THE EUROPEAN
PHYSICAL JOURNAL B

# Inefficiency in Latin-American market indices

L. Zunino[1,2,3,a], B.M. Tabak[4], D.G. Pérez[5], M. Garavaglia[1,3], and O.A. Rosso[6,7]

[1] Centro de Investigaciones Ópticas, C.C. 124 Correo Central, 1900 La Plata, Argentina
[2] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[3] Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, 1900 La Plata, Argentina
[4] Banco Central do Brasil, SBS Quadra 3, Bloco B, 9 andar, DF 70074-900, Brazil
[5] Instituto de Física, Pontificia Universidad Católica de Valparaíso (PUCV), 23-40025 Valparaíso, Chile
[6] Chaos & Biology Group, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina
[7] Centre for Bioinformatics, Biomarker Discovery and Information-Based Medicine, School of Electrical Engineering and Computer Science, The University of Newcastle, University Drive, Callaghan NSW 2308, Australia

**Abstract.** We explore the deviations from efficiency in the returns and volatility returns of Latin-American market indices. Two different approaches are considered. The dynamics of the Hurst exponent is obtained via a wavelet rolling sample approach, quantifying the degree of long memory exhibited by the stock market indices under analysis. On the other hand, the Tsallis $q$ entropic index is measured in order to take into account the deviations from the Gaussian hypothesis. Different dynamic rankings of inefficieny are obtained, each of them contemplates a different source of inefficiency. Comparing with the results obtained for a developed country (US), we confirm a similar degree of long-range dependence for our emerging markets. Moreover, we show that the inefficiency in the Latin-American countries comes principally from the non-Gaussian form of the probability distributions.

**PACS.** 05.45.Tp Time series analysis – 89.65.Gh Economics; econophysics, financial markets, business and management – 05.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion – 89.75.-k Complex systems

## 1 Introduction

According to the Efficient Market Hypothesis (EMH), at any given time, the price of an asset fully reflects all available information. It should follow a martingale process in which each price change is unmodified by its predecessor. Therefore, the existence of autocorrelation between distant observations violates the market efficiency because past prices can help to predict future prices. This question motivates the research on the subject, especially by portfolio managers and analysts who try to generate higher returns. There are many examples of empirical results for testing for long-range dependence, also known as long memory, in the literature. Bęben and Orłowski [1] have found that emerging markets have greater correlation than developed markets, suggesting more predictability. Then, emerging economies seem to be less efficient than developed ones. More recently, Di Matteo et al. [2,3] have also shown that deviations from efficiency are associated with the degree of development. Another "styl-ized fact" is that returns themselves have less evidence of long memory. However, strong evidence of long-range dependence is found for absolute returns[1] [4–7]. Cajueiro and Tabak have shown that long-range dependence for returns [8] and volatility returns [9] is time-varying and therefore the dynamics of these Hurst exponents, that quantify the degree of correlations, should be explored. These authors have also introduced a rank for market efficiency by considering the median Hurst exponent as a measure of it [10,11]. Other recent approaches to this issue seem to show that financial markets are increasing their efficiency over time [12,13]. These increasing efficiency can be mainly attributed to the variation of the effects of (a) speed of information, (b) capital flows, and (c) nonsynchronous trading. More recently, Cajueiro and Tabak [14] studied the dynamics of euro bilateral exchange

[a] e-mail: lucianoz@ciop.unlp.edu.ar

---

[1] Following the arguments mentioned in reference [4], one reason to study the behavior of the absolute returns is that the investors are influenced not only by the evolution of the sign of the investment, but also by the measure of the amplitude of its change.

rates by estimating their Tsallis $q$ entropic index. They provide statistical evidence that this parameter can also be used as a measure of efficiency. In the standard EMH model the returns follow an uncorrelated Gaussian process (white noise). This process guarantees the basic properties of an efficient market. Deviations from this model are possible violating either the independence or the Gaussian assumptions[2]. So, there are two different sources of inefficiency. The Hurst exponent, as a measure of long-range dependence, quantifies the former. It is widely known that the Tsallis $q$ parameter measures the deviations from the Gaussian hypothesis. Then, it takes into account the latter source of inefficiency. Summarizing, the Hurst and Tsallis parameters provide us two independent ranks for market efficiency. It should be remarked that both ranks are necessary for a properly characterization of the same stock market.

The aim of this study is to find out if long memory and deviations from Gaussianity exist in the returns and volatility returns of Latin-American market indices. We put special attention into the efficiency question. For this purpose we use the two approaches previously mentioned.

This paper is organized in the following way. The methodologies used to estimate the Hurst exponent and the Tsallis $q$ entropic index are discussed in Sections 2.1 and 2.2, respectively. In Section 3, the data used in this study are detailed. In Section 4 we discuss the results, and in Section 5, we present the conclusions. Finally, the appendix briefly reviews the main properties of two well-known stochastic process, namely, fractional Brownian motion (fBm) and fractional Gaussian noise (fGn). They are widely used to model long-range dependence.

## 2 Methodology

### 2.1 Estimation of Hurst exponent

The Wavelet Analysis is one of the most useful tools when dealing with data samples. Any signal can be decomposed by using a wavelet dyadic discrete family $\psi_{j,k} = \{2^{-j/2}\psi(2^{-j}t - k)\}$, with $j, k \in \mathbb{Z}$ (the set of integers) — an *orthonormal* basis for $L^2(\mathbb{R})$ consisting of finite-energy signals — of translations and scaling functions based on a function $\psi$: the mother wavelet [15,16]. In the following, given a continuous stochastic process $s(t)$ its associated *discrete signal* is assumed to be given by the sampled values $\mathcal{S} = \{s(n), n = 1, \ldots, M\}$. Its wavelet expansion has associated wavelet coefficients given by

$$C_j(k) = \langle \mathcal{S}, \psi_{j,k} \rangle, \qquad (1)$$

with $j = 1, \ldots, J$, and $J = \log_2 M$. The number of coefficients at each resolution level is $N_j = 2^{-j}M$. Note that this correlation gives information on the signal at scale $2^j$

and time $2^j k$. The set of wavelet coefficients at level $j$, $\{C_j(k)\}_k$, is also a stochastic process where $k$ represents the discrete time variable. It provides a direct estimation of local energies at different scales.

