Available online at www.sciencedirect.com
SCIENCE DIRECT
Physica A 355 (2005) 461–474
PHYSICA A
www.elsevier.com/locate/physa

Long correlations and truncated Levy walks applied to the study Latin-American market indices

Sebastian Jaroszewicz[a], M. Cristina Mariani[b,*], Marta Ferraro[a,b,c]

[a] *Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria Pab. I (1428), Buenos Aires, Argentina*
[b] *Department of Mathematical Sciences, New Mexico State University, Las Cruces, New Mexico 88003-8001, USA*
[c] *CONICET*

Received 7 March 2003
Available online 11 May 2005

Abstract

This work is devoted to the study of long correlations and other statistical properties of Latin-American market indices. We concluded that the behavior of the return is compatible with a slow convergence to a Gaussian distribution. We also detected long-range correlations in the absolute value of the return analyzing the effects of working with short data series. This fact has relevant consequences in the volatility dynamics.

Keywords: Levy flight; Econophysics; Detrended fluctuation analysis; Stock market prices; Latin-American indices

*Corresponding author.
E-mail addresses: sebasj1@df.uba.ar (S. Jaroszewicz), mmariani@nmsu.edu (M.C. Mariani), marta@df.uba.ar (M. Ferraro).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2005.04.003

1. Introduction

In recent years, there has been a growing literature in financial economics that analyzes the major Stock indices in developed countries [1–5].

One of the main problems is the analysis of the existence of long-term or short-term correlations in the behavior of financial markets.

The statistical properties of the temporal series got analyzing the evolution of the different markets have been of a great importance in the study of financial markets.

The empirical characterization of stochastic processes usually requires the study of temporal correlations and the determination of asymptotic probability density distributions (pdf). The first model that describes the evolution of option prices is the Brownian motion. This model assumes that the increment of the logarithm prices follows a diffusive process with Gaussian distribution [6]. However, the empirical study of temporal series of some of the most important indices shows that in short time intervals the associated pdf have larger kurtosis than a Gaussian distribution [5]. The first step in order to explain this behavior was done in 1963 by Mandelbrot [7]. He developed a model for the evolution of the cotton price by a stable stochastic non-Gaussian Levy process [8]. However, these distributions are not appropriate for working in long-range correlation scales. These problems can be avoided considering that the temporal evolution of financial markets is described by a truncated Levy flight (TLF) [9].

The detrended fluctuation analysis (DFA) method is an important technique that allows detecting the presence of long-range correlations in non-stationary temporal series. This method was developed by Peng, and has been applied to the study of the DNA, heart rate dynamics, solid state Physics, and economic series.

Most of the studies mentioned before have been done with indices of developed markets that have a great volume of transactions. This work is devoted to the study of the principal market indices in Latin America.

Our main interest is detecting long-range correlations in emerging markets economic indices. The presentation is described as follows: In the first section, we give a short introduction to the Levy distributions. In the second section, we present the indices that will be analyzed and we give a relation between them and the TLF. In the third section, we give a brief description of the fundamentals of the technique known as DFA and we applied this technique to the indices previously analyzed in order to determine the existence of long-range correlations. Finally, we conclude that our results are compatible with a power law, and also with a TLF distribution. We also detected the presence of long-range correlations in several emerging markets economic indices.

2. The truncated Levy flight

Levy [10] and Khintchine and Levy [11] solved the problem of the determination of the functional form that all the stable distributions must follow. They found that the most general representation is through the characteristic functions $\varphi(q)$, by the following equation:

$$\ln \varphi(q) = \begin{cases} \mathrm{i}\mu q - \gamma|q|^2\left[1 - \mathrm{i}\beta\dfrac{q}{|q|}\tan\left(\dfrac{\pi}{2}\alpha\right)\right] & (\alpha \neq 1)\,, \\ \mu q - \gamma|q|\left[1 + \mathrm{i}\beta\dfrac{q}{|q|}\dfrac{2}{\pi}\ln|q|\right] & (\alpha = 1)\,, \end{cases}$$

where $0<\alpha\leqslant 2$, γ is a positive scale factor, μ is a real number and β is an asymmetry parameter that takes values in the interval $[-1, 1]$.

