Mutual Information: a dependence measure for nonlinear time series

Andreia Dionísio* Rui Menezes**
Diana A. Mendes**
*University of Évora, Largos Colegiais, 2; Management Department, 7000 Evora Portugal, andreia@uevora.pt
**ISCTE, Av. Forças Armadas, 1649-026 Lisboa, Portugal, rui.menezes@iscte.pt; diana.mendes@iscte.pt
First draft

November 10, 2003

Abstract

This paper investigates the possibility to analyse the structure of unconditional or conditional (and possibly nonlinear) dependence in financial returns without requiring the specification of mean-variance models or a theoretical probability distribution.

Abstract

The main goal of the paper is to show how mutual information can be used as a measure of dependence in financial time series. One major advantage of this approach resides precisely in its ability to account for nonlinear dependencies with no need to specify a theoretical probability distribution or use of a mean-variance model.

1 Introduction

Most of the theoretical literature in finance is based on arguments of market efficiency which imply unpredictability and independence of returns, leading to no profit opportunities. For a long period, the economists considered that the financial returns are independent. However, during the 1980s, it became widely accepted the fact that (linear) autocorrelations show some kind of "long-range dependence" and nonlinear dependence effects [Maasoumi *et al.*, (2002); Darbellay *et al.*, (2000)].

The linear autocorrelations of the returns are not statistically different from zero, except possibly for very short time lags, although the empirical evidence is mixed and the linear model results seem to be inconclusive. However the absence of linear autocorrelations is not synonymous of independence. Some recent studies indicate the presence of nonlinear dependence in financial time series [Hsieh, (1989); Cont, (2000); Darbellay *et al.*, (2000); Granger *et al.*, (1994); Granger *et al.*, (2002)].

The residuals of empirical economic models may incorporate nonlinearities, heterogeneity and serial dependence for many and varied reasons. Most studies deal with nonlinearities in the basis of the conditional mean and conditional variance [see e.g. Engle (1982); Hsieh (1989); Soares (1994); Qi (1999) and Curto, (2003)].

The use of linear or neural-network models, pose a number of problems because we test not only for the dependence of the model but also for its functional specification. Thus, some authors have explored nonlinear, nonparametric and semiparametric approaches. For instance White (1988) and Stengos (1995) considered nonlinear models and nonparametric regressions for returns on certain equities and precious metals to evaluate serial nonlinear dependence. Díaz, Grau-Carles and Mangas (2002) consider that nonlinearities in financial returns can be explained in two ways: they arise from a deterministic process that seems to be random (e.g. a chaotic process), or the returns are nonlinear stochastic functions of their own past. They conclude that the nonlinearities in foreign exchange rate returns can be the product of shifts in the variance, which is in contrast with the conclusion of other authors.

The most known measure of dependence between two random variables is the coefficient of linear correlation, but its application requires a pure linear relationship, or at least a linear transformed relationship [see e.g. Granger *et al.*, (1994); Bernhard *et al.*, (1999)], because it is nothing but a normalized covariance and only accounts for linear relationships. However, this statistics may not be helpful in determining serial dependence if there is some kind of nonlinearity in the data.

In this context, it seems that a measure of global dependence is required, that is, some measure that captures linear and nonlinear dependencies, without requiring the specification of any kind of model of dependence. Urbach (2000) defends a strong relationship between entropy, dependence and predictability. This relation has been studied by several authors, namely Granger and Lin (1994); Maasoumi and Racine (2002); Darbellay and Wuertz (2000).

On the basis of the above arguments we try to find out a rationale to the following question: "*Is it possible to inquire about any unconditional, or conditional (and possibly nonlinear) dependence structure in returns without requiring the specification of mean-variance models and theoretical distribution probabilities?*"

The main goal of this paper is to show that mutual information may be used as a measure of dependence in financial time series. One important advantage of this approach resides in its ability to account for nonlinear dependences, without any request about the theoretical probability distribution or mean-variance models.

We apply those concepts to some international stock indexes, in order to verify the possible existence of (linear and nonlinear) dependence and draw some conclusions about the efficiency of those markets.

This paper is organized as follows: in Section 2 we present the basic concepts of mutual information and the mathematical background for its estimation. Section 3 presents the main results and comprises 3 subsections: daily data, weekly data and monthly data. Finally, in Section 4 we present the concluding of this paper.

2 Information and predictability

A measure that takes the value 0 when there is total independence and 1 for total dependence is one of the most practical ways to evaluate (in)dependence between two vectors of random variables $\overrightarrow{X}, \overrightarrow{Y}$. Let $P_{\overrightarrow{X},\overrightarrow{Y}}(A \times B)$ be the joint probability distribution of $\left(\overrightarrow{X}, \overrightarrow{Y}\right)$ and $P_{\overrightarrow{X}}(A)$, $P_{\overrightarrow{Y}}(B)$ the underlying marginal probability distributions, where A is a subset of the observation space of $\overrightarrow{X}$ and B a subset of a observation space of $\overrightarrow{Y}$, such that we can evaluate the following expression:

$$\ln \frac{P_{\overrightarrow{X},\overrightarrow{Y}}(A \times B)}{P_{\overrightarrow{X}}(A) \times P_{\overrightarrow{Y}}(B)}. \tag{1}$$

If the two events are independent, then $P_{\overrightarrow{X},\overrightarrow{Y}}(A \times B) = P_{\overrightarrow{X}}(A) \times P_{\overrightarrow{Y}}(B)$, and so equation (1) will take the value zero.

Granger, Maasoumi and Racine (2002) consider that a good measure of dependence should satisfy the following six "ideal" properties:

(a) Must be well defined for both continuous and discrete variables;

(b) Must be normalized to zero if $\overrightarrow{X}$ and $\overrightarrow{Y}$ are independent, and lying between -1 and $+1$, in general;

(c) The modulus of the measure should equal 1 if there is an exact nonlinear relationship between the variables;

(d) Must be similar or simple related to the linear correlation coefficient in the case of a bivariate normal distribution;

(e) Must be metric in the sense that it is a true measure of "distance" and not just a measure of "divergence";

(f) Must be an invariant measure under continuous and strictly increasing transformations.

