPC: C vs time

(c) DAX: C vs time

(d) Nikkei: C vs time

Figure 9: Most plausible symmetry point C time evolution only for daily returns. Same time window of 252 trading days was used and graphs are displayed separately. Some important market extreme events are pointed out.

most plausible symmetry point C of financial variations fluctuates around and close to zero. Furthermore, there exist intervals containing possible symmetry points, varying over time and, although the mean value μ of financial variations may belong to the interval of symmetry, $\mu \subseteq (C_{min}, C_{max})$, it is not always necessarily the most plausible symmetry point. See figures 4 to 6.

Results of our test for symmetry around zero, with a significance level $\alpha = 0.05$ and a time window of 252 trading days can be seen in table 4, showing that for all the analyzed data samples, with the exception of Nikkei index, Returns and TReturns are not zero symmetric and that TVReturns are zero symmetric according to the symmetry test applied to all data samples. In the same table, the intervals where these observables, Returns and the constructed multi-scale returns, have symmetry points, together with their mean values and their corresponding most plausible symmetry points found with our method are displayed. We find that, in the same way as log-returns have greater plausibility of being symmetric than plain returns (defined as just price differences) [3], TVReturns have greater plausibility of being symmetric than TReturns.

We have also examined the behavior of the most plausible symmetry point C

19

of our samples of financial variations around dates with extreme market movements, see figure 9, showing that C displays a good response and sensitivity to market fluctuations. In our opinion, the sensitivity of C to these fluctuations make our methodology a good candidate to construct automatic trading systems and may be attractive to people interested in this aspect of financial markets.

According to our statistical test, it is not always possible to find symmetry intervals around zero for certain significance levels, as can be seen in figure 7. The same can be said more generally for other points that are different but close to zero.

In fact, at the confidence limits and time window resolution used in our analyses, most of the time markets seem to be close to efficiency, and then the most plausible symmetry point C as well as μ must adopt very small values and increase when the markets moves away from efficiency. Of course, using a small enough time windows would result in periods of time around big market movements where would not be possible to find an interval of symmetry points and neither a most plausible symmetry point. This issue deserves a proper and additional study and as we mentioned before, we consider that this observation could be applied to study quantitatively the efficiency of financial markets.

On the other hand, and under a different research perspective, such as studies performed using agents simulation models, our results are of interest because all acceptable agent models or of any other kind must reproduce the empirical facts of our findings.