The analytic form for a stable Levy distribution is known only in the following cases:

$$\alpha = 1/2, \quad \beta = 1 \quad \text{(Levy}-\text{Smirnov distribution)}\,,$$

$$\alpha = 1, \quad \beta = 0 \quad \text{(Lorentz distribution)}\,,$$

$$\alpha = 2 \text{ (Gaussian distribution)}\,.$$

We consider symmetric distributions ($\beta = 0$) with zero mean value ($\mu = 0$). In this case the characteristic function takes the form:

$$\varphi(q) = \mathrm{e}^{-\gamma|q|^{\alpha}}\,.$$

As the characteristic function of a distribution is its Fourier transform, the stable distribution of index α and scale factor γ is

$$P_L(x) \equiv \frac{1}{\pi}\int_0^{\infty} \mathrm{e}^{-\gamma|q|^{\alpha}}\cos(qx)\,\mathrm{d}q\,.$$

The asymptotic behavior of the distribution for large values of the absolute value of x is given by

$$P_L(|x|) \approx \frac{\gamma\Gamma(1+\alpha)\sin(\pi\alpha/2)}{\pi|x|^{1+\alpha}} \approx |x|^{-(1+\alpha)}$$

and the value at zero $P_L(x = 0)$ by

$$P_L(x = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha\gamma^{1/\alpha}}\,.$$

The fact that the asymptotic behavior for large values of x is a power law has the following consequence: the stable Levy processes have infinite variance. In order to avoid the problems arising in the infinite second moment, Mantegna and Stanley [9] considered a stochastic process with finite variance that follows scale relations called TLF. The TLF distribution is defined by

$$P(x) = \begin{cases} 0, & x > l\,, \\ cP_1(x), & -l < x < l\,, \\ 0, & x < l\,, \end{cases}$$

where $P_1(x)$ is a symmetric Levy distribution and c is a normalization constant. The only stable distributions are the Levy distributions. The TLF distribution is not stable, but it has finite variance, so it converges to the Gaussian distribution, but unlike another pdf, its convergence is very slow. Another property of the TLF is that the cut it presents in its tails is very abrupt. Koponen [12] considered a TLF in which the cut function is a decreasing exponential characterized by a parameter l. The characteristic function of this distribution is defined by

$$\varphi(q) = \exp\left\{ c_0 - c_1 \frac{(q^2 + 1/l^2)^{\alpha/2}}{\cos(\pi\alpha/2)} \cos[\alpha \arctan(l|q|)] \right\}$$

with c_1 a scale factor and

$$c_0 \equiv \frac{l^{-\alpha}}{\cos(\pi\alpha/2)} .$$

If we discretize in time with steps Δt, we obtain that $T = N\Delta t$.

At the end of each interval, we must calculate the sum of N stochastic variables that are independent and identically distributed. The new characteristic function will be

$$\varphi(q, N) = \exp\left\{ c_0 N - c_1 \frac{N(q^2 + 1/l^2)^{\alpha/2}}{\cos(\pi\alpha/2)} \cos[\alpha \arctan(l|q|)] \right\} .$$

For small values of N, the return probability will be very similar to the stable Levy distribution:

$$P_L(x = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha(\gamma N)^{1/\alpha}} .$$

3. Methods and data analyses

We studied several Latin-American market indices using data of five countries (daily close values): Argentina (MERVAL), from October 8th, 1996 to July 18th, 2002; Chile (IPSA), from June 9th, 1997 to July 18th, 2002; Mexico (IPC), from November 8th, 1991 to July 15th, 2002; Peru (Lima General), from April 28th, 1998 to July 18th, 2002; and Brazil (BOVESPA), from April 27th, 1993 to July 15th, 2002.

The number of points for the MERVAL is 1,407, for IPSA is 1,262, for IPC is 2,654, for Lima General is 1,043 and finally for BOVESPA is 2,276. We also used data of the SP500, an index of the New York Stock Exchange. In the last case, the data correspond to two different periods: the first one from April 14th, 1950 to July 15th, 2002 (13,145 points) and the second one from November 8th, 1991 to July 15th, 2002 (2,694 points), in order to compare the results for developed countries with the results for emergent countries. For this reason, we considered two periods: one short period, including the periods corresponding to the emergent countries indices; and a longer one, in order to check the consequences of working with a small number of points.