2.1 Mutual information

The concept of mutual information comes from the theory of communication and measures the information of a random variable contained in another random variable. The definition of mutual information goes back to Shannon (1948) and the theory was extended and generalized by Gelfand, Kolmogorov e Yaglom (1956) [*in* Darbellay, (1998a)] and Perez (1957).

The properties of mutual information appears to confirm its importance as a measure of dependence [Perez, (1957); Klan *et al.*, (1997); Soofi, (1997); Darbellay *et al.*, (1999), (2000); Darbellay, (1998a), (1998b), (1999); Bernhard *et al.*, (1999)]. Some of those properties will be presented and explored in this section.

Broadly speaking, there are two ways of estimating the mutual information: The first consists in a direct estimation and the second require previously the computation of the entropies in order to obtain mutual information[1].

The entropy of a continuous distribution, with probability density function (*pdf*) p_X of the random variable $X \in \overrightarrow{X}$ is defined by[2]:

$$H(X) = -\int p_X(x) \log p_X(x)\, dx. \tag{2}$$

The entropy of a continuous distribution can be negative [Shannon, (1948)] and may change if we change the coordinates of X [Fieldman, (1998)]. If we have two arguments X and Y, where then the *pdf* of Y is p_Y and $p_{X,Y}$ is the joint *pdf*, the joint entropy is given by:

$$H(X,Y) = -\int\int p_{X,Y}(x,y) \log p_{X,Y}(x,y)\, dxdy. \tag{3}$$

The conditional entropy is defined by:

$$\begin{aligned} H(Y|X) &= H(X,Y) - H(X) \\ &= -\int\int p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)} dxdy, \end{aligned} \tag{4}$$

or, in a similar way:

$$H(X|Y) = H(X,Y) - H(Y). \tag{5}$$

The mutual information can be defined by the following expression:

$$\begin{aligned} I(X,Y) &= H(Y) - H(Y|X) \\ &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X,Y) \\ &= \int\int p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)\, p_Y(y)} dxdy. \end{aligned} \tag{6}$$

[1] Note that the mutual information is the difference of entropies.

[2] The selection of the base of the logarithm is irrelevant, but is convenient to distinguish among results: $\log_2$- entropy measure in bits; $\log_{10}$- entropy measure in dits; $\log_e = \ln$ - entropy measure in nats.

Since $H(Y) \geq H(Y|X)$, we have $I(X,Y) \geq 0$, assuming equality if and only if X and Y are statistically independent. So the mutual information between the random variables X and Y can be considered a measure of dependence between these variables, or better yet, the statistical correlation of X and Y. Although, we can not say that is X is causing Y or vice-versa.

In the case of continuous distribution, the mutual information assumes non-negative values. If we have a discrete distribution, then $0 \leq I(X,Y) \leq \min\{H(X), H(Y)\}$. Pompe (1998) presents some of the properties of mutual information to the discrete case, namely:

(i) $I(X,Y) = 0$ iff X and Y are statistically independent, in the sense that $p(A \cap B) = p(A)\, p(B)$;

(ii) $I(X,Y) = H(X)$ iff X is a function of Y;

(iii) $I(X,Y) = H(Y)$ iff Y is a function of X.

The statistics defined in equation (6) satisfies some of the desirable properties of a good measure of dependence described in the previous section, namely (a) and after some transformations, will satisfies also properties (b), (c) and (d) [Granger *et al.*, (2002)].[3]

In order to satisfy the properties b) and d) it is convenient to define a measure that can be compared to the linear correlation coefficient. In equation (6), we have $0 \leq I(X,Y) \leq +\infty$, which difficult comparisons between different samples. In this way, we can compare mutual information with the covariance, since both are dependence measures and for both, comparisons between different samples can be inconclusive.

To obtain a statistic that satisfies property (d) without loosing the properties from (a) to (c) it is convenient to define an equation similar to that in (7). In this context Granger and Lin (1994), Darbellay (1998a) and Soofi (1997), among others, use a standard measure for the mutual information, global correlation coefficient, defined by:

$$\lambda = \sqrt{1 - e^{-2I\left(\overrightarrow{X}, \overrightarrow{Y}\right)}}. \tag{7}$$

This measure varies between 0 and 1 being thus directly comparable to the linear correlation coefficient, based in the relationship between the measures of information theory and variance analysis [Garner *et al*, (1956)].

The function $\lambda\left(\overrightarrow{X}, \overrightarrow{Y}\right)$ captures the overall dependence, both linear and nonlinear, between $\overrightarrow{X}$ and $\overrightarrow{Y}$ and it can be interpreted as the predictability of $\overrightarrow{Y}$ by $\overrightarrow{X}$. This measure of predictability is based on empirical probability distributions, but it does not depend on the particular model used to predict $\overrightarrow{Y}$. In this particular case, the properties mentioned above assume the following form:

[3] The demonstration of some theorems about mutual information properties can be found in Kullback, S. (1968). *Information Theory and Statistics,* Dover, New York.

- $\lambda\left(\overrightarrow{X},\overrightarrow{Y}\right) = 0$, if and only if $\overrightarrow{X}$ contains no information on $\overrightarrow{Y}$, which implies that $\overrightarrow{Y}$ cannot be predicted by means of $\overrightarrow{X}$;
- $\lambda\left(\overrightarrow{X},\overrightarrow{Y}\right) = 1$, if exists a perfect relationship between the vectors $\overrightarrow{X}$ and $\overrightarrow{Y}$. This is the limit case of determinism;
- When modelling the input-output pair $\left(\overrightarrow{X},\overrightarrow{Y}\right)$, by any model with input $\overrightarrow{X}$ and output $\overrightarrow{U} = f\left(\overrightarrow{X}\right)$, where f is some function of $\overrightarrow{X}$, the predictability of $\overrightarrow{Y}$ by $\overrightarrow{U}$ cannot exceed the predictability of $\overrightarrow{Y}$ by $\overrightarrow{X}$, i.e.

$$\lambda\left(\overrightarrow{X},\overrightarrow{Y}\right) \geq \lambda\left(\overrightarrow{U},\overrightarrow{Y}\right). \tag{8}$$