Acknowledgments

Appendix A. Time evolution of C and its confidence interval

Below, we display the plots of evolution on time of the most plausible symmetry point C and its confidence interval for all our samples and observables. We include these plots in this appendix to make our article more readable. A time window of 252 trading days and an significance level $\alpha = 0.05$ were set up in all figures.
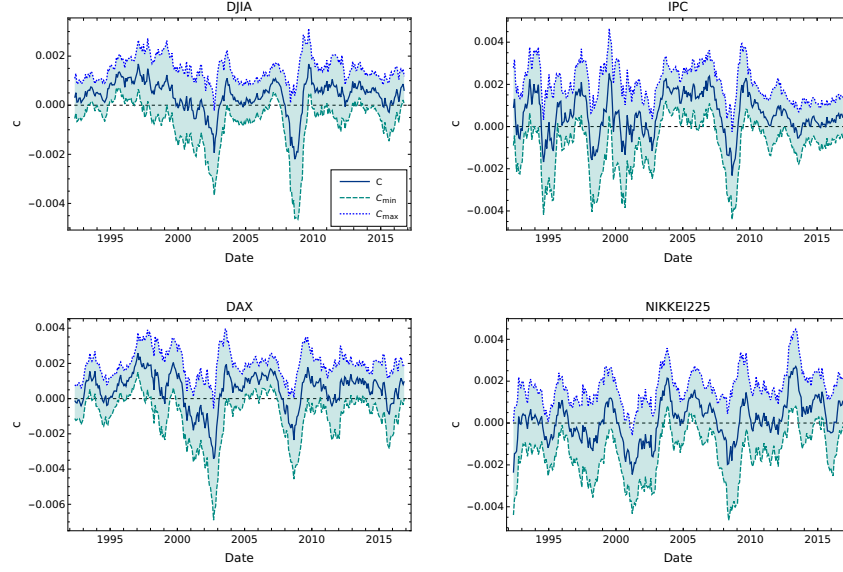
Figure A.10: Plots of the most plausible symmetry point C and the confidence interval from the lower bound C_{min} to the upper bound C_{max} for the time series of the Returns.
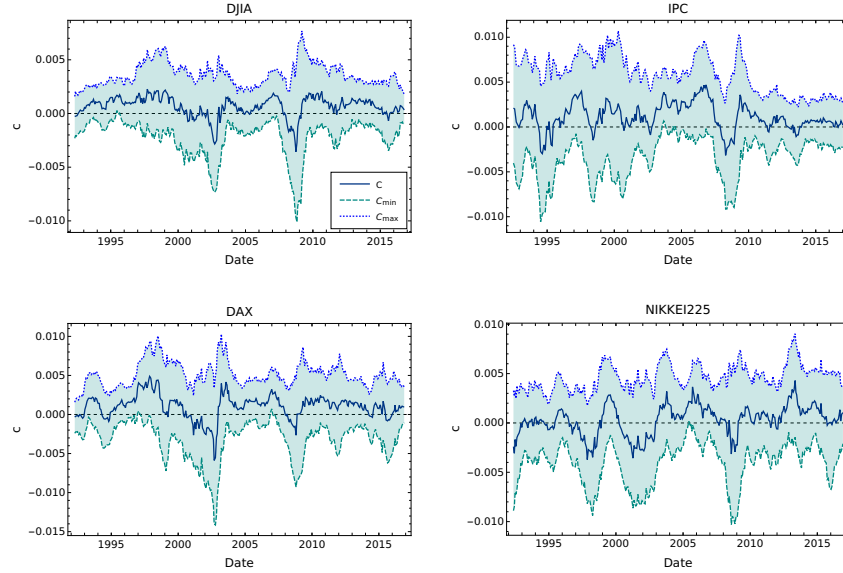

Figure A.11: Plots of the most plausible symmetry point C and the confidence interval from the lower bound C_{min} to the upper bound C_{max} for the time series of the TReturns.
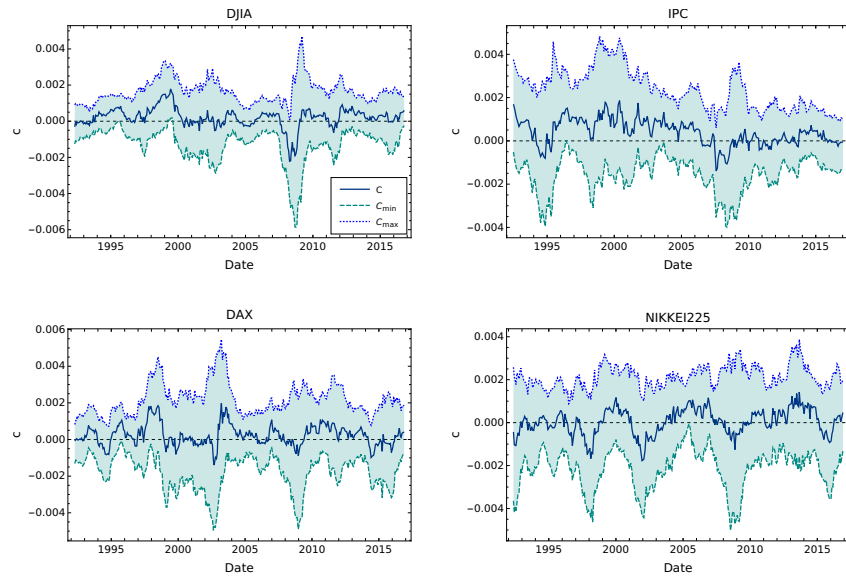
Figure A.12: Plots of the most plausible symmetry point C and the confidence interval from the lower bound C_{min} to the upper bound C_{max} for the time series of the TVReturns.

[1] N. Kaldor, Capital accumulation and economic growth, in: The theory of capital, International Economic Association, Palgrave Macmillan, London, 1961, pp. 177–222. `doi:10.1007/978-1-349-08452-4_10`.