Following the arguments of Abry et al. [17], the wavelet based approach allows the threefold objective: detection, identification and measurements of long-range dependence. These features are due to the fact that the wavelet family $\psi_{j,k}$ exhibits scale invariance[3]. Moreover, the wavelet analysis is able to recognize quite different kinds of scaling. In particular, it is well-suited to fBm and fGn — for details about these stochastic processes see Appendix A. Properties of the wavelet coefficients of them can be gathered into a unified framework [17]:

1. $\{C_j(k)\}_k$ is a stationary process if $N \geq (\alpha - 1)/2$, with $N$ the number of vanishing moments associated to the mother wavelet[4], and $\alpha = 2H + 1$ for fBm or $\alpha = 2H - 1$ for fGn;
2. $\{C_j(k)\}_k$ has only short term residual correlations, i.e., it does not have long-range dependence, on condition that $N \geqslant \alpha/2$;
3. the underlying scaling behaviour of the data are reproduced by the variance of the $\{C_j(k)\}_k$ process, within a given range of resolution levels $j_1 \leqslant j \leqslant j_2$:

$$\mathsf{E}\left[C_j^2(k)\right] \propto 2^{\alpha j}, \qquad (2)$$

where $\mathsf{E}$ stands for the average using some, at first, unknown probability distribution associated to the coefficients. It should be stressed that these results are valid for any self-similar process with stationary increments ($\alpha = 2H + 1$) and long-range dependent processes ($\alpha = 2H - 1$). fBm and fGn are the archetypes of these two different kinds of scaling.

These properties are widely used for estimating $H$ or the related spectral exponent $\alpha = 2H \pm 1$ [18–20]. Note that $1 < \alpha < 3$ for fBm and $-1 < \alpha < 1$ for fGn. The variance of the wavelet coefficients is estimated as usual:

$$\mu_j = \frac{1}{N_j} \sum_{k=1}^{N_j} |C_j(k)|^2, \qquad (3)$$

This is a non-parametric, unbiased estimator of the variance of the process $\{C_j(k)\}_k$. Then, plotting the $\log_2(\mu_j)$ versus $j$ and fitting a minimum square line, the estimated Hurst exponent, $\widehat{H}$, is obtained from the slope. Basically the estimation problem turns into a linear regression slope estimation and a fitting range of resolutions levels is required. Using these issues, Abry and Veitch introduced a wavelet-based analysis tool named the *Logscale Diagram*, which is defined as the plot of $y_j = \log_2 \mu_j$ against $j$ together with confidence intervals assigned to each $y_j$. The associated Matlab routine can be found in reference [21].

---

[2] The occurrence of extreme events such as political disturbances, terrorist attacks, bankruptcies of leading companies or wars imply the existence of power-law behaviors in the probability distributions of returns ("fat tails").

[3] As it is widely known the long-range dependence is intimately related to scale invariance and self-similarity [18].

[4] The mother wavelet $\psi$ has $N$ vanishing moments if $\int t^k \psi(t) dt \equiv 0$ for $k = 0, 1, 2, \ldots, N - 1$. It is shown that any polynomial trend of the signal can be rigorously eliminated by varying this number of vanishing moments [18].

We calculate the time-dependent Hurst exponent by using the Logscale Diagram and a rolling sample approach previously introduced [8,22,23]. The Hurst exponent is estimated on the ensemble of points obtained from the intersection of the signal and a sliding window of size $N_s$. Then, the time window is rolled $\delta_s$ points forward eliminating the first $\delta_s$ observations and including the next ones, and the Hurst exponent is re-estimated. These procedure is repeated until the end of the time serie.

*Multifractional Brownian motion* (mBm) [24] was formalized as a class of processes where the exponent $H$ is no longer constant, but a continuous function of the time $t$ ($H \to H(t)$)[5]. Artificial mBm were analyzed in order to test the quality of our estimator. In Figure 1 one can compare the theoretical and experimental results for a mBm with $H$ changing linearly from 0.1 to 0.9 with $t$. The Matlab code introduced by Coeurjolly [26] was implemented to simulate the mBm. $N_s = 1024$ and $\delta_s = 16$ were chosen for the rolling sample approach. Hereafter, we used the Daubechies ($N = 3$) as mother wavelet. We have also implemented the special prefiltering required in the analysis of intrinsically discrete time series[6], included into the Abry and Veitch Matlab routine (MRA initialization). Resolution levels from $j = 2$ to $j = 7$ were used for the linear regression in each time window of this mBm simulation. It should be remarked that the Hurst exponent is always underestimated. It was shown that the wavelet estimator is asymptotically unbiased and Gaussian and the variance of the estimation obtained from a sliding window of size $N_s$ is of order $1/N_s$ [17,18,28]. As it is an asymptotically efficient estimator, we assume that small window lengths are a potential source of bias in the estimates. We need to estimate the Hurst exponent over windows that are short enough so that they can be taken as constant but long enough so that their statistical estimates are robust. Furthermore, the finite segmentation lengths introduce noise into the estimates that should be removed [29].

## 2.2 The Tsallis q entropic index

The maximization of the non-extensive entropy subject to appropriate constraints, for a fixed entropic index $q$, yields the Tsallis distribution [30,31]:

$$p_q(x) = \frac{\left[1 + \beta(q-1)(x-\overline{x})^2\right]^{\frac{-1}{q-1}}}{Z_q}, \qquad (4)$$

where

$$Z_q = \left[\frac{\pi}{\beta(q-1)}\right]^{1/2} \frac{\Gamma((3-q)/2(q-1))}{\Gamma(1/(q-1))} \qquad (5)$$



**Fig. 1.** Top: synthetically generated mBm signal with $H$ changing linearly from 0.1 to 0.9 with $t$. Bottom: theoretical (dashed curve) and measured (continuous curve) Hurst exponent for this simulated mBm.

is the appropriate normalization factor and $\beta$ is the so-called Lagrange parameter associated with the constraints. In the limit of $q \to 1$, the Gaussian distribution is recovered. From equation (4) the ordinary variance can be easily obtained

$$\sigma^2 = \begin{cases} \frac{1}{\beta(5-3q)}, & q < 5/3 \\ \infty, & q \geq 5/3. \end{cases}$$

It is also easy to show that the kurtosis coefficient $K$ depends only on $q$ and is given by

$$K = \frac{3(5-3q)}{7-5q}. \qquad (6)$$

This equation is particularly useful to calculate the entropic index from real data. Observe that with $q = 1$ we obtain $K = 3$, the expected value for the kurtosis of a Gaussian process.

## 3 Data

We have collected daily closing prices for Argentina, Brazil, Chile, Colombia, Mexico, Peru, Venezuela and the

---

[5] Moreover, this generalized fractional Brownian motion was suggested as an alternative theoretical framework to explain the multifractal behavior observed in financial time series [22,25].