It is well known that the Gaussian distribution maximizes Shannon entropy for given first and second moments. This implies that the entropy of Shannon of any distribution is bounded upwards by the normal mutual information (NMI), and depends on the covariance matrix [Kraskov et al., (2003)). Let us consider a normal probability distribution, defined in a Euclidian space with dimension d. Then the normal mutual information for $\left(\overrightarrow{X},\overrightarrow{Y}\right)$ is given by:

$$\begin{aligned} I\left(\overrightarrow{X},\overrightarrow{Y}\right) &= \frac{1}{2}\log\frac{\det V_X \det V_Y}{\det V} \quad (9)\\ &\equiv NMI\left(\overrightarrow{X},\overrightarrow{Y}\right), \quad (10)\end{aligned}$$

where V is the covariance matrix of $\left(\overrightarrow{X},\overrightarrow{Y}\right)$ and V_X and V_Y are the respectively covariances matrices of $\overrightarrow{X}$ and $\overrightarrow{Y}$. It can be shown that the argument of the logarithm in the right-hand side of (9) depends only on the coefficients of linear correlation [see e.g. Darbellay, (1998a)]. When $d = 2$, that is, for $\left(\overrightarrow{X},\overrightarrow{Y}\right) = (X,Y)$ equation (9) takes the form [Kullback, (1968)]:

$$I(X,Y) = -\frac{1}{2}\log\left(1 - r^2(X,Y)\right). \tag{11}$$

If the empirical distribution is normal, the mutual information can be calculated by equation (11), because normal distribution is a "linear" distribution, in the sense that the linear correlation coefficient captures the overall dependence. In this case, any empirical mutual information must be greater or equal to the normal mutual information [Kraskov *et al.*, (2003)].

Intuitively, one would like to have the measure of predictability larger than the measure of linear predictability, i.e. $\lambda \geq r$. Unfortunately, this not always true [Darbellay, (1998b)][4]. It is important to refer that the difference $(\lambda - r)$

[4] A situation that can induce $\lambda < r$ is the small size of the sample. A small size, in this context, sample is a sample with $n \leq 500$.

cannot be equated to the nonlinear part of the predictability. Nevertheless, in the majority of cases, we do have $\lambda\left(\overrightarrow{X},\overrightarrow{Y}\right)=\left|r\left(\overrightarrow{X},\overrightarrow{Y}\right)\right|$, and in $\mathbb{R}^2$ we have $\lambda\left(X,Y\right)=\left|r\left(X,Y\right)\right|$ [Granger *et al.*, (1994); Darbellay, (1998a)].

Maasoumi (1993) shows that the mutual information doesn't satisfy property (e). In this case, mutual information is just a measure of divergence, because it does not satisfy the triangular inequality.

Another important property of the mutual information is the additivity, and it says that can be decomposed into hierarchical levels [Shannon, (1948); Kraskov *et al.*, (2003)], that is:

$$I\left(\overrightarrow{X},\overrightarrow{Y},\overrightarrow{Z}\right)=I\left(\left(\overrightarrow{X},\overrightarrow{Y}\right),\overrightarrow{Z}\right)+I\left(\overrightarrow{X},\overrightarrow{Y}\right). \tag{12}$$

It follows that $I\left(\overrightarrow{X},\overrightarrow{Y},\overrightarrow{Z}\right)$ will be always greater or equal to $I\left(\overrightarrow{X},\overrightarrow{Y}\right)$. By the same token, the coeficient of linear determination and the coeficient of linear correlation cannot decrease when one adds more variables to the model.

According to properties presented by mutual information, and because independence is one of the most valuables concepts in econometry, we can construct a independence test based on the following hypothesis:

$$\begin{aligned} H_0 &: \; p_{X,Y}\left(x,y\right)=p_X\left(x\right)p_Y\left(y\right), \\ H_1 &: \; p_{X,Y}\left(x,y\right)\neq p_X\left(x\right)p_Y\left(y\right). \end{aligned}$$

If $\left(p_{X,Y}\left(x,y\right)=p_X\left(x\right)p_Y\left(y\right)\right)$, then $I\left(X,Y\right)=0$ and the independence between the variables is found. If $\left(p_{X,Y}\left(x,y\right)\neq p_X\left(x\right)p_Y\left(y\right)\right)$ then $I\left(X,Y\right)>0$ and we reject the null hypothesis of independence. The above hypothesis can be reformulated in the following way:

$$\begin{aligned} H_0 &: \; I\left(X,Y\right)=0, \\ H_1 &: \; I\left(X,Y\right)>0. \end{aligned}$$

In order to test adequately the independence between variables (or vectors of variables) we will need to calculate the critical values. There are three approaches to obtaining critical values for our test under this null: asymptotic approximations to the null distribution; simulated critical values for the null distribution and permutation-based critical values for the null distribution.

The critical values calculated in this paper for mutual information are based upon simulated critical values for the null distribution or the percentile approach (see *Appendix A*). These values have been found through the simulation of critical values based upon a white noise, for a number of sample sizes. Given that the distribution of mutual information is skewed, we can adopt a percentile approach to obtain critical values.

Appendix A lists the 90^{th}, 95^{th} and 99^{th} percentiles of the empirical distribution of the mutual information for the process $y_t=\epsilon_t$ with $\epsilon_t \sim i.i.d.N\left(0,1\right)$, having been made 5000 simulations for each critical value. This methodology was applied as proposed by Granger, Maasoumi and Racine (2002), and according to these authors, the critical values can be used as the base to test for time series serial independence.

According to Pompe (1998), mutual information is very useful to analyze statistical dependences in scalar or multivariate time series as well as for detecting fundamental periods, detecting optimal time combs for forecasting, modelling and analyzing the (non)stationarity of data. Some of those potentialities have been explored by some authors, namely Granger and Lin (1994) and Darbellay and Wuertz (2000), whose results reveal that mutual information varies in a nonstationary time series framework.

2.2 Estimation from data - *marginal equiquantisation*

One difficulty for calculating the mutual information from empirical data lies in the fact that the underlying *pdf* is unknown. There are, essentially, three different methods to estimate mutual information:

- Histogram-based estimators;
- Kernel-based estimators;
- Parametric methods.