In this work, the analyzed stochastic variable is the return (r_t), defined as the difference of the logarithm of two consecutive index prices:

$$G_t = \log I_t - \log(I_{t-T}) ,$$

where T is the difference (in labor days) between two values of the index. In order to compare the returns for different values of T, we define the normalized return by

$$g \equiv \frac{G - \langle G \rangle_T}{\sigma} ,$$

where $\sigma^2 \equiv \langle G^2 \rangle_T - \langle G \rangle_T^2$, and $\langle \cdots \rangle_T$ denotes the arithmetic average over the series of temporal scale T. The empirical study of temporal series of some of the most important indices shows that in short time intervals the associated pdf have larger kurtosis than a Gaussian distribution. In Fig. 1, we show the accumulated probability distribution for the two versions of the normalized return of the SP500 index and in Figs. 2–6, we show the accumulated probability distribution of the normalized returns of the Latin-American indices. These graphs depict an asymptotic behavior corresponding to power law:

$$P(g > x) \approx \frac{1}{x^{\alpha}} .$$

In Table 1, we give the exponents α calculated for each index, all the values are strictly greater than 2. These values are greater than the values corresponding to a Levy distribution; however, Weron [13] concluded that these values could be a consequence of working with finite samples. The behavior is compatible with a slow convergence to a Gaussian distribution. We can verify this statement by studying the

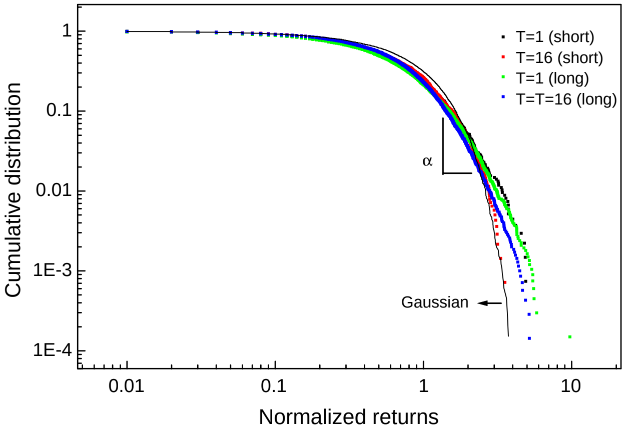

Fig. 1. For two different data series and for two different values of time scale T log–log plot of the cumulative distribution of the normalized returns of SP500 index. Short series contains 2,694 records and long series contain 13,145 records.

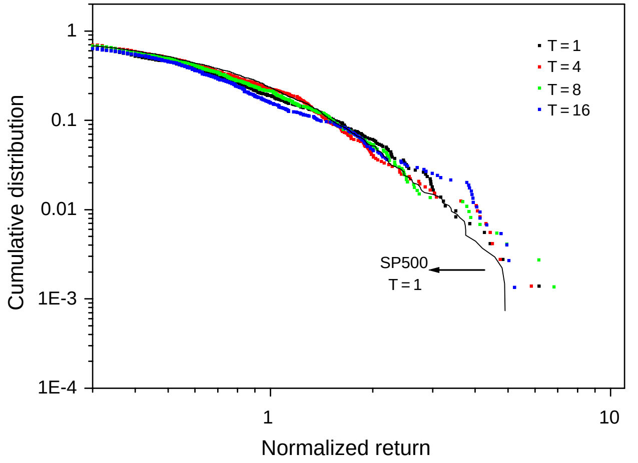

Fig. 2. For four different values of time scale T log–log plot of the cumulative distribution of the normalized returns of MERVAL index (Argentina).