[6] The discrete wavelet transform (DWT) is defined only for continuous-time processes. So, the discrete data are associated to a continuous process, and then the usual wavelet method is applied [27].
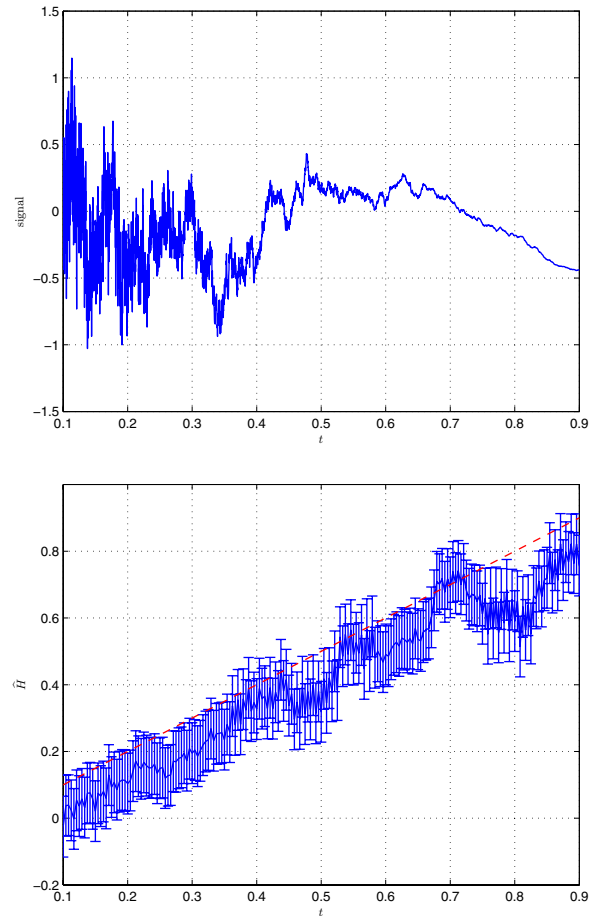
US. This last index is included for comparison purposes. The first three markets together with Mexico and the US correspond to large stock markets. On the other hand, Colombia, Peru and Venezuela can be considered smaller stock markets. The data samples cover the period from January 1995 to February 2007 with 3169 observations for all the countries except Venezuela with 3168 observations and the US with 3141 observations. All series were gathered from the Morgan Stanley Capital Index (MSCI). This style index was employed in order to disentangle the effects of autocorrelation and infrequent trading, which are specially severe in emerging market. They are value-weighted indices of a large sample of companies in each country and the indices do not include foreign companies and are computed consistently across markets, thereby allowing for a close comparison across countries.

Let $x(t)$ be the price of a stock on a time $t$, the returns, $r_t$, are calculated as its logarithmic difference, $r_t = \ln(x(t+1)/x(t))$.

Volatility is of very interest to traders because it quantifies the risk and is directly related to the amount of information arriving in the financial market at a given time. The volatility of stock price changes is a measure of how much the market is liable to fluctuate. It is a "subsidiary process" estimated from the original one. The series of absolute returns $|r_t|$ and squared returns $r_t^2$ are usually used as proxies for volatility returns. It should be noted that there are evidence of stronger long-range dependence for absolute returns than for squared returns [4,5,9]. We follow Ding, Granger and Engle [32] who suggest measuring volatility directly from absolute returns. Furthermore, Davidian and Carroll [33] show absolute returns volatility specification is more robust against asymmetry and nonnormality.

## 4 Empirical results

We have estimated the Hurst exponent through the Logscale Diagram. Figure 2 illustrates this wavelet-based analysis tool in the case of Argentina[7]. These global values are reported in Table 1. It should be stressed that the fitting range $2 \leq j \leq 8$ was always selected. Note that $\widehat{\alpha}_{\min} = -0.114$ and $\widehat{\alpha}_{\max} = 0.206$ for returns, and $\widehat{\alpha}_{\min} = 0.254$ and $\widehat{\alpha}_{\max} = 0.558$ for volatility returns. Thus, we conclude that both processes should be modeled with a fGn within the Fractal Market Hypothesis. Obviously, its associated cumulative process $y(t) = \ln(x(t))$ will be a fBm.

Since the long-range dependence of financial time series seems to evolve over time [8,9], we measure this exponent statically and also dynamically via the wavelet rolling sample approach described in Section 2.1. In order to estimate a time-varying Hurst exponent two opposite requirements must be satisfied [22]: to determine a *local* exponent $H$ at a particular time $t$, a small window around

---

[7] The plots for the daily returns and volatility returns of the other countries are available upon request from the corresponding author.
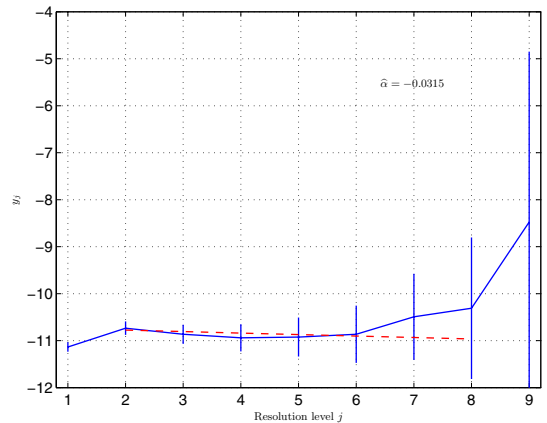


**Fig. 2.** Logscale diagram for the daily returns of Argentina and the regression line for $2 \leq j \leq 8$ (dashed curve). The vertical bar at each resolution level gives 95% confidence intervals for the $y_j$.

**Table 1.** Global Hurst exponent for the daily returns and volatility returns, considering the time series as a whole.

| Country | Returns | Volatility |
|---|---|---|
| Argentina | $0.484 \pm 0.042$ | $0.689 \pm 0.042$ |
| Brazil | $0.505 \pm 0.042$ | $0.784 \pm 0.042$ |
| Chile | $0.565 \pm 0.042$ | $0.627 \pm 0.042$ |
| Colombia | $0.603 \pm 0.042$ | $0.779 \pm 0.042$ |
| Mexico | $0.488 \pm 0.042$ | $0.652 \pm 0.042$ |
| Peru | $0.524 \pm 0.042$ | $0.628 \pm 0.042$ |
| Venezuela | $0.509 \pm 0.042$ | $0.679 \pm 0.042$ |
| US | $0.443 \pm 0.042$ | $0.641 \pm 0.042$ |

this time has to be considered but still sufficiently large to contain enough points for a significant statistics. Following the line of reasoning of Cajueiro and Tabak [8], we choose a sliding window of size $N_s = 1024$ (about 4 years) because it reflects political cycles in most countries and it is sufficiently large to give precise estimates. The time window was rolled $\delta_s = 16$ points. For smaller steps we find redundant information. Figures 3 and 4 present the time-varying Hurst exponents for the returns and volatility returns, respectively, of Latin-American and the US market indices. Resolution levels from $j = 2$ to $j = 7$ were used for the linear regression in each time window of the wavelet rolling sample approach. Also the global Hurst exponent, considering the time series as a whole, are included in these plots.