According to Moddemeijer (1999), histogram-based estimators are divided in two groups: equidistant cells (see e.g. Moddemeijer, 1999) and equiprobable cells, i.e. marginal equiquantisation [see e.g. Darbellay, (1998a)]. The second approach presents some advantages, since it allows for a better adequacy to the data and maximizes mutual information [Darbellay, (1998a)].

The kernel-based estimators have too many adjustable parameters such as the optimal kernel width and the optimal kernel form, and a non-optimal choice of those parameters may cause a large bias the in results [Granger *et al.*, (2000)]. Moreover, this kind of estimators can only deal with bivariate distributions. For the application of parametric methods one needs to know the specific form of the stochastic process.

The definition of mutual information is expressed in an abstract way and it is based on space partitions. To simplify, let us consider a finite dimension in an Euclidian space, $\mathbb{R}^d = \mathbb{R}^{d_X} \times \mathbb{R}^{d_Y}$, and let $\Gamma_X = \{A_i\}_{i=1}^{n_1}$ $\Gamma_Y = \{B_j\}_{j=1}^{n_2}$ be two generic partitions of the spaces $\mathbb{R}^{d_X}$ and $\mathbb{R}^{d_Y}$. Then the mutual information is a positive number defined as:

$$I\left(\overrightarrow{X}, \overrightarrow{Y}\right) \equiv \sup_{\{A_i\}\{B_j\}} \sum_{i,j} P_{\overrightarrow{X},\overrightarrow{Y}}\left(A_i \times B_j\right) \log \frac{P_{\overrightarrow{X},\overrightarrow{Y}}\left(A_i \times B_j\right)}{P_{\overrightarrow{X}}\left(A_i\right) P_{\overrightarrow{Y}}\left(B_j\right)}. \qquad (13)$$

The supreme is taken over all the finite partitions of $\mathbb{R}^{d_X}$ and $\mathbb{R}^{d_Y}$. The conventions $0 \ln\left(\frac{0}{z}\right) = 0$ for $z \geq 0$ and $z \ln\left(\frac{z}{0}\right) = +\infty$ are used. Darbellay (1998a) shown that mutual information is finite if and only if the measure $P_{\overrightarrow{X},\overrightarrow{Y}}$ is absolutely continuous with respect to the product measure $P_{\overrightarrow{X}} \times P_{\overrightarrow{Y}}$. The system

$\Gamma = \Gamma_X \times \Gamma_Y$ is a partition of $\mathbb{R}^d = \mathbb{R}^{d_X} \times \mathbb{R}^{d_Y}$ and is the product of two marginal partitions, one of $\mathbb{R}^{d_X}$ and another of $\mathbb{R}^{d_Y}$. Dobrushin (1959) shows that this restriction to the product partitions is not necessary [*in* Darbellay, *(*1998a)].

Let D be a sequence of numbers from a partition, so:

$$D_\Gamma \equiv \sum_{k=1}^{m} D_{C_k} \equiv \sum_{k=1}^{m} P_{\overrightarrow{X},\overrightarrow{Y}}(C_k) \log \frac{P_{\overrightarrow{X},\overrightarrow{Y}}(C_k)}{P_{\overrightarrow{X}}(C_k)\, P_{\overrightarrow{Y}}(C_k)}, \tag{14}$$

never decreases, as the partition $\Gamma = \{C_k = A_k \times B_k, k = 1, ..., m\}$ is made finer and finer.

Lemma 1 *Let* $\Gamma = \{C_k\}$ *be a finite partition of* $\mathbb{R}^d$ *and* $\Lambda = \{C_{k,l}\}$ *be a refinement of* Γ*, then*

$$D_\Gamma \leq D_\Lambda, \tag{15}$$

and the equality holds if and only if, for every cell C_k *of the partition* Γ*:*

$$\frac{P_{\overrightarrow{X},\overrightarrow{Y}}(C_{k,l})}{P_{\overrightarrow{X}}(C_{k,l})\, P_{\overrightarrow{Y}}(C_{k,l})} = \frac{P_{\overrightarrow{X},\overrightarrow{Y}}(C_k)}{P_{\overrightarrow{X}}(C_k)\, P_{\overrightarrow{Y}}(C_k)}, \forall l. \tag{16}$$

The inequality follows from $\ln z \leq z - 1$, $\forall z > 0$*, with* $\ln z = z - 1$ *if and only if* $z = 1$.

The lemma tells us that if we construct finer and finer partitions of Γ, then the sequence of numbers will monotonically increase until for every finite subpartitions $\{C_{k,l}; l = 1, ..., n_k\}$ of an arbitrary cell C_k of Γ is possible. This fact, according to Darbellay (1998a, 1999) shows that mutual information is a finite measure. The condition (16) means that if the random vectors $\overrightarrow{X}$ and $\overrightarrow{Y}$ are conditionally independent, then there is *local independence*. If this is true for every cell C_k of Γ, then we can set $I\left(\overrightarrow{X}, \overrightarrow{Y}\right) = D_\Gamma$.

Darbellay (1998a) and Bernhard and Darbellay (1999) show how to proceed for homogeneous partitions. These authors defend the use of equiprobable cells, given the flexibility and adequacy of this approach and in accordance with the invariance of mutual information under one-to-one transformations of its component variables:

$$\begin{aligned} & I\left(\left(f_1(X_1), ..., f_{d_a}(X_{d_a})\right), \left(f_{d_a+1}(X_{d_a+1}), ..., f_d(X_d)\right)\right) \\ = \; & I\left((X_1, ..., X_{d_a}), (X_{d_a+1}, ..., X_d)\right) \end{aligned} \tag{17}$$

It is possible to compute the marginal equiquantisation through different algorithms. Let t be a variable ranging from 1 to β, where β is the number of subpartitions and α^{td} the number of cells. The algorithms can be formulated according to the following rules:

Algorithm A:

1. Let $\mathbb{R}^d$ be the initial one-cell partition;

2. A subpartion of all cells into α^{td} subcells can be obtained by dividing each edge into α *equidistant intervals*;

3. Stop the subpartitioning of a cell if the vectors of random variables $\overrightarrow{X}$ and $\overrightarrow{Y}$ are *uniformly distributed*;

Algorithm B:

1. Let $\mathbb{R}^d$ be the initial one-cell partition;

2. A subpartion of all cells into α^{td} subcell can be obtained by dividing each edge into α *equiprobable intervals*;

3. Stop the subpartitioning of a cell if the vectors of random variables $\overrightarrow{X}$ and $\overrightarrow{Y}$ are *conditionally independent* on it.

