Analysis of efficiency in high-frequency digital markets using the Hurst exponent

M. López[a] and R. Mansilla[b]

[a]*Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México.*
[b]*Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México.*

Received 03 June 2021; accepted 23 July 2021

In this paper, we analyze the Efficient Market Hypothesis for automated high-frequency stock markets. Using the Hurst exponent as a measure of efficiency, we show that the time series of high-frequency stock prices do not follow random walks, rejecting then (as we discuss in the text) the EMH for these markets.

Keywords: Efficiency; high-frequency trading; Hurst exponent.

PACS: 89.65.Gh; 89.70.+c; 05.90.+m

DOI: https://doi.org/10.31349/RevMexFis.67.061402

1. Introduction

The history of the relation between Physics and Finance dates back at least to the pioneering work of L. Bachelier [1], published more than a century ago. However, it was not until recent decades that a growing and well-established body of research using tools from Physics, and particularly statistical mechanics, to understand economic phenomena through time series analysis had emerged [2, 3]. To cite just a few recent examples, in Refs. [4,5] the authors apply random matrix theory to study cross-correlations in financial markets in order to describe various market states and state transitions, while in Ref. [6], a microscopic model for automatically done high-frequency transactions is presented, using tools from the kinetic theory of gases. Our work is framed in this tradition, branded as Econophysics, from which the problem of the so-called market efficiency, here discussed, has by no way escaped.

It is important to establish that when speaking of the efficiency of the markets, two great perspectives of analysis must be distinguished: the so-called distributive efficiency and the informational efficiency. It is to the second approach to which we dedicate this work. The informational efficiency of prices is defined as the immediate incorporation of all relevant information for the formation of prices through the interaction in the markets of highly sophisticated economic agents.

Introduced by Fama in Ref. [7], it has important implications for the theory, in particular, that the differences of the logarithms of the consecutive prices of assets in a market must follow a (Gaussian) random walk. This hypothesis has been widely questioned, for example, in Refs. [8, 9].

One of the most common explanations for the inefficiency of real markets has been the “animal spirit” of Keynes [10], that is, the psychological and emotional factors that lead investors to make their decisions in capital markets when there is uncertainty, how human emotions can drive making financial decisions in uncertain and volatile environments.

Another explanation comes from the work of H. Simon [11], through the concept of limited rationality, that postulates that most people are only partially rational and act on emotional impulses without rational foundations in many of his actions.

Various authors, such as Lo in Ref. [12] and McCauley in Ref. [9] for example, recover the elementary fact that financial markets are, like all social phenomena, of a historical and dynamic nature, that is, agents respond to their specific social, political, psychological and technical conditions, which is why it is inappropriate to postulate general and anti-historical hypotheses about their behavior.

As indicated in Ref. [8], in recent years, most of the operations in the large financial markets have been automated and are now computers and not human beings who make decisions by executing certain algorithms. Although in the last decades, evidence has accumulated against the efficiency of the traditional markets, one could imagine that, with the execution of orders controlled by computer algorithms, devoid of feelings and emotional decisions, efficiency could be achieved in financial markets. Thus, the following question arises: are automated markets more or less efficient than classical ones? That is, what happens to efficiency if, in the process of incorporating the relevant information for price formation, human beings are replaced by algorithms?

The objective of this work is to offer evidence, through the analysis of time series of automated transactions in the US and Mexican markets, that the use of computational algorithms in automated high-frequency markets has not led these markets to the efficiency prescribed by the neoclassical theory. The organization of the paper is as follows: Section 2 develops the theoretical framework in which the analysis will be carried out. Section 3 describes the characteristics of the data and the methodology used for our analysis. In Sec. 4, the results obtained are discussed. Section 5 contains the conclusions. Section 6 contains the bibliography used.

2. Theoretical framework

The informational efficiency of a market implies that the consecutive price differences must be independent. Indeed, if there were any correlation between consecutive prices, it could be used to perform arbitrage, an action that would be in contradiction with the assumed efficiency. Thus, one can examine the short-term movement patterns that describe the returns of the assets in the market in question and attempt to identify the process underlying those returns. If the market is efficient, the model will not be able to identify a pattern, and we will conclude that the returns follow a random walk process. If a model is able to establish a pattern, past market data can be used to predict future market movements and the market is therefore inefficient.

Note that the observation of a random walk is a necessary condition for efficiency. There are studies that show that this condition is not sufficient [13,14]. Consequently, the deviations of a random walk allow rejecting the informational efficiency of the assets under study.

The method that we will use has its origins in the work of Hurst [15], framed in the context of his studies in hydrology and later refined by Mandelbrot and Wallis [16]. Given a time series $x_n, n = 1, \ldots, m$, with mean $\mu = (1/m)\sum_{i=1}^{m} x_i$ and variance $\sigma^2 = (1/m)\sum_{i=1}^{m}(x_i - \mu)^2$, we define its partial centered sums as $y_n = \sum_{i=1}^{n}(x_i - \mu)$ and its R/S statistics or rescaled range as the ratio between the range of the partial centered sums series and the standard deviation of the original series:

$$R/S = \frac{\max_{n\leq m} y_n - \min_{n\leq m} y_n}{\sigma}.$$

Hurst noted that the rescaled range of the time series of annual flows of the Nilo river as a function of the length n of the series was asymptotically a power law when n tends to infinity: $\mathrm{E}(R/S(n)) \sim n^H$ for n sufficiently large, where $\mathrm{E}(R/S(n))$ is the mean of the R/S statistics calculated on the subseries of length n of the original series. The exponent H of the power law is known as the *Hurst exponent*. It is known that under the hypothesis of the original series being a random walk, the exponent is $H = 0.5$ [16]; instead, Hurst found $H > 0.5$.