In order to rank these countries we build the following measure of inefficiency:

$$Ineff = \frac{\left|\widehat{H} - 0.5\right|}{\sigma_{\widehat{H}}}. \qquad (7)$$

This measure gives the distance, in terms of standard deviations, of estimated Hurst exponents from the critical value of 0.5. Thus, it can be used to asses whether returns or volatility returns possess long-range dependence. If this score is larger than 2 then we can confirm the hypothesis of long-range dependence with a 95% of confidence.
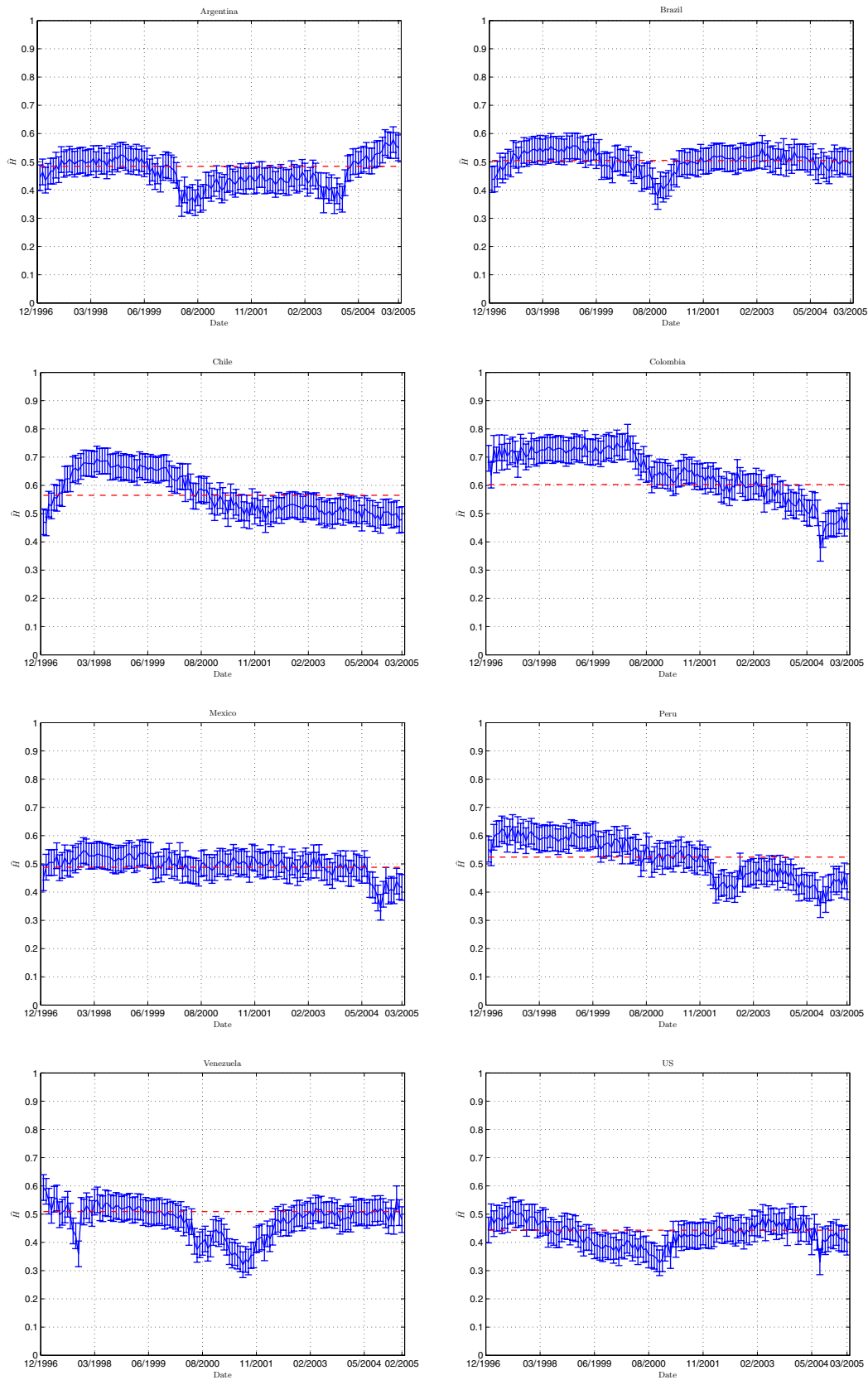
**Fig. 3.** Plots of estimated Hurst exponent by using the wavelet rolling sample approach with a sliding window of size $N_s = 1024$ and step $\delta_s = 16$ for the daily returns of Latin-American and the US market indices. Dashed lines correspond to the global Hurst exponent considering the time series as a whole.
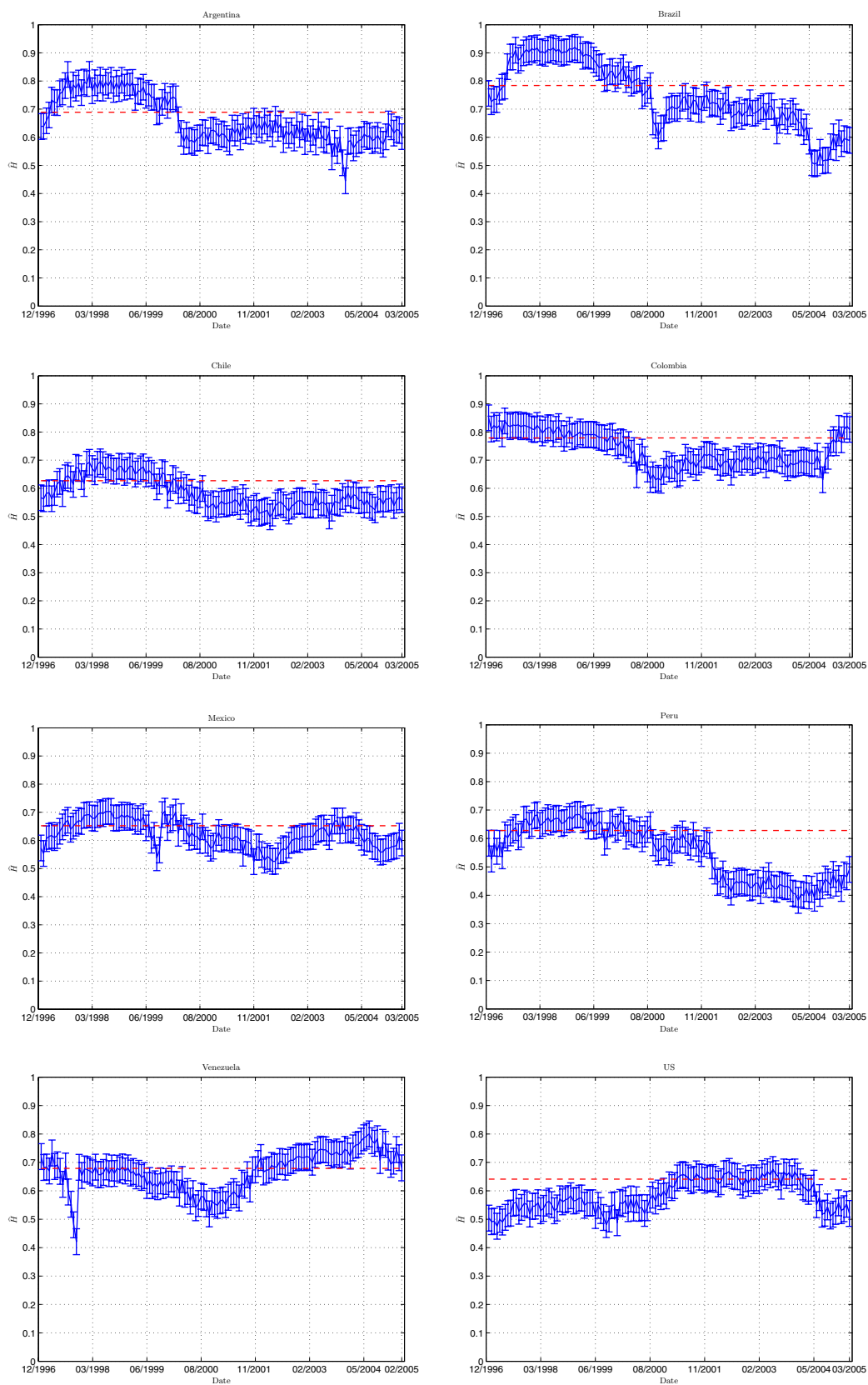
**Fig. 4.** Plots of estimated Hurst exponent by using the wavelet rolling sample approach with a sliding window of size $N_s = 1024$ and step $\delta_s = 16$ for the daily volatility returns of Latin-American and the US market indices. Dashed lines correspond to the global Hurst exponent considering the time series as a whole.