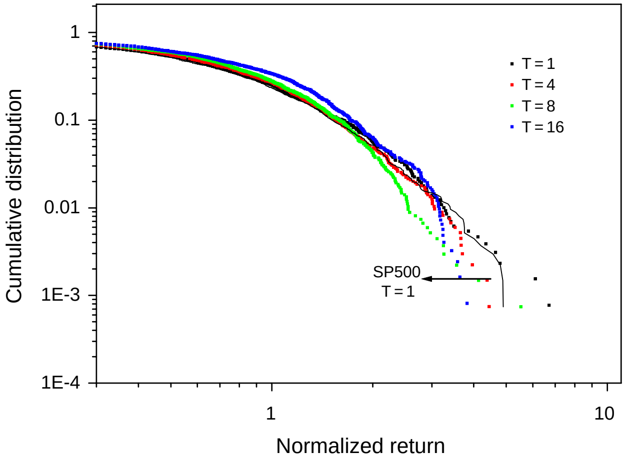

Fig. 3. For four different values of time scale T log–log plot of the cumulative distribution of the normalized returns of IPC index (Mexico).

momentum evolution. The momentum μ_k is defined by

$$\mu_k = \langle |g|^k \rangle_T \,.$$

The graph of the normalized return momentum for the SP500 for different values of T (Fig. 7) corresponds to the results of Gopikrishnan et al. [14], and has the same behavior as the one presented by Latin-American indices momentum, see, for example, the moments of the MERVAL index (Fig. 8).

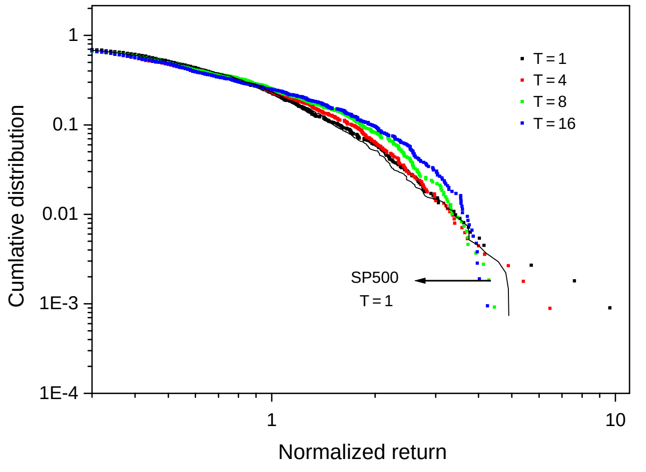

Fig. 4. For four different values of time scale T log–log plot of the cumulative distribution of the normalized returns of BOVESPA index (Brazil).

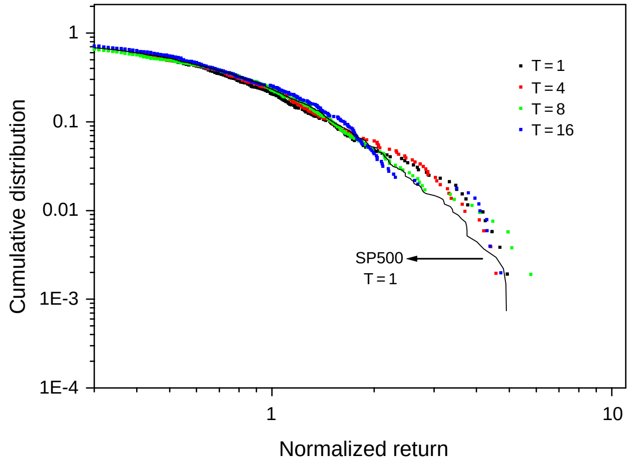

Fig. 5. For four different values of time scale T log–log plot of the cumulative distribution of the normalized returns of Lima General index (Peru).

4. Long-range correlations in market indices

A direct estimation of the correlation function defined by

$$C(\tau) \equiv \frac{\langle G(t)G(t+\tau)\rangle - \langle G(t)\rangle^2}{\langle G^2(t)\rangle - \langle G(t)\rangle^2}$$

is not appropriate in general, due to possible presence of unknown influences in the data series that would introduce problems in the calculation of the temporal mean