In general, a Hurst exponent H greater than 0.5 is associated with the long-term persistence of the series: the range grows faster than expected from a random walk, that is, movements in one direction follow, with greater probability, movements in the same direction; while $H < 0.5$ is associated with long-term antipersistence: the range grows more slowly than that of a random walk, that is, movements in one direction are more likely to follow movements in the other direction, this on average and for large enough lengths. In both cases, the deviation of H from its hypothetical value 0.5 can be taken as a long-term memory measure: the movements of the series are not independent of the remote past.

The study of long-term correlations measured by Hurst exponent has been applied in Physics, for example, in Ref. [17] to the ion saturation current fluctuations and Ref. [18] to gamma, ray data. In the context of the financial time series that concerns us here, this long-term memory translates into deviations from market efficiency.

The remaining part of this section will be devoted to briefly discuss the approaches in the literature that will be used in the next section to define our methodology, as well as some results related to those of this work.

In Ref. [19], it is argued that, given a sequence of independent and identically distributed random variables, the shape of the probability distributions of the random variables affects the Hurst exponent of the series. The authors calculate the Hurst exponent H_{stock} for the daily series of the S&P500 index. Then, they shuffle the series in order to remove its memory, after which they calculate the Hurst index of the shuffled series, denoted by H_{perm} (actually, this process of shuffling and calculating is repeated a certain number of times, H_{perm} is defined as the mean and the standard deviation is reported). The difference between H_{stock} and H_{perm} is an indicator of the memory of the original series, while if H_{perm} is different of 0.5. This is attributed to the distributions of the variables, since the shuffled series are, by construction, memoryless. It is proposed then to use H_{perm} as an indicator of lack of memory, instead of the canonical value 0.5; that is, $H_{\text{stock}} > (<) H_{\text{perm}}$ would be an indicator of (anti)persistency. In other words, the significant null hypothesis will be not $H_{\text{stock}} = 0.5$, but $H_{\text{stock}} = H_{\text{perm}}$.

In Ref. [20], it is argued that the R/S statistic is sensitive to short-term correlations so that if for a time series is obtained $H \neq 0.5$, this is not enough to conclude the presence of long-term memory. In Ref. [21], the authors propose that, to ensure that the results of the R/S analysis are due to long-term correlations, the following experiment is carried out: the series is divided into blocks of, for example, 50 elements each one, and the elements within each block are permuted to destroy the short-range correlations. With this new series, the previous analyzes are repeated, and the long-term memory is corroborated if the change in the Hurst exponent is insignificant.

In Refs. [22, 23], the need to observe the evolution of the efficiency, measured by the Hurst exponent, over time is stated. The first article uses one-minute resolution data from 1983 to 2009 from the SP500. The Hurst exponent of the daily subseries is calculated, and a decreasing evolution is observed from 0.8 towards 0.5, with a statistically insignificant difference for the period 2005-2009. Similar results are obtained for the monthly exponents. For the purposes of this project, it is important to underline their explanation of the phenomenon: they attribute it to the growth of algorithmic trading. In the second article, daily exponent series are studied of eleven emerging markets between 1992 and 2002, with similar results.

Other perspectives on market efficiency by studying Hurst exponent had been proposed in the past. Very interesting is the discussion in Ref. [24], in which the authors study a measure of quantitative correlation between theoretical inefficiency and empirical predictability for 60 financial indices from different countries. The Hurst exponent is taken as a measure of market inefficiency, while to measure predictability, they use one of the most basic techniques of supervised machine learning: Nearest Neighbor (NN) and its proportion of correct answers. A considerable positive correlation (around 60%) between inefficiency and predictability is reported.

In Ref. [25], it is shown that before the great economic collapses of 1929, 1987, and 1998 a clear decrease in the Hurst exponent from the persistence regime ($H > 0.5$) to the anti-persistence regime ($H < 0.5$) is observed.

In Ref. [26], Peters carries out an extensive analysis of the R/S statistic, the relevance of which he argues through the Fractal Market Hypothesis (FMH) as an alternative to the Efficient Market Hypothesis (EMH). The fractal properties of the financial time series would be due to the differences in the time horizons of the financial agents, who, according to their interests, incorporate certain pieces of relevant information into the price of an asset that is not relevant for other time horizons. Market stability is attributed to the dynamic interaction of these different scales.

	Market Code	Name
1	ABT	Abbott Laboratories
2	BAC	Bank of America Corporation
3	BMY	Bristol-Myers Squibb Company
4	C	Citigroup Inc.
5	CSCO	Cisco Systems, Inc.
6	F	Ford Motor Company
7	FB	Facebook, Inc.
8	FOXA	Tweenty-First Century Fox, Inc.
9	GE	General Electric Company
10	GM	General Motors Company
11	HPQ	HP Inc.
12	INTC	Intel Corporation
13	KO	The Coca-Cola Company
14	MDLZ	Mondelez International, Inc.
15	MO	Altria Group, Inc.
16	MS	Morgan Stanley
17	MSFT	Microsoft Corporation
18	ORCL	Oracle Corporation
19	PFE	Pfizer Inc.
20	T	AT&T Inc.
21	TWTR	Twitter, Inc.
22	USB	U.S. Bancorp
23	VZ	Verizon Communication, Inc.
24	WFC	Wells Fargo & Company

FIGURE 1. Assets of the US market.

3. Nature of the data and methodology used

The data used in this investigation are the time series of prices of the automated (algorithmic) operations that occurred from March 7, 2018, to March 7, 2019, in the Mexican and US markets (251 trading days). For the US market, there were 539,834,024 records, and for the Mexican market, 78,863,574 records.

The importance for this work that our data comes from fully-automated transactions cannot be overstated: it is for this alone that we can test efficiency specifically for digital markets.