The number α of equiprobable intervals is arbitrary. However, in order to simplify computation, we may choose $\alpha = 2$, because a large α will complicate unnecessarily the calculus [Darbellay, (1999)].

Marginal equiquantisation consists of dividing each edge of a cell into α intervals with approximately the same number of points. The approximativeness of the division has two causes: the number of points in a cell may not be exactly divisible by α, or some X may take repeating values. The lower (L) and upper (U) bounds and $x_i^k(L)$ and $x_i^k(U)$ $(i-th$ edge of the hyperrectangle final points of the cell) are found through marginal equiquantisation.

Our goal is to estimate $I\left(\overrightarrow{X},\overrightarrow{Y}\right)$ from a finite sample of N points $\overrightarrow{x} = (x_1, ..., x_d)$ in $\mathbb{R}^d$, then:

$N_{\overrightarrow{X,Y}}(C_k)$ represents the number of points $\overrightarrow{x}$ such that $x_i^k(L) < x_i < x_i^k(U), \forall i = 1, ..., d$. Then, the underlying marginal number of points is: $N_{\overrightarrow{X}}(C_k)$ which represents the number of points $\overrightarrow{x}$ such that $x_i^k(L) < x_i < x_i^k(U)$, $\forall i = 1, ..., d_X$ and $N_{\overrightarrow{Y}}(C_k)$ is the number of points $\overrightarrow{x}$ such that $x_i^k(L) < x_i < x_i^k(U)$, $\forall i = d_X + 1, ..., d$.

The probabilities are estimated by the underlying frequencies, i.e.,

$$P_{\overrightarrow{X},\overrightarrow{Y}}(C_k) \approx \frac{N_{\overrightarrow{X}\overrightarrow{Y}}(C_k)}{N} \tag{18}$$

$$P_{\overrightarrow{X}}(C_k) \approx \frac{N_{\overrightarrow{X}}(C_k)}{N} \tag{19}$$

$$P_{\overrightarrow{Y}}(C_k) \approx \frac{N_{\overrightarrow{Y}}(C_k)}{N}. \tag{20}$$

Then the local independence condition becomes:

$$N_{\overrightarrow{X},\overrightarrow{Y}}(C_{k,l}) \approx N_{\overrightarrow{X},\overrightarrow{Y}}(C_k)\frac{N_{\overrightarrow{X}}(C_{k,l})\,N_{\overrightarrow{Y}}(C_{k,l})}{N_{\overrightarrow{X}}(C_k)\,N_{\overrightarrow{Y}}(C_k)}, \qquad l = 1, ..., \alpha^{td}. \tag{21}$$

For each t, the subpartition $\{C_{k,l}\}$ of C_k is clearly finer, and therefore different. The $N_{\overrightarrow{X},\overrightarrow{Y}}(C_k)$ observations contained in cell C_k are classified into α^{td} mutually exclusive classes, the subcells. The right hand side of (21) is the expected number in class l, and the left hand side is the observed number in class l. These observed numbers follow a multinomial distribution whose probabilities are given by the expected numbers, which is a standard situation in statistics. In this context Darbellay (1998a) suggest the use of a χ^2 statistics rather then analyze all conditions about α^{td} in an individual way.

If we use *algorithm A*, the distribution is tested using the χ^2 test. We test whether the data are uniformly distributed in $(A \times B)$ by comparing the estimated probability of each subcube with the test probability $P_t(A_i \times B_j) = \frac{1}{\alpha^{td}}$:

$$\chi^2 = \sum_{l=1} \frac{\left(N(A_i \times B_j) - N(A \times B) P(A_i \times B_j)\right)^2}{N(A \times B) P(A_i \times B_j)}. \tag{22}$$

Here, $N(A \times B)$ denotes the number of points $\left(\overrightarrow{X}, \overrightarrow{Y}\right)$ following in the hyper-rectangle $A\times B$, A is a subset of the observation space of $\overrightarrow{X}$ and B a subset of the observation space of $\overrightarrow{Y}$.

If $\chi^2 < \chi^2_c$, being χ^2_c the critical value of the statistic, then the null hypothesis, i.e., the hypothesis of a uniform distribution holds with an error probability of 1% or 5%.

If we use *algorithm B*, the χ^2test is used to search for local independence, that is:

$$\chi^2 \equiv \chi^2(\{C_{k,l}\}) \equiv \frac{N_{\overrightarrow{X}}(C_k) N_{\overrightarrow{Y}}(C_k)}{N_{\overrightarrow{X},\overrightarrow{Y}}(C_k)} \sum_{l=1}^{\alpha^{td}} \frac{D^2(\{C_{k,l}\})}{N_{\overrightarrow{X}}(C_{k,l}) N_{\overrightarrow{Y}}(C_{k,l})}, \tag{23}$$

where

$$D^2(\{C_{k,l}\}) \equiv \left[N_{\overrightarrow{X},\overrightarrow{Y}}(C_{k,l}) - N_{\overrightarrow{X},\overrightarrow{Y}}(C_k) \frac{N_{\overrightarrow{X}}(C_{k,l}) N_{\overrightarrow{Y}}(C_{k,l})}{N_{\overrightarrow{X}}(C_k) N_{\overrightarrow{Y}}(C_k)}\right]^2, \tag{24}$$

or we can use the log-likelihood ratio statistic:

$$L \equiv L(\{C_{k,l}\}) \equiv \sum_{l=1}^{\alpha^{td}} N_{\overrightarrow{X},\overrightarrow{Y}}(C_{k,l}) \log\left[\frac{N_{\overrightarrow{X},\overrightarrow{Y}}(C_{k,l}) N_{\overrightarrow{X}}(C_k) N_{\overrightarrow{Y}}(C_k)}{N_{\overrightarrow{X},\overrightarrow{Y}}(C_k) N_{\overrightarrow{X}}(C_{k,l}) N_{\overrightarrow{Y}}(C_{k,l})}\right]. \tag{25}$$

The statistical test will be applied to each one of the β subpartitions into α^{td} subcells, with $t = 1, ..., \beta$. If the test does not reject local independence anymore on any cell, then we have the final partition $\Gamma = \{C_{k;} k = 1, ..., m\}$, upon which the sample's mutual information will be evaluated, i.e.,

$$I\left(\overrightarrow{X},\overrightarrow{Y}\right)=\frac{1}{N}\sum_{k=1}^{m}\left(N_{\overrightarrow{X},\overrightarrow{Y}}\left(C_k\right)\log\frac{N_{\overrightarrow{X},\overrightarrow{Y}}\left(C_k\right)}{N_{\overrightarrow{X}}\left(C_k\right)N_{\overrightarrow{Y}}\left(C_k\right)}\right)+\log N. \tag{26}$$

The value of β has a direct impact on the consistency of the estimator. A higher value of β prevents an early stop in the partioning of a cell, avoiding spurious results [Darbellay, (1998a)].