**Table 2.** Inefficiency ranking for the daily returns considering the relative efficiency approach of Lim [34] (long-range dependence).

| Country | Total number of rolling windows | Total number of significant windows | Percentage of significant windows (%) |
|---|---|---|---|
| Brazil | 135 | 3 | 2.22 |
| Mexico | 135 | 6 | 4.44 |
| Argentina | 135 | 19 | 14.07 |
| Venezuela | 134 | 21 | 15.67 |
| Peru | 135 | 33 | 24.44 |
| US | 133 | 37 | 27.82 |
| Chile | 135 | 47 | 34.81 |
| Colombia | 135 | 102 | 75.56 |

**Table 3.** Inefficiency ranking for the daily volatility returns considering the relative efficiency approach of Lim [34] (long-range dependence).

| Country | Total number of rolling windows | Total number of significant windows | Percentage of significant windows (%) |
|---|---|---|---|
| Chile | 135 | 47 | 34.81 |
| US | 133 | 56 | 42.11 |
| Peru | 135 | 69 | 51.11 |
| Mexico | 135 | 101 | 74.81 |
| Venezuela | 134 | 112 | 83.58 |
| Argentina | 135 | 113 | 83.70 |
| Brazil | 135 | 123 | 91.11 |
| Colombia | 135 | 135 | 100.00 |

Figures 5 and 6 show the time-varying inefficiency measure for the returns and volatility returns, respectively, of Latin-American and the US market indices by considering the Hurst exponent sequences derived previously from the wavelet rolling sample approach. Following the line of reasoning of Lim [34], a comparative analysis is achieved by comparing the total time windows these markets exhibit significant long-range dependence. A window is considered as significant if its associated inefficiency measure is greater than or equal to 2, the threshold level chosen in our case. So, the percentage of significant windows could be a meaningful indicator for assessing the relative inefficiency of stock markets. In Tables 2 and 3 we rank the returns and volatility returns of the Latin-American and the US market indices by considering this relative inefficiency approach.

If the returns and volatility returns are normally distributed the Tsallis $q$ entropic index would be close to one. The closer to one the more efficient is the time serie. In Table 4 we rank the returns and volatility returns by using the Tsallis $q$ entropic index. We use equation (6) to estimate this parameter. The entropic index can be easily derived:

$$q = \frac{15 - 7K}{9 - 5K}. \qquad (8)$$

**Table 4.** Inefficiency ranking for the daily returns and volatility returns according to the Tsallis $q$ entropic index (deviations from the Gaussian hypothesis).

| Country | Returns | Volatility |
|---|---|---|
| US | 1.3027 | 1.3487 |
| Chile | 1.3071 | 1.3589 |
| Peru | 1.3298 | 1.3604 |
| Brazil | 1.3368 | 1.3695 |
| Colombia | 1.3622 | 1.3804 |
| Mexico | 1.3692 | 1.3861 |
| Argentina | 1.3793 | 1.3906 |
| Venezuela | 1.3964 | 1.3981 |

As you can observe, the returns and its associated proxy for volatility (absolute returns) provide the same rank. We also conclude that volatility returns have stronger deviations from the Gaussian hypothesis ($q$ farther from one) than the associated returns. This should be expected as volatility is bounded from below at zero.

To the best of our knowledge this is the first paper that employs a ranking methodology that does not depend on whether we are focusing on returns or volatility (moment independent). Furthermore, it provides reasonable results. Therefore, it is worth mentioning that the Tsallis $q$ entropic index seems to be an interesting method that can be used to characterize inefficiency.

## 5 Conclusions

We provide empirical evidence of long-range dependence in the daily returns and volatility returns of Latin-American market indices. Long memory is categorically showed in the volatility measures, while there is little evidence of it in the returns. Portfolio and risk managers are interested in finding patterns of predictability. Since volatility has strong long-range dependence, it should be predicted more easily and better than returns. These results also imply that commonly used models for volatility such as GARCH models are flawed and should be replaced by models that incorporate long memory.