The data belongs to 59 assets; of these, 35 correspond to companies listed on the Mexican stock exchange: AC, ALSEA, ALPEK, ALPHA, AMX, ASUR, BIMBO, BSMX, CEMEX, CUERVO, ELEKTRA, FEMSA, GAP, GCARSO, GENTERA, GFINBUR, GFNORTE, GMEXICO, GMXT, GRUMA, IENOVA, KIMBER, KOF, LALA, LIVEPOL, MEGA, MEXCHEM, NEMAK, OMA, PENOLES, PINFRA, RA, TLEVISA, VOLAR, WALMEX; while the other 24 are from the US market: ABT, BAC, BMY, C, CSCO, F, FB, FOXA, GE, GM, HPQ, INTC, KO, MDLZ, MO, MS, MSFT, ORCL, PFE, TWTR, T, USB, WFC, VZ. The details can be seen in Figs. 1 and 2.

The following analysis is carried out for the series of logarithmic returns

$$r(t, \tau) = x(t + 1, \tau) - x(t, \tau),$$

where $x(t, \tau)$ is the t-th term of the series of means of τ seconds of logarithms of the prices of a given asset.

The Hurst exponent of a time series is obtained as follows: Given n less than the length of the series, we calculate R/S of all subseries of length n of the original series and define $E(R/S(n))$ as the average of these calculations. Finally, the Hurst exponent of the series is calculated as the exponent of the function $c \cdot n^H$ that best fits (in the least-squares sense) the function $n \mapsto E(R/S(n))$ for n large enough. Taking into account the asymptotic nature of the Hurst exponent, as well as the sensitivity of the R/S statistic to the length of the time series [27], uniformly spaced values of n are taken on a logarithmic scale, with a minimum n of the order of 2^9.

To analyze the evolution of the Hurst exponent throughout the period under study, which is one year, we calculate the Hurst exponent H_{stock} of subseries of a certain number N of days, slid one day at a time [23]. For example, if $N = 5$, the Hurst exponent of the first five days of the series is calculated, then that of the series that goes from the second to the sixth day, etc., and the last calculation is for the series of the last five days, with which the evolution of the weekly Hurst exponent H_{stock} throughout the year is obtained.

There is no satisfactory analytical theory for the R/S statistic; most of the results on the subject are derived from computer simulations, which implies that they depend on particular models. Thus, although R/S is non-parametric, it is usually used to test the null hypothesis of Gaussian random walk [26], so its rejection may be due to non-Gaussianity or short-term memory. That is why the methodology that will

	Market Code	Name
1	AC	Arca Continental, S.A.B. de C.V.
2	ALFA	Alfa, S.A.B. de C.V.
3	ALPEK	Alpek, S.A.B. de C.V.
4	ALSEA	Alsea, S.A.B. de C.V.
5	AMX	América Móvil, S.A.B. de C.V.
6	ASUR	Grupo Aeroportuario del Sureste, S.A.B. de C.V.
7	BIMBO	BIMBO & Grupo Bimbo, S.A.B. de C.V.
8	BSMX	Grupo Financiero Santander, S.A.
9	CEMEX	Cemex, S.A.B. de C.V.
10	CUERVO	Becle, S.A.B. de C.V.
11	ELEKTRA	ELEKTRA & Grupo Elektra, S.A.B. de C.V.
12	FEMSA	Fomento Econóomico Mexicano, S.A.B. de C.V.
13	GAP	Grupo Aeroportuario del Pacífico, S.A.B. de C.V.
14	GCARSO	Grupo Carso, S.A.B. de C.V.
15	GENTERA	Gentera, S.A.B. de C.V.
16	GFINBUR	Grupo Financiero Inbursa, S.A.B. de C.V.
17	GFNORTE	Grupo Financiero Banorte, S.A.B. de C.V.
18	GMEXICO	Grupo México, S.A.B. de C.V.
19	GMXT	Gruma, S.A.B. de C.V.
20	GRUMA	Infraestructura Energética Nova, S.A.B. de C.V.
21	IENOVA	Kimberly Clark de México, S.A.B. de C.V.
22	KIMBER	U.S. Coca-Cola Femsa, S.A.B. de C.V.
23	KOF	GMexico Transportes, S.A.B. de C.V.
24	LALA	Grupo Lala, S.A.B. de C.V.
25	LIVEPOL	El Puerto de Liverpool, S.A.B. de C.V.
26	MEGA	Megacable Holdings, S.A.B. de C.V.
27	MEXCHEM	MexicHEM, S.A.B. de C.V.
28	NEMAK	Nemak, S.A.B. de C.V.
29	OMA	Grupo Aeroportuario del Centro Norte, S.A.B. de C.V.
30	PENOLES	Industrias Peñoles, S.A.B. de C.V.
31	PINFRA	Promotora y Operadora de Infraestructura, S.A.B. de C.V.
32	RA	Regional, S.A.B. de C.V.
33	TLEVISA	Grupo Televisa
34	VOLAR	Controladora Vuela Compañía de Aviación, S.A.B. de C.V.
35	WALMEX	Walmart de México, S.A.B. de C.V.

FIGURE 2. Assets of the Mexican market.

be used below to establish the statistical significance of our calculations, inspired by the proposals discussed in Sec. 2, is based on global and local permutations of the series in question [19, 21, 23].

Continuing with the previous example with $N = 5$, for each subseries of five days of the original series, we shuffle its terms to destroy its memory, and the Hurst exponent is calculated for this new randomized subseries. The process of shuffling and calculating the Hurst exponent is repeated one hundred times, thus obtaining a statistical sample, which we will call H_{perm}, of the Hurst exponent of the subseries under the null hypothesis of lack of long-term memory, so we can use its quantiles to test the statistical significance of the difference between H_{stock} and H_{perm}.