[2] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues, Quantitative Finance 1 (2001) 223–236. `doi:10.1080/713665670`.

[3] W. L. Beedles, On the asymmetry of market returns, The Journal of Financial and Quantitative Analysis 14 (3) (1979) 653–660. `doi:10.2307/2330196`.

[4] H. R. Fogler, R. C. Radcliffe, A note on measurement of skewness, Journal of Financial and Quantitative Analysis 9 (3) (1974) 485489. `doi:10.2307/2329875`.

[5] J.P. Bouchaud, M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management, 2nd Edition, Cambridge University Press, 2009.
URL `https://books.google.com.mx/books?id=4IBjPwAACAAJ`

[6] P. T. Heden Ahlgren, M. H. Jensen, I. Simonsen, R. Donangelo, K. Sneppen, Frustration driven stock market dynamics: Leverage effect and asymmetry, Physica A: Statistical Mechanics and its Applications 383 (1) (2007) 1–4. `doi:10.1016/j.physa.2007.04.081`.

[7] A. Dutta, Modelling volatility: Symmetric or asymmetric garch models?, Journal of Statistics: Advances in Theory and Applications 12 (2) (2014) 99–108.
URL `https://pdfs.semanticscholar.org/5c18/f7ade84ed77319381a45db4f74bde243eec8.pdf?_ga=2.35273953.1115894547.1542995825-2123431015.1542995825`

[8] D. d. Almeida, L. K. Hotta, The leverage effect and the asymmetry of the error distribution in GARCH based models: The case of Brazilian market related series, Pesquisa Operacional 34 (2014) 237 – 250. `doi:10.1590/0101-7438.2014.034.02.0237`.

[9] Lillo, F., Mantegna, R. N., Symmetry alteration of ensemble return distribution in crash and rally days of financial markets, Eur. Phys. J. B 15 (4) (2000) 603–606. `doi:10.1007/s100510051162`.

[10] X. Jiang, T. Chen, B. Zheng, Time-reversal asymmetry in financial systems, Physica A: Statistical Mechanics and its Applications 392 (21) (2013) 5369 – 5375. `doi:10.1016/j.physa.2013.07.006`.

[11] M. H. Jensen, A. Johansen, I. Simonsen, Inverse statistics in economics: the gain–loss asymmetry, Physica A: Statistical Mechanics and its Applications 324 (1) (2003) 338–343. `doi:10.1016/S0378-4371(02)01884-8`.

[12] K. Karpio, M. A. Zauska-Kotur, A. Orowski, Gainloss asymmetry for emerging stock markets, Physica A: Statistical Mechanics and its Applications 375 (2) (2007) 599 – 604. `doi:10.1016/j.physa.2006.10.003`.

[13] J.-J. Tseng, S.-P. Li, Asset returns and volatility clustering in financial time series, Physica A: Statistical Mechanics and its Applications 390 (7) (2011) 1300 – 1314. `doi:10.1016/j.physa.2010.12.002`.

[14] R. Savona, M. Soumare, J. V. Andersen, Financial symmetry and moods in the market, PLOS ONE 10 (4) (2015) 1–21. `doi:10.1371/journal.pone.0118224`.

[15] M. B. Graczyk, S. M. Duarte Queirós, Intraday seasonalities and non-stationarity of trading volume in financial markets: Individual and cross-sectional features, PLOS ONE 11 (11) (2016) 1–25. `doi:10.1371/journal.pone.0165057`.

[16] B. Sándor, I. Simonsen, B. Z. Nagy, Z. Néda, Time-scale effects on the gain-loss asymmetry in stock indices, Phys. Rev. E 94 (2016) 022311. `doi:10.1103/PhysRevE.94.022311`.

[17] L. Bil, D. Grech, M. Zienowicz, Asymmetry of price return–analysis and perspectives from a non-extensive statistical physics point of view, PLOS ONE 12 (11) (2017) 1–24. `doi:10.1371/journal.pone.0188541`.