The global results, considering the time series as a whole, suggest that there is stronger long-range dependence in returns for emerging markets than for developed ones. What is more, the US return has a clear antipersistent behaviour (see Tab. 1). However, on average long-range dependence in global volatility returns seems to be similar in mature liquid markets than in less developed ones. This result is in accord with previous findings [2,35].

In some countries the Hurst exponent changes substantially over time. This suggests that such countries are facing structural changes in the dynamics of their stock markets, which could be due to changes in their regulation. Thus, we find that the efficiency of stock markets should be examined as a dynamic characteristic. It is evolving in time, with periods of inefficiency alternate with those of efficiency. The rolling sample approach is an excellent approach to this complex issue. We also conclude that different rankings can be derived, each of them taking into account different sources of inefficiency.
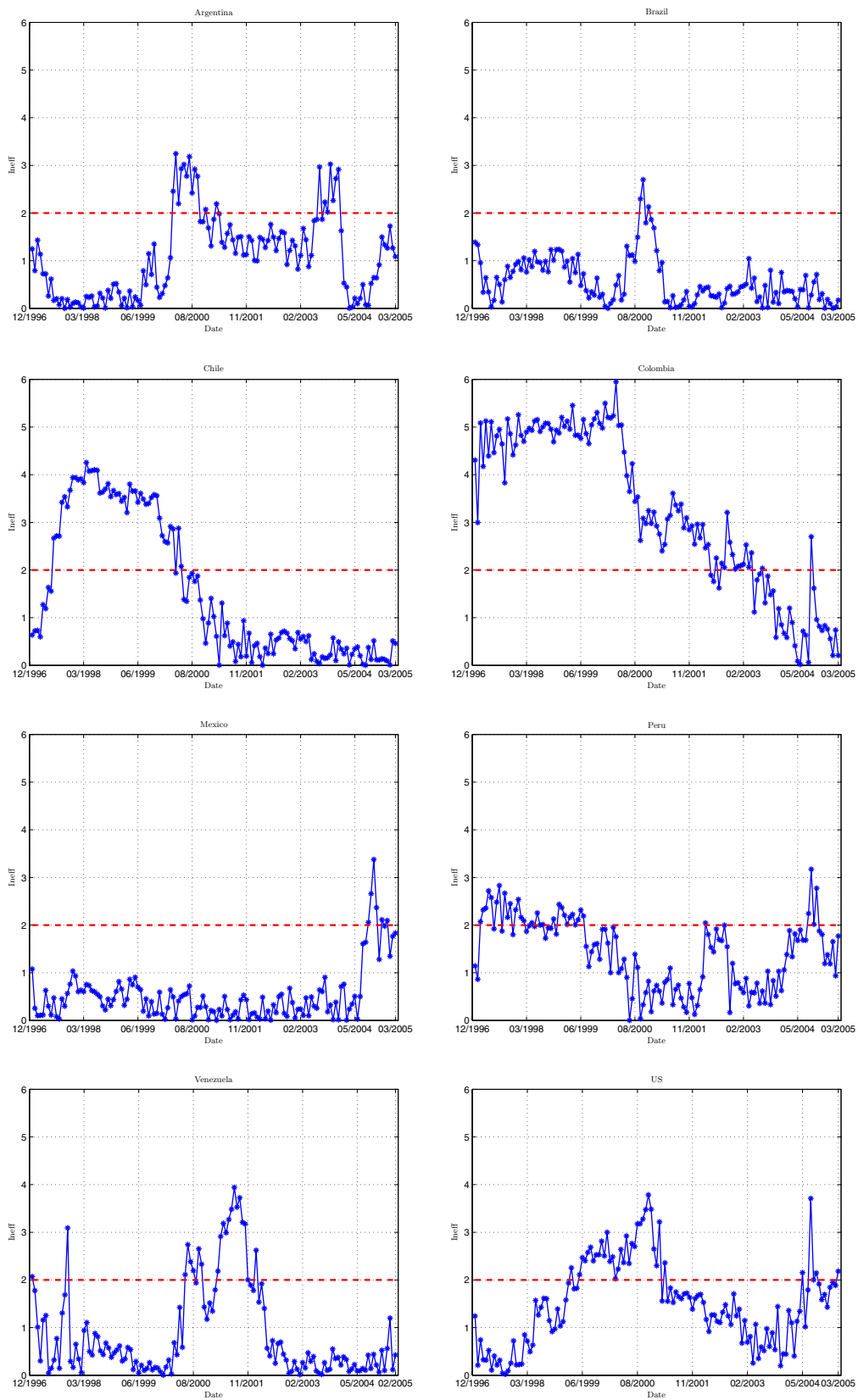
**Fig. 5.** Plots of the inefficiency measure (Eq. (7)) by using the wavelet rolling sample approach with a sliding window of size $N_s = 1024$ and step $\delta_s = 16$ for the daily returns of Latin-American and the US market indices. The horizontal lines correspond to the threshold level.
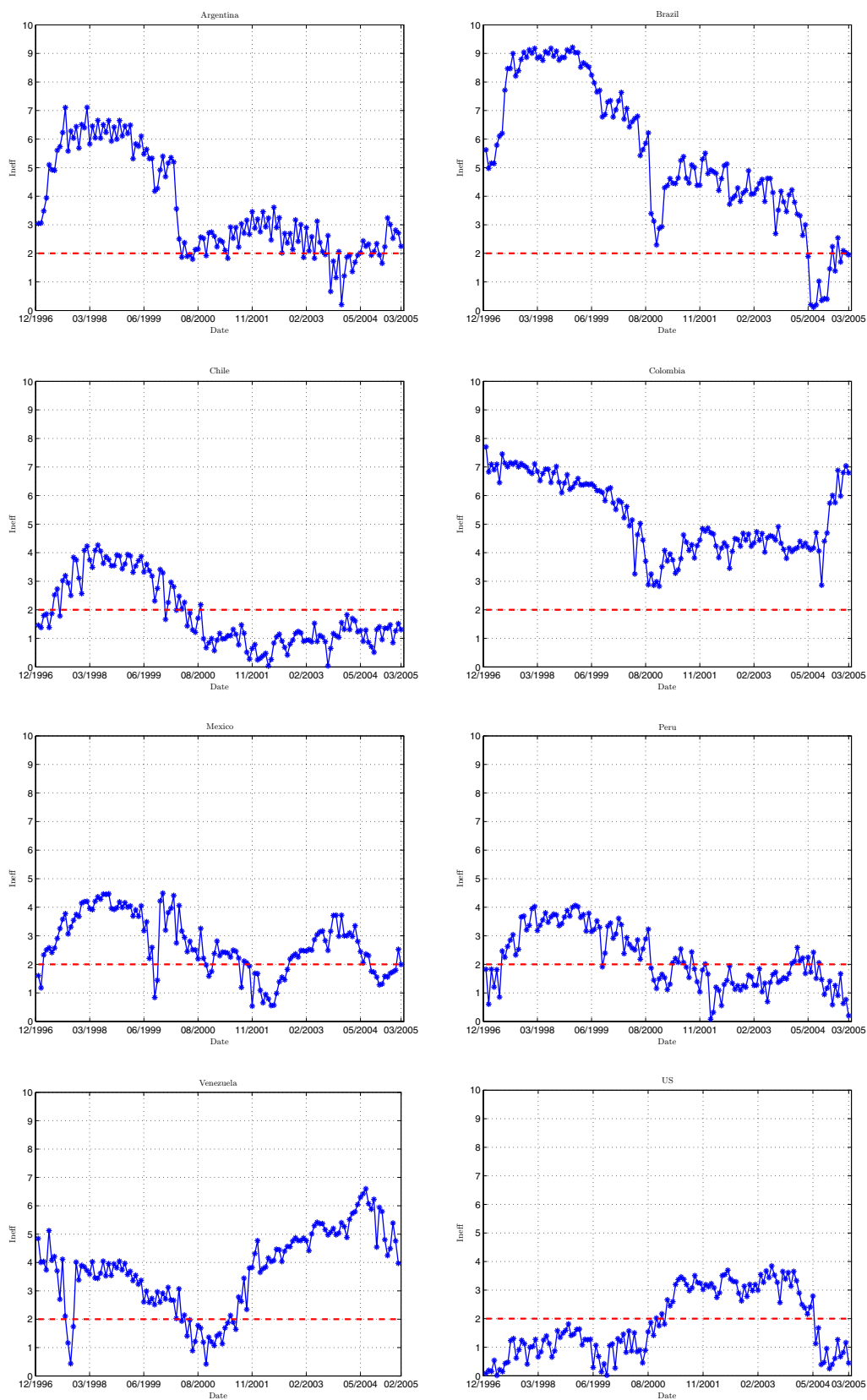
**Fig. 6.** Plots of the inefficiency measure (Eq. (7)) by using the wavelet rolling sample approach with a sliding window of size $N_s = 1024$ and step $\delta_s = 16$ for the daily volatility returns of Latin-American and the US market indices. The horizontal lines correspond to the threshold level.