To rule out that the results thus obtained are due to short-term memory, we obtain similarly a statistical sample of locally randomized Hurst exponents. Given a weekly subseries ($N = 5$) and a fixed length l, the subseries is divided into blocks of l elements, and the elements within each block are shuffled to destroy short-term correlations without altering the long-range memory structure. This process is repeated a hundred times to get a statistical sample which we will call H_{locperm}. Thus, if not only H_{stock} but also H_{locperm} is statistically different from H_{perm}, then it is ruled out that the rejection of the null hypothesis is due to short-range correlations.

We make all these calculations for each subseries of N days, so we can observe the evolution of H_{stock}, H_{perm}, and H_{locperm} throughout the year.

4. Results and discussion

In Figs. 3 and 4 we plot for $\tau = N = 1$ and $\tau = N = 5$, respectively, the evolution of H_{stock} (blue curve) and the area between the 0.1 and 0.9 quantiles of H_{perm} (purple zone). Thus, when the blue curve passes outside this area, it is concluded that the correspondent original (daily or weekly) subseries has long-term memory: the difference between H_{stock} and H_{perm} is statistically significant; while when the H_{stock} curve passes inside, the randomness of the subseries cannot be ruled out: the difference between H_{stock} and H_{perm} is not statistically conclusive.

Figure 3 shows that for the US market, it is not possible in general to reject the null hypothesis $H_{\text{stock}} = H_{\text{perm}}$ when $\tau = N = 1$ (daily series of one-second averages), while the inspection of Fig. 4 allows concluding the existence of a clear tendency to anti-persistence ($H_{\text{stock}} < H_{\text{perm}}$) for $\tau = N = 5$.

In what follows, we will focus on the latter case. Figure 5 shows for $l = 300$ the effect of locally shuffling the series to destroy their short-term memory. As before, we plot H_{stock} and the 0.1 and 0.9 quantiles of H_{locperm}. Although the Hurst exponent tends to increase slightly after the local shuffling, the global H_{locperm} shape is considerably similar to H_{stock}, reinforcing the idea that its behavior reflects well the long-term memory from the original series. Note that this tendency to increase H_{locperm} is not valid for all assets, for example, F.

Once we have visually detected the general trend towards anti-persistence and the effect of local shuffling, we can define two annual inefficiency indices, one of them given by the percentage of windows whose Hurst exponent H_{stock} is below the 0.1 quantile of H_{perm}, the second by the percentage of windows such that the mean of H_{locperm} mean does the same. We will call each of the weak and strong anti-persistence indices, and we will denote them by I_d, I_f respectively.

To be more specific, if H^q_{perm} is the q-quantile of H_{perm} and

$$N_d = \#\big\{\text{sliding } N\text{-days windows such that } H_{\text{stock}} < H^{0.1}_{\text{perm}}\big\},$$

then, since our data consist of 251 trading days and therefore we have $251 - N + 1$ sliding N-days windows, we set $I_d := N_d/(251 - N + 1)$. Analogously we define $I_f := N_f/(251 - N + 1)$, where