[18] H. F. Coronel-Brizio, A. R. Hernández-Montoya, R. Huerta-Quintanilla, M. E. Rodríguez-Achach, Assessing symmetry of financial returns series, Physica A: Statistical Mechanics and its Applications 383 (1) (2007) 5 – 9, econophysics Colloquium 2006 and Third Bonzenfreies Colloquium. `doi:10.1016/j.physa.2007.04.122`.

[19] J. V. Siven, J. T. Lins, Temporal structure and gain-loss asymmetry for real and artificial stock indices, Phys. Rev. E 80 (2009) 057102. `doi:10.1103/PhysRevE.80.057102`.

[20] A. M. Yamashita Rios de Sousa, H. Takayasu, M. Takayasu, Detection of statistical asymmetries in non-stationary sign time series: Analysis of foreign exchange data, PLOS ONE 12 (5) (2017) 1–18. `doi:10.1371/journal.pone.0177652`.

[21] A. Peiró, Asymmetries and tails in stock index returns: are their distributions really asymmetric?, Quantitative Finance 4 (1) (2004) 37–44. `doi:10.1088/1469-7688/4/1/003`.

[22] A. Peiró, Skewness in financial returns, Journal of Banking & Finance 23 (6) (1999) 847 – 862. `doi:10.1016/S0378-4266(98)00119-8`.

[23] H. Coronel-Brizio, A. Hernández-Montoya, H. Olivares-Sánchez, E. Scalas, Analysis of short term price trends in daily stock-market index data, arXiv preprint arXiv:1211.3060.
URL `https://arxiv.org/abs/1211.3060`

[24] J. H. Einmahl, I. W. Mckeague, Empirical likelihood based hypothesis testing, Bernoulli 9 (2) (2003) 267–290. `doi:10.3150/bj/1068128978`.

[25] J. P. Imhof, Computing the distribution of quadratic forms in normal variables, Biometrika 48 (3-4) (1961) 419–426. `doi:10.1093/biomet/48.3-4.419`.

[26] J. Einmahl, I. McKeague, Empirical likelihood based on hypothesis testing, Workingpaper, Econometrics, pagination: 25 (2002).
URL `https://pure.uvt.nl/ws/portalfiles/portal/543315/92.pdf`

[27] S. Wilks, Mathematical Statistics, Mathematical Statistics, Princeton University Press, 1943.
URL `https://books.google.com.mx/books?id=k38pAQAAMAAJ`

[28] K. Prause, The generalized hyperbolic model: Estimation, financial derivatives, and risk measures, Ph.D. thesis, PhD thesis, University of Freiburg (1999).

[29] H. F. Coronel Brizio, Regression tests of fit and some comparisons, Ph.D. thesis, Theses (Dept. of Mathematics and Statistics)/Simon Fraser University (1994).

[30] H. F. Coronel-Brizio, A. R. Hernández-Montoya, On fitting the pareto-levy distribution to stock market index data: Selecting a suitable cutoff value, Physica A: Statistical Mechanics and its Applications 354 (Supplement C) (2005) 437 – 449. `doi:10.1016/j.physa.2005.03.001`.

[31] Y. Zhang, H. Li, Investors risk attitudes and stock price fluctuation asymmetry, Physica A: Statistical Mechanics and its Applications 390 (9) (2011) 1655 – 1661. `doi:10.1016/j.physa.2011.01.002`.

[32] Y. At-Sahalia, L. P. Hansen, J. A. Scheinkman, {Chapter} 1 - operator methods for continuous-time markov processes, in: Y. Ait-Sahalia, L. P. Hansen (Eds.), Handbook of Financial Econometrics: Tools and Techniques, Vol. 1 of Handbooks in Finance, North-Holland, San Diego, 2010, pp. 1 – 66. `doi:10.1016/B978-0-444-50897-3.50004-3`.

[33] J. Durbin, Distribution Theory for Tests Based on the Sample Distribution Function, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, 1973. `doi:10.1137/1.9781611970586`.

[34] C. M. Rodríguez-Martínez, H. F. Coronel-Brizio, Tn symmetry statistic [computer program] (2016).
URL `http://github.com/CarlosManuelRodr/TnSymmetryStatistic`