Therefore, they do not provide the same rank of inefficiency for Latin-American market indices. For example, Chile is one of the worst ranked for the daily returns considering the relative efficiency approach (see Tab. 2) but at the same time is one of the best ranked countries according to the Tsallis $q$ entropic index (see Tab. 4). So, we can conclude that the Chilean market index has stronger long-range dependence and weaker deviations from the Gaussian hypothesis.

Since the dynamics inefficiency rankings considered (see Tabs. 2, 3 and 4) we conclude that all the Latin-American countries under study experiment important deviations from the Gaussian hypothesis, this is a common source of inefficiency.

It was shown that magnitude series contain information regarding the nonlinear properties of the original time series [36]. Moreover, long-range dependence in the volatility is a proof for nonlinearity in the associated time serie [37]. Then, our results confirm nonlinear dependence (non-zero bicorrelation). Actually, this is another stylized fact that can be considered as a source of inefficiency, and a new rank will emerge from it [34]. Nonlinearity was previously detected in the rate of returns series for seven Latin-American stock market indices [38]. Its influence should be quantified in future works. Further study should also focus on the economic determinants of these different sources of inefficiency.

The intention is to repeat the same analysis for high-frequency returns and volatility time series [28,39–41] in order to understand the fine structure of the stock market. In this kind of studies we will have more observations. Thus, we expect more accuracy in the results obtained via our asymptotically efficient wavelet rolling sample approach. We will also remove the noise due to finite segmentation lengths through the Papanicolaou and Sølna approach [29].

## Appendix A: Fractal Market Hypothesis

Mandelbrot and van Ness [42] introduced the fractional Brownian motion (fBm) as a generalization of the ordinary Brownian motion by considering correlations between the increments of the process. Thus, this fractal stochastic model contemplates the presence of long-range dependence, the first of the two sources of inefficiency mentioned at the end of Section 1. The fBm and its corresponding generalized derivative process, the fractional Gaussian noise (fGn), are the benchmark of the so-called Fractal Market Hypothesis (FMH).

### A.1 Fractional Brownian motion

This is the only one family of processes which is Gaussian, self-similar, and endowed with stationary increments. The normalized family of these Gaussian processes, $\{B^H(t), t > 0\}$, is the one with $B^H(0) = 0$ almost surely (with probability 1), $\mathbb{E}\left[B^H(t)\right] = 0$ (zero mean), and covariance

$$\mathbb{E}\left[B^H(t)B^H(s)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t - s|^{2H}\right), \quad \text{(A.1)}$$

for $s, t \in \mathbb{R}$. Here $\mathbb{E}[\cdot]$ refers to the average with Gaussian probability density. The power exponent $H$, commonly known as Hurst exponent or Hurst parameter, has a bounded range between 0 and 1. These processes exhibit memory, as can be observed from equation (A.1), for any Hurst exponent but $H = 1/2$, for which one recovers the classical Brownian motion. In this case successive Brownian motion increments are as likely to have the same sign as the opposite, and thus there is no correlation among them. Precisely, this Hurst exponent defines two distinct regions in the interval $(0, 1)$. When $H > 1/2$, consecutive increments tend to have the same sign so that these processes are *persistent*. For $H < 1/2$, on the other hand, consecutive increments are more likely to have opposite signs, and it is said that these processes are *anti-persistent*. Fractional Brownian motions are continuous but non-differentiable processes (in the classical sense). As a nonstationary process, fractional Brownian motion does not possess a spectrum defined in the usual sense; however, it is possible to define a *generalized power spectrum* of the form [43]:

$$\Phi_{B^H}(f) \propto \frac{1}{|f|^\alpha}, \quad \text{(A.2)}$$

with $\alpha = 2H + 1$ and $1 < \alpha < 3$. Remember that this equation does not represent a valid power spectrum in the theory of stationary processes since it yields a non-integrable function (in the classical sense).