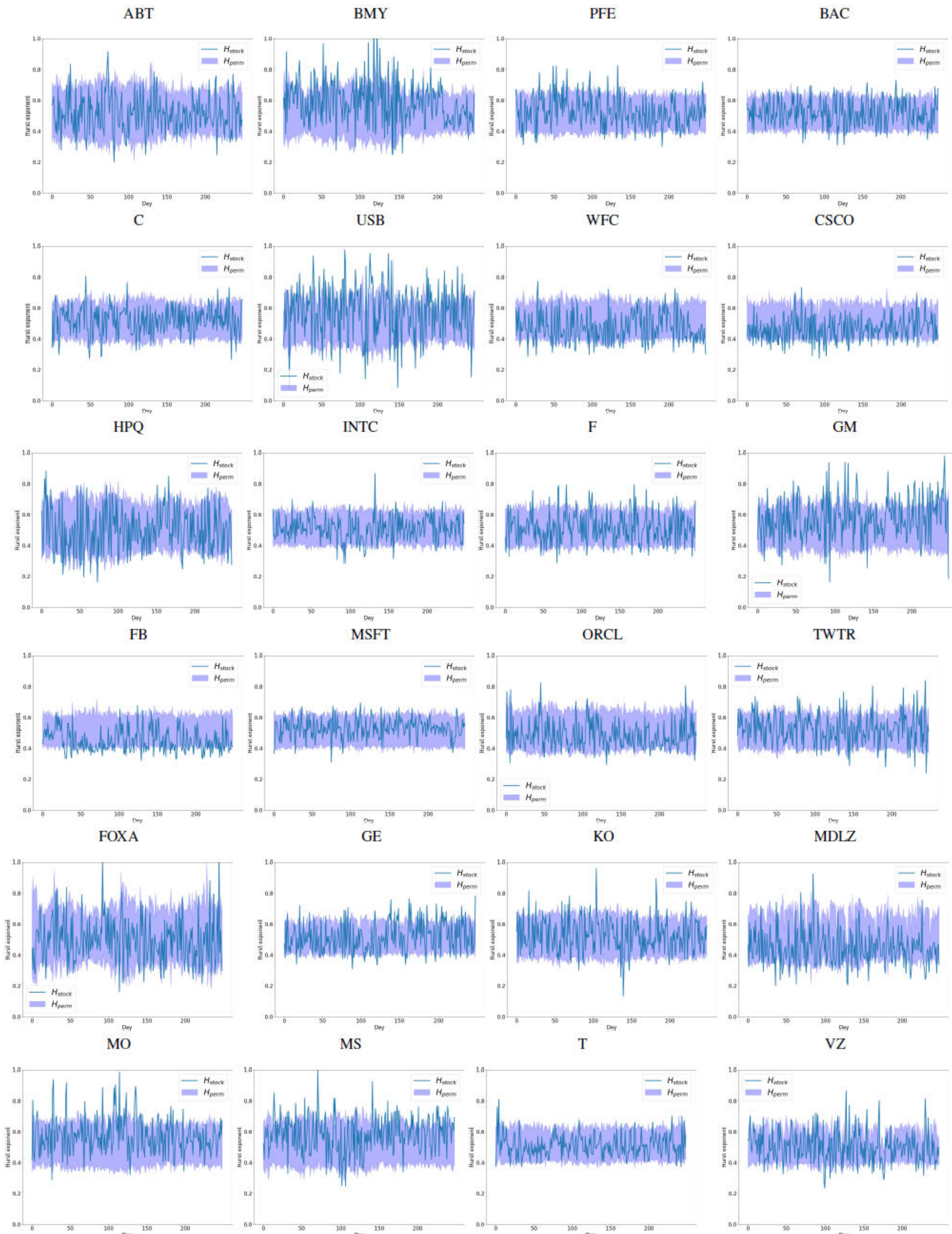

FIGURE 3. Evolution of H_{stock} and 0.1 and 0.9 quantiles of H_{perm} for the series of US market for $\tau = 1$ y $N = 1$.

FIGURE 4. Evolution of H_{stock} and 0.1 and 0.9 quantiles of H_{perm} for the series of US market for $\tau = 5$ y $N = 5$.

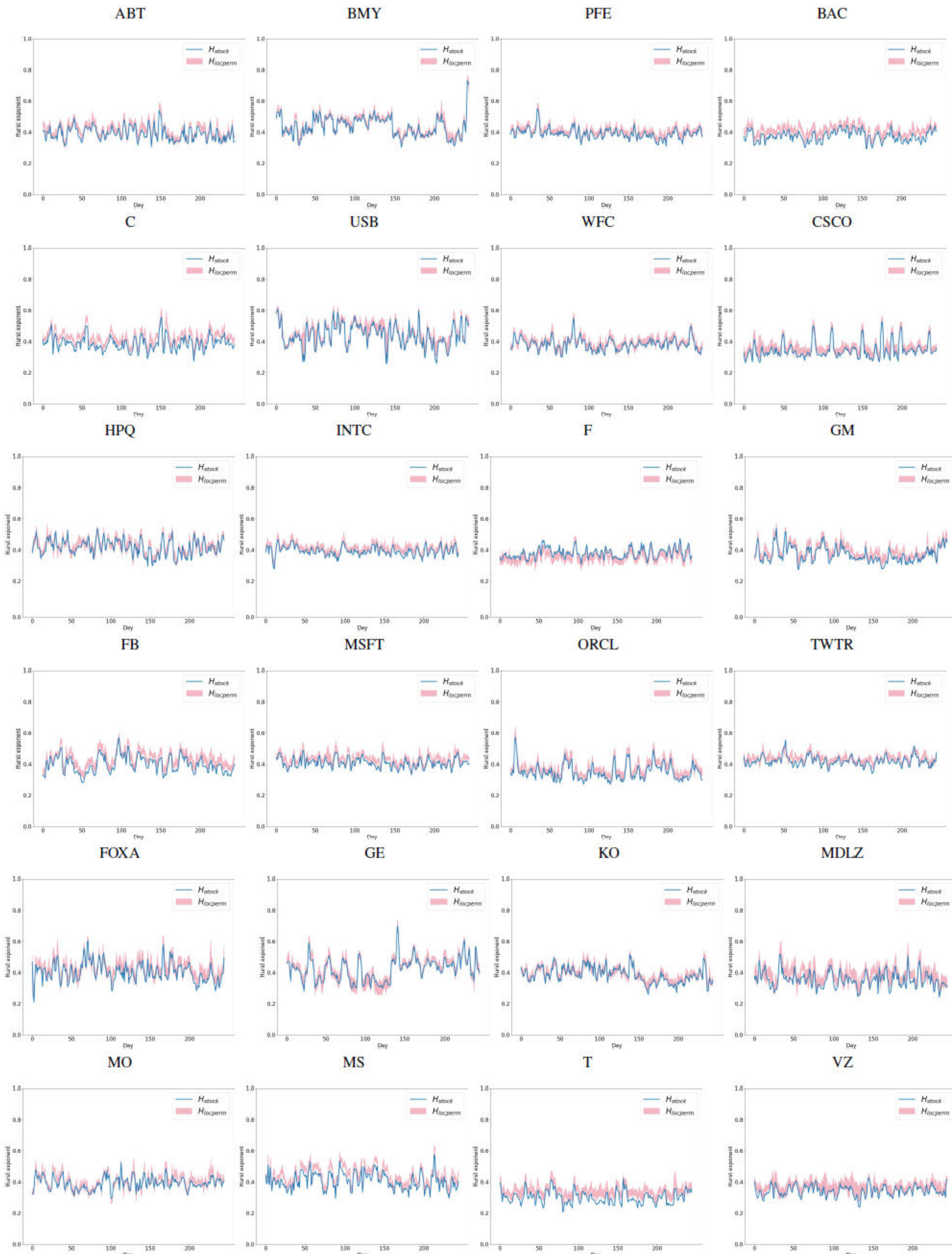

FIGURE 5. Evolution of H_{stock} and 0.1 and 0.9 quantiles of H_{locperm} for the series of the US market for $\tau = 5$, $N = 5$ y $l = 300$.