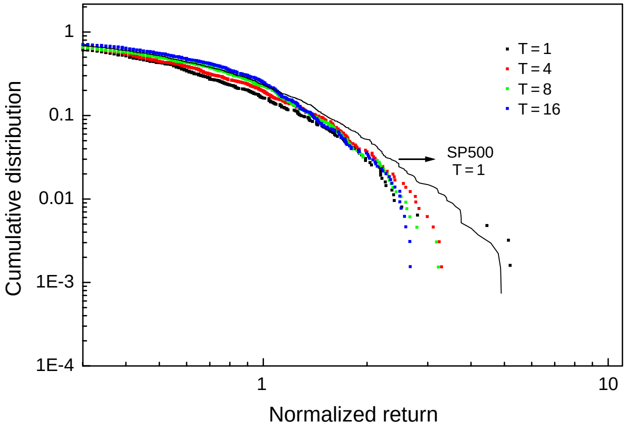

Fig. 6. For four different values of time scale T log–log plot of the cumulative distribution of the normalized returns of IPSA index (Chile).

Table 1
Values of the exponent α for four different values of time scale T (see also Figs. 1–6)

Index	SP500	MERVAL	IPC	BOVESPA	Lima General	IPSA
$T=1$	2.78	2.35	2.64	2.46	2.22	2.75
$T=4$	3.15	2.64	2.96	2.62	2.32	2.98
$T=8$	3.19	2.43	3.33	2.50	2.46	3.33
$T=16$	3.39	2.04	3.72	2.39	2.58	3.46

values. The spectral analysis is not appropriate even for the computation of the correlation function, because this method is applicable only to linear temporal series that are stationary or strictly periodic.

The DFA method is an important technique that allows detecting the presence of long-range correlations in non-stationary temporal series. This method was developed by Peng [15], and has been applied to the study of the DNA [16,17], the cardiac dynamics [18,19], climatic studies [20,21], solid state Physics [22,23], and economic series [24–27]. The DFA method is based in the theory of the random walks. In the determination of correlations, it is of fundamental importance to detect the difference between different tendencies arising in external agents, and intrinsic long-range correlations. Strong tendencies in the data may produce the detection of false long-range correlations, if a careful analysis of the results is not done [28]. On the other hand, the DFA method eliminates the tendencies of different order.

In order to apply the method to the study of the returns, first we have to integrate the absolute value of $g(t)$:

$$y(t) = \sum_{i}^{t} |g(i)| \,.$$

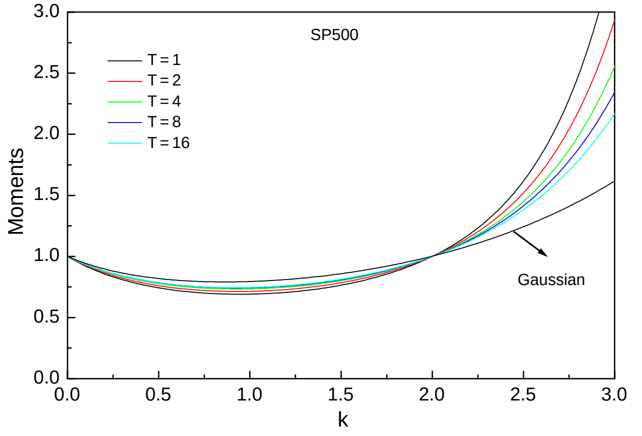

Fig. 7. Moments of the distribution of long data series of SP500.

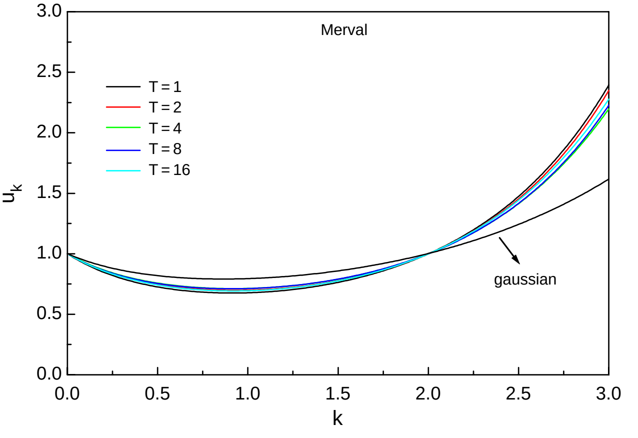

Fig. 8. Moments of the distribution of MERVAL index.