### A.2 Fractional Gaussian noise

We denoted by $\{W^H(t), t > 0\}$ the process derived from the increment of fractional Brownian motion, namely

$$W^H(t) = B^H(t) - B^H(t + 1). \quad \text{(A.3)}$$

We face a stationary Gaussian process with mean zero and covariance given by

$$\rho(k) = \mathbb{E}\left[W^H(t)W^H(t + k)\right]$$
$$= \frac{1}{2}\left[(k + 1)^{2H} - 2k^{2H} + |k - 1|^{2H}\right], \; k > 0. \quad \text{(A.4)}$$

The last expression has the following asymptotic behavior as $k \to \infty$ [44]

$$\frac{\rho(k)}{H(2H-1)k^{2H-2}} \to 1. \qquad (A.5)$$

Therefore, when $H > 1/2$ this correlation decays to zero so slowly that the sum $\sum_{k=-\infty}^{k=\infty} \rho(k) = \infty$ diverges [44]; this sub-family of processes has long-memory. On the other hand, for $H < 1/2$ the correlations of the increments are summable [44], and this sub-family exhibits short-memory. Equation (A.5) also allows one to corroborate the assertions about the persistent or anti-persistent behavior mentioned above. Note that for $H = 1/2$ all correlations at non-zero lags vanish and $\{W^{1/2}(t), t > 0\}$ is *white noise*. Naturally, time series of cumulative Gaussian white noise constitute samples of classical Brownian motions. The power spectrum associated to fractional Gaussian noise reads

$$\Phi_{W^H}(f) \propto \frac{1}{|f|^{\alpha}}, \qquad (A.6)$$

with $\alpha = 2H - 1$ and $-1 < \alpha < 1$.

## References

1. M. Bęben, A. Orłowski, Eur. Phys. J. B **20**, 527 (2001)
2. T. Di Matteo, T. Aste, M.M. Dacorogna, Physica A **324**, 183 (2003)
3. T. Di Matteo, T. Aste, M.M. Dacorogna, J. Bank. Fin. **29**, 827 (2005)
4. P. Grau-Carles, Physica A **287**, 396 (2000)
5. T. Lux, Appl. Econom. Lett. **3**, 701 (1996)
6. B.K. Ray, R.S. Tsay, J. Bus. Econ. Stat. **18**, 254 (2000)
7. P. Sibbertsen, Empirical Econ. **29**, 477 (2004)
8. D.O. Cajueiro, B.M. Tabak, Physica A **336**, 521 (2004)
9. D.O. Cajueiro, B.M. Tabak, Physica A **346**, 577 (2005)
10. D.O. Cajueiro, B.M. Tabak, Solitons & Fractals **22**, 349 (2004)
11. D.O. Cajueiro, B.M. Tabak, Chaos, Solitons & Fractals **23**, 671 (2005)
12. B. Tóth, J. Kertész, Physica A **360**, 505 (2006)
13. H.F. Coronel-Brizio, A.R. Hernández-Montoya, R. Huerta-Quintanilla, M. Rodríguez-Achach, Physica A **380**, 391 (2007)
14. B.M. Tabak, D.O. Cajueiro, Physica A **367**, 319 (2006)
15. I. Daubechies, Ten lectures on wavelets, SIAM, Philadelphia, 1992
16. S. Mallat, *A wavelet tour of signal processing*, 2nd edn. (Academic Press, 1999)
17. P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch, Wavelets for the analysis, estimation, and synthesis of scaling data, in: *Self-similar Network Traffic and Performance Evaluation*, edited by K. Park, W. Willinger (Wiley, 2000), pp. 39–87
18. P. Abry, D. Veitch, IEEE Trans. Inform. Theory **44**, 2 (1998)
19. L. Zunino, D.G. Pérez, M. Garavaglia, O.A. Rosso, Fractals **12**, 223 (2004)
20. L. Zunino, D.G. Pérez, M. Garavaglia, O.A. Rosso, Physica A **364**, 79 (2006)
21. http://www.cubinlab.ee.mu.oz.au/~darryl/secondorder_code.html
22. R.L. Costa, G.L. Vasconcelos, Physica A **329**, 231 (2003)
23. A. Carbone, G. Castelli, H.E. Stanley, Physica A **344**, 267 (2004)
24. R.F. Peltier, J. Lévy-Véhel, Multifractional Brownian motion: definition and preliminary results, Research Report RR-2645, INRIA (1995)
25. S.V. Muniandy, S.C. Lim, R. Murugan, Physica A **301**, 407 (2001)
26. J.-F. Coeurjolly, *Statistical inference for fractional and multifractional Brownian motions*, Ph.D. thesis, Laboratoire de Modélisation et Calcul – Institut d'Informatique el Mathématiques Appliquées de Grenoble, 2000
27. D. Veitch, M.S. Taqqu, P. Abry, Sign. Process. **80**, 1971 (2000)
28. E. Bayraktar, H. Vincent Poor, K. Ronnie Sircar, Int. J. Theor. Appl. Finance **7**, 615 (2004)
29. G. Papanicolaou, K. Sølna, *Wavelet based estimation of local Kolmogorov turbulence, Long-Range Dependence: Theory and Applications*, edited by P. Doukhan, G. Oppenheim, M.S. Taqqu (Birkhäuser, 2003), pp. 473–505
30. C. Tsallis, S.V.F. Levy, A.M.C. Souza, R. Maynard, Phys. Rev. Lett. **75**, 3589 (1995)
31. C. Tsallis, D.J. Bukman, Phys. Rev. E **54**, R2197 (1996)
32. Z. Ding, C. Granger, R. Engle, J. Empir. Fin. **1**, 83 (1993)
33. M. Davidian, R.J. Carroll, J. Am. Stat. Association **82**, 1079 (1987)
34. K.P. Lim, Physica A **376**, 445 (2007)
35. D.O. Cajueiro, B.M. Tabak, Chaos, Solitons & Fractals, in press
36. Y. Ashkenazy, P.C. Ivanov, S. Havlin, C.-K. Peng, A.L. Goldberger, H.E. Stanley, Phys. Rev. Lett. **86**, 1900 (2001)
37. T. Kalisky, Y. Ashkenazy, S. Havlin, Phys. Rev. E **72**, 011913 (2005)
38. C.A. Bonilla, R. Romero-Meza, M.J. Hinich, Appl. Econ. Lett. **13**, 195 (2006)
39. M. Raberto, E. Scalas, G. Cuniberti, M. Riani, Physica A **269**, 148 (1999)
40. C. Morana, Appl. Finan. Econ. Lett. **2**, 361 (2006)
41. R.T. Baillie, Y.-W. Han, R.J. Myers, J. Song, J. Futures Mark. **27**, 643 (2007)
42. B.B. Mandelbrot, J.W. Van Ness, SIAM Rev. **4**, 422 (1968)
43. P. Flandrin, IEEE Trans. Inform. Theory **IT-35**, 197 (1989)
44. J. Beran, Statistics for long-memory processes, in: *Monographs on Statistics and Applied Probability* (Chapman & Hall, 1994), Vol. 61