	I_d	I_f
F	0.82	0.92
ABT	0.67	0.45
KO	0.67	0.61
FB	0.65	0.36
HPQ	0.42	0.33
MSFT	0.57	0.26
GE	0.44	0.4
MO	0.64	0.52
MS	0.5	0.26
VZ	0.98	0.91
INTC	0.74	0.45
BAC	0.83	0.48
MDLZ	0.85	0.69
GM	0.76	0.61
C	0.71	0.34
FOXA	0.51	0.36
USB	0.31	0.21
ORCL	0.88	0.77
WFC	0.78	0.66
CSCO	0.88	0.86
PFE	0.75	0.57
T	0.99	0.95
BMY	0.35	0.32
TWTR	0.4	0.12
Average US	0.67	0.52

FIGURE 6. Values of I_d and I_f for the US market for $\tau = 5$ y $N = 5$.

$$N_f = \#\left\{\text{sliding } N\text{-days windows such that } \mathrm{E}(H_{\text{locperm}}) < H^{0.1}_{\text{perm}}\right\}$$

and $\mathrm{E}(H_{\text{locperm}})$ is the mean of the statistical sample H_{locperm}. Figure 6 shows the table of I_d and I_f results by asset and its average per market (US). It is concluded that most of the assets spend a significant part of the year under the anti-persistence regime.

It is important to note that even in the case of assets with a low level of inefficiency measured with these indices (TWTR, for example), the antipersistence trend is clear given that the H_{stock} and H_{locperm} curves remain in the lower part of the H_{perm} zone throughout the year (recall that the antipersistence threshold that we define: the 0.1 quantile of H_{perm}, is as arbitrary as the more traditional 0.05), a result that is hard due to the statistical sensitivity of the methods used, since it is observed systematically in all assets and throughout the year. Thus, although they are useful as summary indicators, they should not be considered as the ultimate criterion of efficiency. These observations on the qualitative nature of the process are possible thanks to the use of a dynamic approach to observe the evolution of the Hurst exponent [23], as opposed to the more traditional method of calculating a single exponent for each series, a method that reduces the problem to a purely quantitative and static criterion.

To formalize this idea and obtain, also here, a quantitative indicator: considering the subset of the $M = \lfloor 251/N \rfloor$ consecutive weekly series without overlap (where $\lfloor n \rfloor$ is the largest integer less than or equal to n) and given $q \leq 0.5$ and n_d the number of these weekly series such that H_{stock} is less

	$q=0.1, n=n_d$	$q=0.5, n=n_d$	$q=0.1, n=n_f$	$q=0.5, n=n_f$
F	9.38e-27	8.88e-16	4.86e-36	8.88e-16
ABT	2.7e-27	8.88e-14	2.38e-09	2.18e-12
KO	3.41e-23	1.78e-15	4.64e-16	8.88e-14
FB	1.85e-19	2.18e-12	9.56e-08	2.86e-08
HPQ	5.36e-07	3.49e-11	0.008	4.11e-10
MSFT	3.71e-17	8.88e-14	0.00269	3.49e-11
GE	1.57e-08	3.8e-09	1.57e-08	9.82e-07
MO	1.15e-20	1.78e-15	4.64e-16	8.88e-14
MS	2.38e-09	1.78e-15	0.0215	9.82e-07
VZ	1e-49	1.78e-15	1.4e-40	1.78e-15
INTC	6.94e-26	1.78e-15	4.28e-11	1.78e-15
BAC	3.55e-28	8.88e-16	8.87e-12	8.88e-16
MDLZ	7.55e-37	1.78e-15	1.15e-20	1.78e-15
GM	2.7e-27	8.88e-14	1.85e-19	2.18e-12
C	2.7e-27	8.88e-14	0.000226	3.49e-11
FOXA	5.61e-13	3.49e-11	5.69e-05	2.86e-08
USB	1.31e-05	7.1e-05	0.0215	0.0047
ORCL	1.4e-40	8.88e-14	2.95e-30	8.88e-14
WFC	1.98e-33	8.88e-14	6.54e-22	8.88e-14
CSCO	7.55e-37	8.88e-14	7.55e-37	3.49e-11
PFE	2.7e-27	8.88e-14	3.71e-17	8.88e-14
T	1e-49	1.78e-15	9.57e-45	1.78e-15
BMY	5.36e-07	2.86e-08	5.36e-07	9.82e-07
TWTR	9.56e-08	2.18e-12	0.215	2.18e-12
Average US	5.96e-07	2.96e-06	0.0112	0.000196

FIGURE 7. Values of $P^q(n)$ for the US market $\tau = 5$, $N = 5$ y $q = 0.1, 0.5$.

	I_d	I_f
GMEXICO	1.0	1.0
ELEKTRA	0.25	0.35
PENOLES	0.94	0.76
BSMX	0.54	0.92
MEXCHEM	0.23	0.34
GFNORTE	0.44	0.44
AMX	0.93	0.95
ALFA	0.84	0.97
CEMEX	0.88	0.86
PINFRA	0.49	0.5
WALMEX	0.54	0.53
KIMBER	0.67	0.74
GAP	0.47	0.62
RA	0.23	0.33
ALSEA	0.68	0.9
CUERVO	0.77	0.8
IENOVA	0.29	0.41
AC	0.53	0.94
MEGA	0.36	0.52
VOLAR	0.6	0.51
GFINBUR	0.65	0.78
LIVEPOL	0.54	0.61
GMXT	0.65	0.67
NEMAK	0.57	0.64
LALA	0.58	0.54
BIMBO	0.73	0.53
ASUR	0.46	0.7
GRUMA	0.69	0.89
TLEVISA	0.73	0.77
KOF	0.67	0.6
OMA	0.62	0.58
GENTERA	0.15	0.1
ALPEK	0.79	0.78
FEMSA	0.65	0.66
GCARSO	0.37	0.36
Average MX	0.59	0.65