Then the integrated series is divided into intervals of length n, with no intersection between them. As the length of the series is not necessarily a multiple of n, there will remain a little amount at the end of $y(t)$. In order to take account of these values, we repeat the same procedure but beginning from the end, obtaining $N_s = 2N/n$ (with N the number of elements of $y(t)$) intervals with no intersection of equal length. In each interval the mean square linear regression is computed, obtaining a curve $(y_n(t))$. Finally we calculate

$$F(n) = \sqrt{\frac{1}{2N_s} \sum_{t=1}^{2N_s} [y(t) - y_n(t)]^2}\ .$$

This computation must be done over all the intervals of length n. A linear relation between $F(n)$ and n in a log–log plot says that there is a power law scaling with exponent α'. Therefore, the parameter α' that gives the relation between $F(n)$ and n characterizes the fluctuations. For autosimilar processes the exponent α' is related to the exponent β of the power spectrum in the following way:

$$\alpha' = \frac{1+\beta}{2} \; .$$

Then, for data series with no correlation, or with short-range correlation, it is expected that $\alpha' = (1/2)$. On the other hand, the values $(1/2) < \alpha' < 1$ show the presence of large-range correlations that follow power laws.

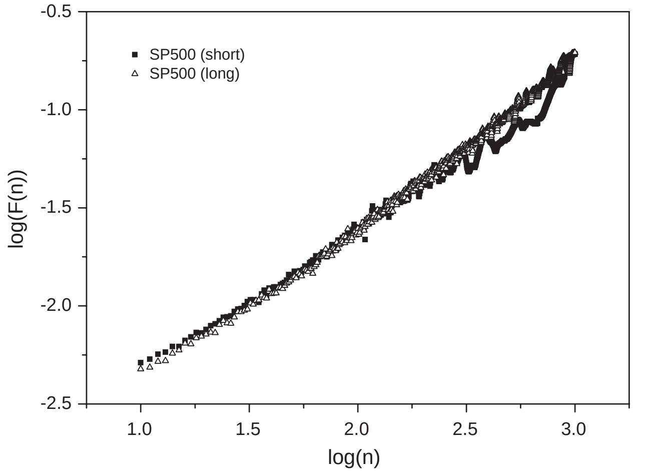

Fig. 9. DFA method applied to the short and long data series of the SP500 index.

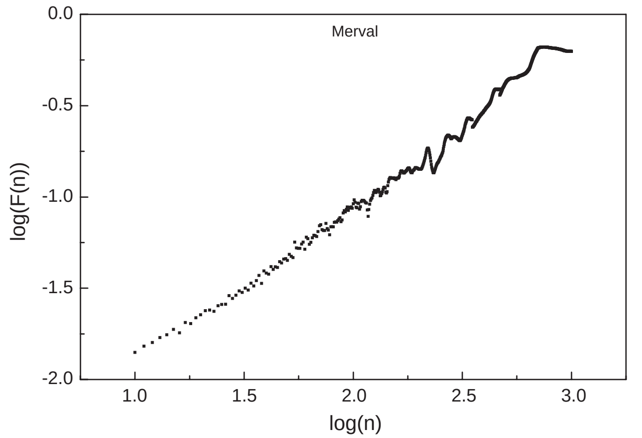

Fig. 10. DFA method applied to the data series of the MERVAL index.

This method was used for measuring correlations in financial series of high frequencies, and in the daily evolution of some of the most relevant indices.

In order to be able to observe the importance of the sample size when applying the DFA, we applied this method to the two series corresponding to the SP500 index (Fig. 9). We can observe that the series with less number of points presents bigger variations in the amplitude for the biggest values of $\log n$.

The graph obtained for Latin-American markets are shown in Figs. 10–14. Finally, Table 2 shows the exponents found for all the indices analyzed.