The level of significance is the probability of partioning a cell, when it shouldn't be done. The significance levels are (as usually) 1% and 5%, and according to Darbellay (1998a) the higher the value of β the higher the significance level, i.e., the significance level for the case where $\beta = 1$ should be smaller than the significance level for $\beta = 2$. The selection of β depends on the sample size as well as the dimension of the observation space.

Tambakis (2000) presents a mutual information estimator, which is based on equidistant cells. This author suggests the determination of a new measure: the *Self-Information Measure (SIM),* through the univariate non-parametric predictability computation, as a function of the mutual information and the number of partitions (K), that is:

$$P_K\left(\overrightarrow{X},\overrightarrow{Y}\right)=\frac{I\left(\overrightarrow{X},\overrightarrow{Y}\right)}{\log K}. \tag{27}$$

Asymptotic predictability is smallest when the entropy reachs the maximum, and is zero for an *iid* sequence. Based on the definition $I\left(\overrightarrow{X},\overrightarrow{Y}\right)=H\left(\overrightarrow{X}\right)-H\left(\overrightarrow{X}|\overrightarrow{Y}\right)$, Tambakis (2000) creates the *Self-Information Measure (SIM)* through the maximization of the entropy $H\left(\overrightarrow{X}\right)$:

$$SIM_K=\frac{H\left(\overrightarrow{X}\right)}{\log K}. \tag{28}$$

Tambakis (2000) applied this measure to several financial time series, in order to evaluate market efficiency, using $K = 100$. When the *SIM* increases, the power of predictability vanishes, and then the market is more efficient.

Paninski (2003) justified the difficulties about the computation of mutual information by the fact that is a nonlinear measure defined for a unknown joint probability space. The same author defends that when the ratio $\frac{N}{m}\rightarrow\infty$, with N being the number of observations and m the number of bins, the bias resulting from the estimation process decays. This author suggests the mutual information estimation based on a sequence of intervals, whose points are calculated by functions over the random variables in study. The "*method of sieves*" is based on a log-likelihood function [Paninski (2003)].

Moddemeijer (1999) points out some problems related with the estimation of mutual information based on histograms, namely:

- Variance;

- Bias caused by the finite number of observations;
- Bias caused by the quantization;
- Bias caused by the finite histogram.

The relative contribution of these factors depends on the empirical application, namely on the number of observations, the configuration of the histogram cells and the smoothness of the *pdf*. The last two factors are independent of the number of observations and are only relevant in the case of continuous variables. According to Darbellay (1998a), Darbellay and Vajda (1999), Kraskov, Stögbauer and Grassberger (2003) the space partition in equiprobable cells minimizes the bias.

3 Empirical evidence

We now apply the concepts of mutual information and global correlation coefficient (equation (7)) as a measures of dependence in financial time series in order to evaluate the use of these measures and extract the advantages of this approach face to the traditional linear correlation coefficient. Mutual information was estimated through marginal equiquantisation, and was applied to some stock market indexes.

>From the data base DataStream we selected the daily closing prices of several stock market indexes: ASE (Greece), CAC 40 (France), DAX 30 (Germany), FTSE 100 (UK), PSI 20 (Portugal), IBEX 35 (Spain) and S&P 500 (USA), spanning the period from 4/01/1993 to 31/12/2002, which corresponds to 2596 observations *per* index, in order to compute the rates of return.

The rates of return were computed in the following way:

$$r_{i,t} = \ln\left[\frac{P_{i,t} + D_{i,t}}{P_{i,t-1}}\right], \tag{29}$$

where $r_{i,t}$ is the stock market index i rate of return at moment t; $P_{i,t}$ is the stock market index i closing price at moment t; $D_{i,t}$ are the dividends and $P_{i,t-1}$ is the stock market index i closing price at moment $t-1$.

Weekly rates of return were computed through the mean of each week and we obtained 517 observations; to monthly rates of return we applied the same method which generate 119 observations. It was necessary to make some adjustments on the stock market index prices, especially in what concerns to the opening market days. To avoid loosing observations when some index does not have a price in some day, we used for that day's the previous price of that index.

3.1 Daily data analysis

In order to evaluate the possible serial dependence in some financial time series, we have tested for the possible linear dependence, in attempting to verify the presence of autocorrelation. To this end the *Ljung-Box* test was applied, and the results are presented in Table 1[5]. It is interesting to note the fact that serial dependence is statistically significant for most indexes, except for DAX 30, IBEX 35 and S&P 500[6].

	ASE	CAC 40	DAX 35	FTSE 100	IBEX 35	PSI20	S&P 500
LBQ(10)	37,197**	24,208**	17,43	45,716**	14,757	54,410**	16,782
ρ_1	0,108**	0,013	-0,025	0,022	0,032	0,129**	0,001
ρ_2	-0,006**	-0,028	-0,031	-0,058**	-0,035	0,039**	-0,025
ρ_3	-0,023**	-0,071**	-0,021	-0,084**	-0,047**	0,005**	-0,046

Table 1: *Ljung-Box* test and autocorrelation coefficients for daily observations.

Firstly, it was calculated the average mutual information and the global correlation coefficient for lags $k = 1, \ldots, 10$ of all indexes in study. The global correlation coefficient (λ) was calculated through equation (7). The indexes does not present the same behaviour in terms of serial dependence, as shown in Figure 1.