FIGURE 8. Values of I_d and I_f for the Mexican market for $\tau = 30$ y $N = 30$.

than H^q_{perm}, the q-quantile of H_{perm}, and assuming the M consecutive weekly series as independent experiments, that is, assuming that these series are statistically independent of each other (efficiency, lack of memory at that scale), what is the probability of observing, as we do, at least n_d windows (realizations) in which H_{stock} is below H^q_{perm}? This problem is equivalent to determining the probability of obtaining at least n_d heads in a sequence of M tosses with an (unfair) coin with probability q of observing a head in each realization. This probability is modeled with the binomial distribution:

$$P^q(n_d) = \sum_{i=n_d}^{M} \binom{M}{i} q^i (1-q)^{M-i}.$$

Thus, this p-value indicates how likely it is to observe the behavior described above in an efficiency scenario given by the independence between non-overlapping weekly series. Once again, we define a strong version of this index given by $P^q(n_f)$, where n_f is the number of weekly series such that the mean of H_{locperm} is less than the quantile q of H_{perm}. The results for $q = 0.1$ and $q = 0.5$ are shown in Fig. 7. The evidence against the null efficiency hypothesis thus formulated is compelling. Except for the maximum value in the table, which is obtained for TWTR with $P^{0.1}(n_f) = 0.215$, all stocks clearly reject the null hypothesis with a 95% level of confidence, almost always by a considerable margin, and even TWTR does it for the other three-parameter combinations.

Similar results were obtained for the Mexican market (Figs. 8 and 9), although due to their lower resolution $\tau = 30$ and $N = 30$ are used. It is observed that the result of the local permutations is more ambiguous in this case, which can be interpreted as short-term memory lack.

5. Conclusions

This paper discusses the efficiency in high-frequency digital markets, quantified by the Hurst exponent measured by the R/S statistic. Results indicate that, in the period from March 7, 2018, to March 7, 2019, and for the 24 assets in the United States market and the 35 in the Mexican market studied here, the Efficient Market Hypothesis is clearly rejected: the presence of long-term memory, particularly of anti-persistence, is clear.

As noted before, the relevance of these results to the question of efficiency in automated digital markets lies like our data, coming from fully-automated (algorithmic) transactions. It is because of this that we can draw the main conclusion of this paper: automated digital markets do not meet the efficiency postulated by neoclassical theory. Thus, classical explanations of the inefficiencies of human markets, based on the psychological or emotional factors of human beings [10] or their limited rationality [11], must be discarded, since the algorithms that have ordered the transactions here studied do not suffer from these human limitations. Therefore, market inefficiency seems to be due to more fundamental factors of economic dynamics. This opens a new line of investigation in the search for the real sources of the lack of efficiency.

	$q=0.1, n=n_d$	$q=0.5, n=n_d$	$q=0.1, n=n_f$	$q=0.5, n=n_f$
GMEXICO	1e-45	2.84e-14	1e-45	2.84e-14
ELEKTRA	8.2e-05	0.00805	3.74e-06	2.71e-07
PENOLES	8.06e-41	2.84e-14	8.45e-24	1.31e-12
BSMX	4.51e-12	4.33e-10	1.04e-38	2.84e-14
MEXCHEM	0.00404	0.0676	6.95e-07	0.0178
GFNORTE	2.62e-09	0.116	2.62e-09	0.0676
AMX	9.88e-37	2.84e-14	1.04e-38	2.84e-14
ALFA	9.5e-30	1.31e-12	8.06e-41	2.84e-14
CEMEX	4.4e-33	4.67e-09	4.4e-33	4.67e-09
PINFRA	2.62e-09	2.94e-11	1.84e-08	2.94e-11
WALMEX	4.55e-13	3.29e-05	4.55e-13	3.29e-05
KIMBER	4.01e-21	3.29e-05	3.3e-25	3.94e-08
GAP	2.62e-09	0.000124	1.97e-17	1.31e-12
RA	0.032	0.116	6.95e-07	0.00805
ALSEA	8.45e-24	1.31e-12	9.88e-37	2.84e-14
CUERVO	1.15e-26	1.31e-12	3.53e-28	2.84e-14
IENOVA	0.00122	0.000412	3.74e-06	3.94e-08
AC	1.97e-17	1.31e-12	4.06e-43	2.84e-14
MEGA	8.2e-05	2.71e-07	4.55e-13	4.33e-10
VOLAR	2.77e-16	3.94e-08	4.1e-11	1.56e-06
GFINBUR	1.97e-17	2.71e-07	3.3e-25	4.33e-10
LIVEPOL	4.21e-14	2.94e-11	1.27e-18	2.94e-11
GMXT	1.27e-18	2.71e-07	1.27e-18	3.94e-08
NEMAK	4.21e-14	4.67e-09	2.77e-16	2.84e-14
LALA	1.97e-17	4.33e-10	4.21e-14	4.33e-10
BIMBO	1.94e-22	2.71e-07	4.51e-12	0.000412
ASUR	4.1e-11	1.56e-06	4.01e-21	4.33e-10
GRUMA	1.94e-22	2.71e-07	4.4e-33	2.84e-14
TLEVISA	3.3e-25	4.67e-09	3.3e-25	2.94e-11
KOF	1.94e-22	1.31e-12	3.57e-15	2.94e-11
OMA	1.97e-17	4.67e-09	4.21e-14	4.33e-10
GENTERA	0.159	0.186	0.671	0.186
ALPEK	3.53e-28	4.67e-09	3.53e-28	4.67e-09
FEMSA	7.5e-20	4.67e-09	1.27e-18	4.33e-10
GCARSO	1.84e-05	2.71e-07	3.74e-06	2.71e-07
Average MX	0.0056	0.0141	0.0192	0.00799

FIGURE 9. Values of $P^q(n)$ for the Mexican market for $\tau = 30$, $N = 30$ y $q = 0.1, 0.5$.