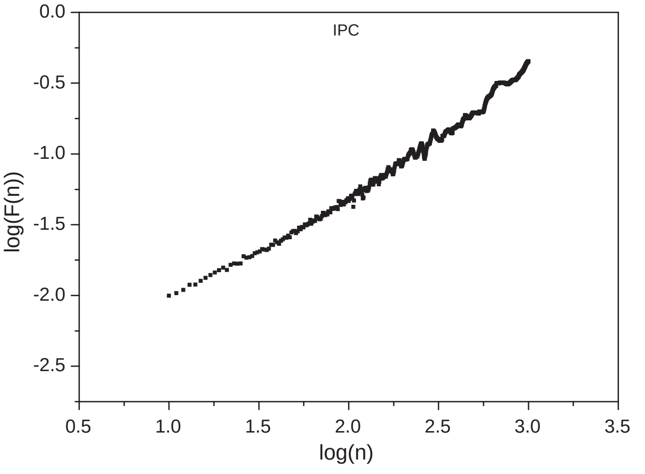

Fig. 11. DFA method applied to the data series of the IPC index.

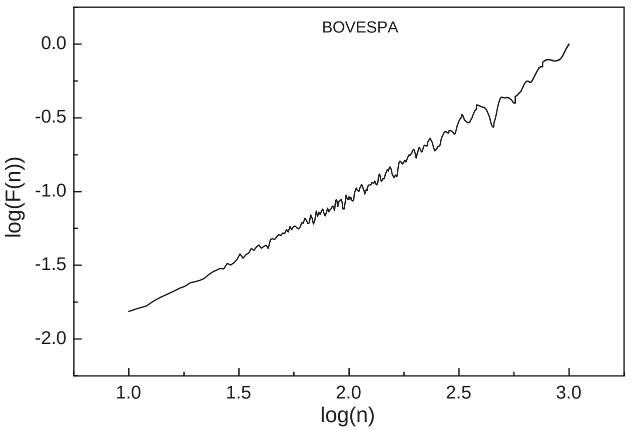

Fig. 12. DFA method applied to the data series of the BOVESPA index.

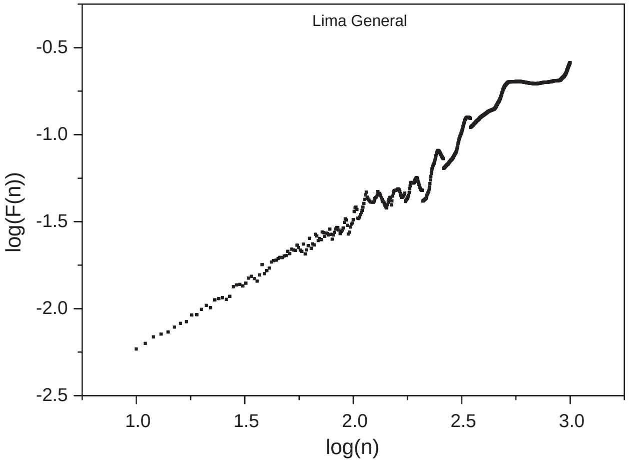

Fig. 13. DFA method applied to the data series of the Lima General index.

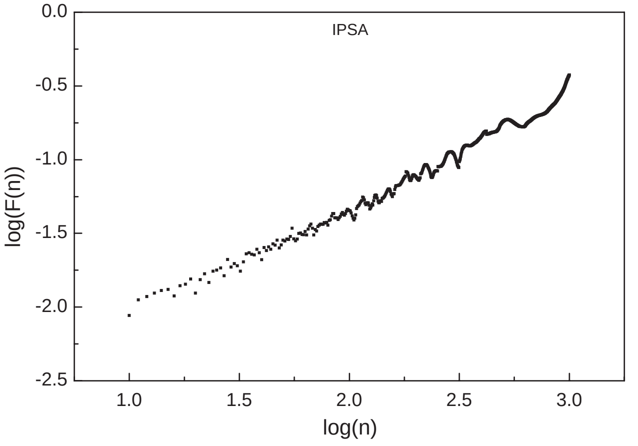

Fig. 14. DFA method applied to the data series of the IPSA index.