In a relatively coarse way we can group the indexes in three groups: the weak serial dependence group, constituted by the indexes CAC 40 and FTSE 100; the average serial dependence group constituted by the IBEX 35 and PSI 20 indexes; and finally, the strong serial dependence group: ASE, DAX 30 and S&P 500 indexes. The S&P 500 index presents the strongest global correlation coefficient, which along with a non-significant linear autocorrelation coefficient, seems to indicate that there exists nonlinear dependence. The higher order lags of the IBEX 35, DAX 30 and S&P 500 indexes have a stronger weight than normal. As noted by Bonanno, Lillo and Mantegna (2001) and Mantegna, Palági and Stanley (1999) the autocorrelation function should be a monotonically decreasing function with the time lag, being actually significant for short periods. This is some way not confirmed by our results.

The property of "*short run memory*" is related to the efficient market hypothesis, where investors can not make systematic profits. However, the lack of autocorrelation does not mean independence; nonlinear dependence can be significant for higher lags. The authors state that the presence of significant nonlinear dependence can be extended for about 20 days [Bonanno *et al.*, (2001); Mantegna *et al.*, (1999)].

We have calculated the average mutual information and the global correlation coefficient (λ) for the time series above mentioned and compared the results with the normal mutual information (equation (9)) and with the linear correlation coefficient (see Tables 2, 3 and 4).

According to the results in Table 2, the linear correlation and global corre-

[5] ** 1% level of significance
* 5% level of significance

[6] We have also applied the *LM* test, but the results were not significantly different from the *Ljung-Box* results.

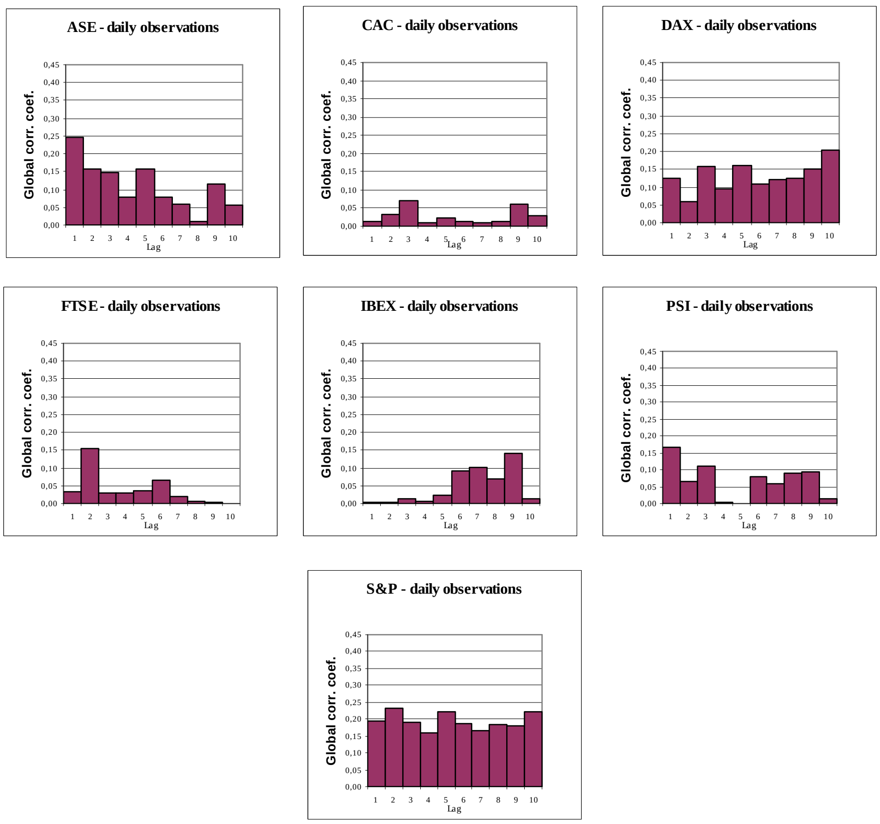

Figure 1: Global correlation coefficient for serial dependence in the stock market index daily returns, for lags $k = 1, ..., 10$.

	Mutual Inf.	*NMI*	λ	r
lag: 1				
ASE	0.03126**	0.00583	0.24618	0.10760
CAC 40	0.00008	0.00009	0.01273	0.01319
DAX 30	0.00772**	0.00031	0.12379	0.02470
FTSE 100	0.00054	0.00023	0.03277	0.02153
IBEX 35	0.00000	0.00051	0.00193	0.03203
PSI 20	0.01412**	0.00841	0.16687	0.12914
S&P 500	0.01901**	0.00000	0.19316	0.00118

Table 2: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of daily data relative to 1 lag.

lation coefficients are not ranked in the same way:

$$r_{S\&P500} < r_{CAC40} < r_{FTSE100} < r_{DAX30} < r_{IBEX35} < r_{ASE} < r_{PSI20},$$
$$\lambda_{IBEX35} < \lambda_{CAC40} < \lambda_{FTSE100} < \lambda_{DAX30} < \lambda_{PSI20} < \lambda_{S\&P500} < \lambda_{ASE}.$$

The index S&P 500 is the one that presents the least linear correlation with regard to the first lag and the largest global correlation, indicating the possible existence of strong nonlinear serial dependence.

	Mutual Inf.	*NMI*	λ	r
lag: 1 e 2				
ASE	0.08042**	0.01191	0.38545	0.10901
CAC 40	0.01115*	0.00087	0.14853	0.0379
DAX 30	0.03701**	0.00105	0.26709	0.04018
FTSE 100	0.01508**	0.00174	0.17238	0.06247
IBEX 35	0.00001	0.00082	0.00488	0.04855
PSI 20	0.03596**	0.01706	0.26343	0.12992
S&P 500	0.07531**	0.00032	0.37392	0.02476

Table 3: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of daily data relative to 2 lags.