1. L. Bachelier, *Théorie de la Spéculation*, (Annales de l'Ecole Normale Supérieure, Ser 3, Gauthier-Villars, Paris, 1900).

2. R. N. Mantegna, H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, (Cambridge University Press, UK, 2000).

3. R. Mansilla, *Una breve introducción a la Econofísica*, (Equipo Sirius, S. A., Madrid, España, 2003).

4. H. K. Pharasi, K. Sharma, R. Chatterjee, A. Chakraborti, F. Leyvraz, T. H. Seligman, *Identifying long-term precursors of financial market crashes using correlation patterns*, New Journal of Physics **20** (2018) 103041, `https://doi.org/10.1088/1367-2630/aae7e0`.

5. A. Chakraborti, K. Sharma, H. K. Pharasi, K. Shuvo Bakar, S. Das, T. H. Seligman, *Emerging spectra characterization of catastrophic instabilities in complex systems*, New Journal of Physics **22** (2020) 063043, `https://doi.org/10.1088/1367-2630/ab90d4`.

6. K. Kanazawa, T. Sueshige, H. Takayasu, M. Takayasu, *Derivation of the Boltzmann Equation for Financial Brownian Motion: Direct Observation of the Collective Motion of High-Frequency Traders*, Physical Review Letters **120** (2018) 138301, `https://doi.org/10.1103/physrevlett.120.138301`.

7. E. Fama, *The Behavior of Stock Market Prices*, The Journal of Business, **38** (1965) 34, `https://doi.org/10.1086/294743`.

8. A. W. Lo, A. C. MacKinlay, *A Non-Random Walk Down Wall Street*, (Princeton University Press, Princeton, New Jersey, 1999), pp. 10.

9. J. L. McCauley, *Dynamics of Markets, The New Financial Economics*, second edition, (Cambridge University Press, New York, 2009), pp. 10-29.

10. J. Keynes, *The General Theory of Employment, Interest and Money*, (Palgrave Macmillan, London, 1936), pp. 161-162.

11. H. Simon, *A Behavioral Model of Rational Choice*, The Quar-

terly Journal of Economics **69** (1955) 99, https://doi.org/10.2307/1884852.

12. A. W. Lo, *The Adaptative Market Hypothesis*, The Journal of Portfolio Management 30th Anniversary **30** (2004) 15, https://doi.org/10.3905/jpm.2004.442611.
13. S. Leroy, *Risk aversion and the martingale property of stocks rerturns*, International Economic Review **14** (1973) 436, https://doi.org/10.2307/2525932.
14. R. Lucas, *Asset prices in an exchange economy*, Econometrika **46** (1978) 1429, https://doi.org/10.2307/1913837.
15. H. E. Hurst, *Long-Term Storage Capacity of Reservoirs*, Transactions of the American Society of Civil Engineers **116** (1951) 770, https://doi.org/10.1061/taceat.0006518.
16. B. B. Mandelbrot, J. R. Wallis, *Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence*, Water Resources Research **5** (1969) 967, https://doi.org/10.1029/wr005i005p00967.
17. G. Wang, G. Antar, P. Devynck, *The Hurst exponent and long-time correlation*, Physics of Plasmas **7** (2000) 1181, https://doi:10.1063/1.873927.
18. C.-F. Li, *Rescaled-range and power spectrum analyses on well-logging data*, Geophysical Journal International **153** (2003) 201, https://doi.org/10.1046/j.1365-246x.2003.01893.x.
19. M. A. Sánchez, J. E. Trinidad, J. García, M. Fernández, *The Effect of the Underlying Distribution in Hurst Exponent Estimation*, PLoS ONE **10** (2015) e0127824, https://doi:10.1371/journal.pone.0127824.
20. A. W. Lo, *Long-Term Memory in Stock Market Prices*, Econometrica **59** (1991) 1279, https://doi.org/10.2307/2938368.
21. W. Willinger, M. S. Taqqu, V. Teverovsky, *Stock market prices and long-range dependence*, Finance Stochast. **3** (1999) 1, https://doi.org/10.1007/s007800050049.
22. J. Barunik, L. Kristoufek, *On Hurst exponent estimation under heavy-tailed distributions*, Physica A: Statistical Mechanics and its Applications **389** (2010) 3844, https://doi.org/10.1016/j.physa.2010.05.025.
23. D. O. Cajueiro, B. M. Tabak, *The Hurst exponent over time: testing the assertion that emerging markets are becoming more effcient*, Physica A: Statistical Mechanics and its Applications **336** (2004) 521, https://doi.org/10.1016/j.physa.2003.12.031.
24. Ch. Eom, S. Choi, G. Oh, W. Jung *Hurst exponent and prediction based on weak-form efficient market hypothesis of stock markets*, Physica A: Statistical Mechanics and its Applications **387** (2008) 4630, https://doi.org/10.1016/j.physa.2008.03.035.
25. D. Grech, Z. Mazur, *Can one make any crash prediction in finance using the local Hurst exponent idea?*, Physica A: Statistical Mechanics and its Applications **336** (2004) 133, https://doi.org/10.1016/j.physa.2004.01.018.
26. E. Peters, *Fractal Market Analysis. Applying Chaos Theory to Investment and Economics*, (John Wiley and Sons, INC., New York, 1994) pp. 39-85.
27. M.A. Sánchez, J.E. Trinidad, J. García, *Some comments on Hurst exponent and the long memory processes on capital markets*, Physica A: Statistical Mechanics and its Applications **387** (2008) 5543, https://doi.org/10.1016/j.physa.2008.05.053.