Table 2
DFA analysis for the absolute value of returns

Index	Length	α'	Error
SP500	13,145	0.858	0.002
SPS	2,694	0.782	0.005
ARG	1,407	0.940	0.004
MEX	2,654	0.893	0.003
BRA	2,276	0.953	0.004
PER	1,043	0.877	0.005
CHI	1,262	0.777	0.003

5. Conclusions

In this work, we did an empirical study of the statistical behavior of the most relevant Latin-American indices. Although we obtained compatible results with a potency law, it is not possible to conclude that the statistical distribution is not a TLF. This fact is because only when working with very large data sets (greater than 10^6 observations) is possible to detect a power behavior, allowing an estimation of the exponent α. Finally, we detected long-range correlations in the absolute value of the return. This fact has relevant consequences in the volatility dynamics, that is a fundamental variable in the option pricing models.

Acknowledgements

Partial support was provided by GRANTS: ADVANCE-NSF, UBACYT X O98 and X 202, CONICET.

References

[1] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, 1999.

[2] H.E. Stanley, L.A.N. Amaral, D. Canning, P. Gopikrishnan, Y. Lee, Y. Liu, Phys. A 269 (1999) 156.

[3] M. Ausloos, N. Vandewalle, Ph. Boveroux, A. Minguet, K. Ivanova, Phys. A 274 (1999) 229.

[4] J.-Ph. Bouchaud, M. Potters, Théorie des riques financiers, Alea-Saclay/Eyrolles, Paris, 1997.

[5] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.

[6] P. Willmott, J.N. Dewynne, S.D. Howison, Option Pricing: Mathematical Models and Computation, Oxford Financial Press, Oxford, 1993.

[7] B.B. Mandelbrot, J. Business 36 (1963) 394.

[8] P. Levy, Théorie de l'Addition des Variables Aléatoires, Gauthier-Villars, Paris, 1937.

[9] R.N. Mantegna, H.E. Stanley, Phys. Rev. Lett. 73 (1994) 2946.

[10] P. Levy, Calcul des probabilités, Gauthier-Villars, Paris, 1925.

[11] A.Ya. Khintchine, P. Levy, Sur les lois stables, C. R. Acad. Sci. Paris 202 (1936) 374.

[12] I. Koponen, Phys. Rev. E 52 (1995) 1197.

[13] R. Weron, Int. J. Modern Phys. C 12 (2) (2001) 209–223.

[14] P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Mayer, H.E. Stanley, Phys. Rev. E 60 (5) (1999) 5305.

[15] C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E 49 (1994) 1685.

[16] S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Matsa, C.K. Peng, M. Simons, H.E. Stanley, Phys. Rev. E 51 (1995) 5084.

[17] C.K. Peng, S.V. Buldyrev, A.L. Goldberger, R.N. Mantegna, M. Simons, H.E. Stanley, Phys. A 221 (1995) 180.

[18] C.K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, Phys. Rev. Lett. 70 (1993) 1343.

[19] C.K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, Chaos 5 (1995) 82.

[20] E. Koscienly-Bunde, H.E. Roman, A. Bunde, S. Havlin, H.J. Schellnhuber, Philos. Mag. B 77 (1998) 1331.

[21] E. Koscienly-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, H.J. Schellnhuber, Phys. Rev. Lett. 81 (1998) 729.

[22] J.W. Kantelhardt, R. Berkovits, S. Havlin, A. Bunde, Phys. A 266 (1999) 461.
[23] N. Vandewalle, M. Ausloos, M. Houssa, P.W. Mertens, M.M. Heyns, Appl. Phys. Lett. 74 (1999) 1579.
[24] Y.H. Liu, P. Cizeau, M. Meyer, C.K. Peng, H.E. Stanley, Phys. A 245 (1997) 437.
[25] P. Cizeau, Y.H. Liu, M. Meyer, C.K. Peng, H.E. Stanley, Phys. A 245 (1997) 441.
[26] M. Ausloos, N. Vandewalle, P. Boveroux, A. Minguet, K. Ivanova, Phys. A 274 (1999) 229.
[27] M. Ausloos, K. Ivanova, Phys. A 286 (2000) 353.
[28] L.A.N. Amaral, P. Cizeau, P. Gopikrishnan, Y. Liu, M. Meyer, C.K. Peng, H.E. Stanley, Comput. Phys. Commun. 121–122 (1999) 145.