Table 3 shows again that the linear and global correlation coefficients are not dispose in the same way:

$$r_{S\&P500} < r_{CAC40} < r_{DAX30} < r_{IBEX35} < r_{FTSE100} < r_{ASE} < r_{PSI20},$$
$$\lambda_{IBEX35} < \lambda_{CAC40} < \lambda_{FTSE100} < \lambda_{PSI20} < \lambda_{DAX30} < \lambda_{S\&P500} < \lambda_{ASE}.$$

Once more, the S&P 500 index is the one that presents least linear correlation face to lags 1 and 2; however is the one having the highest global correlation, suggesting the existence of possible strong serial nonlinear dependence. The ASE index does not lead to the same conclusions. It is the index that presents the highest evidence of simultaneously linear and global dependence. The PSI 20 index is the one that presents the highest value for the linear correlation coefficient (r), but its global correlation coefficient (λ) is less than those for the DAX 30, ASE and S&P 500 indexes.

	Mutual Inf.	*NMI*	λ	r
lag: 1, 2 e3				
ASE	0.10195**	0.01745	0.42948	0.11092
CAC 40	0.00893	0.00359	0.13308	0.07683
DAX 30	0.05254**	0.00341	0.31584	0.04821
FTSE 100	0.04030**	0.00808	0.27829	0.10231
IBEX 35	0.00073	0.00223	0.03826	0.06566
PSI 20	0.06559**	0.02500	0.35063	0.12990
S&P 500	0.11376**	0.00152	0.45111	0.05229

Table 4: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of daily data relative to 3 lags.

The analysis of Table 4 show some differences about the way the two correlation coefficients are ranked:

$$
\begin{aligned}
r_{DAX30} &< r_{S\&P500} < r_{IBEX35} < r_{CAC40} < r_{FTSE100} < r_{ASE} < r_{PSI20}, \\
\lambda_{IBEX35} &< \lambda_{CAC40} < \lambda_{FTSE100} < \lambda_{DAX30} < \lambda_{PSI20} < \lambda_{ASE} < \lambda_{S\&P500}.
\end{aligned}
$$

The global correlation coefficient (λ) is higher than the linear correlation coefficient (r) for the majority of the indexes, denoting the existence of possible serial nonlinear dependence. The S&P 500 index presents a very strong nonlinear dependence (already proven by the values shown in Figure 1). According to the *Ljung-Box* test results this index does not show evidence of the existence of significant linear autocorrelation, but such conclusion is not true for the global autocorrelation, whose values are above $0,19$ in all analysed situations.

Relatively to the analysis of the serial dependence face to the two first lags ($t-1$, $t-2$) and to the three first lags ($t-1$, $t-2$, $t-3$) it is possible to ascertain that the nonlinear dependence tends to grow when the number of lags included increases, while the linear correlation coefficient tends to assume a significantly constant value in all analyses. The results allow us to verify the mutual information additivity property, since when we include more lags in the analysis, mutual information tends to increase. It is important to remember that the difference ($\lambda - r$) does not correspond exactly to the nonlinear part of the measure of dependence.

Our results point to the existence of global dependence larger than the one that is captured by the linear correlation coefficient. It becomes interesting to analyze the filtered series, in order to capture the exclusively nonlinear dependence. For the effect, the filtered series was calculated as an $ARMA(3,0)$ process and applied the *Ljung-Box* test (Table 5). The filtered series for the indexes DAX 30 and S&P 500 was not calculated because they do not present, for any lag, a statistically significant linear correlation coefficient.

	ASE	CAC 40	FTSE 100	IBEX 35	PSI20
LBQ(10)	7,484	11,702	37,572	9,012	9,4532
ρ_1	0,002	0,011	0,018	0,031	-0,002
ρ_2	-0,015	-0,03	-0,001	-0,034	0,013
ρ_3	-0,02	-0,002	-0,085	-0,001	-0,002

Table 5: *Ljung-Box* test for the filtered daily series.

The filtered series of all indexes do not show any evidence of linear autocorrelation. In this context, to verify if there exists nonlinear dependence, we applied the BDS test [Hsieh, (1989)] where the results point clearly to the rejection of the null hypothesis, or either, is rejected the hypothesis of that the time series in observation, are nonlinear independent. The measures of information theory were also applied to the filtered series, and the results confirm the previous BDS tests results. Mutual information also allow the knowledge of the relations intensity between ϵ_t e a $\epsilon_{t-1.}$ Table 6 shows that $\lambda \geq r$ for all cases (except for the IBEX 35 index), with the evidence of the ASE and PSI 20 indexes which present a very high global correlation coefficient and statistically significant, and that could be a clear indicator of nonlinear dependence.

	$Mutual\ Inf.$	NMI	λ	r
lag: 1				
ASE	0.01142**	0.00000	0.15025	0.00190
CAC 40	0.00224*	0.00006	0.06690	0.01130
FTSE 100	0.00051	0.00014	0.03205	0.01673
IBEX 35	0.00001	0.00047	0.00541	0.03074
PSI 20	0.00535**	0.00000	0.10321	0.00166

Table 6: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of filtered daily data.

The levels of global correlation are very high, especially if we take into account that the linear expression is not significant. Our results are similar to other author's results in similar studies, namely Darbellay and Wuertz (2000) and Maasoumi and Racine (2002).

The presence of serial dependence must provide some strategy to generate systematic profits, for that market can be considered not efficient, being a efficient market some market where is not possible to make profits in a systematic way, because there is any investor with more information than others.

According to Fama (1970, 1991) the presence of serial statistically significant dependence does not mean that the market is not efficient, since the existence of transaction costs would eliminate any attempts of systematic profits. In this case, the investors must found which function is behind the global and nonlinear dependence to use that to take decisions and try to make profits, even with the presence of transaction costs.

3.2 Weekly data analysis

The weekly data analysis is based on weekly averages of the daily rate of returns, for about 517 observations. The *Ljung-Box* test applied to the weekly stock market indexes' time series reveals the existence of significant linear dependence for all the indexes (see Table 7). It is important to note that the linear dependence in weekly observations is stronger than the one observed for daily observations. That might be explained by the fact that in the present analysis the observations correspond to average weekly data. This methodology allows to not loose observations, or at least, the new time series is more representative than the daily one, but, at the same time there is a smooth on the weekly time series, which can increase the autocorrelation values. Another explanation for the higher linear autocorrelation is that weekly averages are not isolated values of a day in a week or month, and also because the daily movement can contain some bias resulting from slow dilutions of information in the markets.

	ASE	CAC 40	DAX 30	FTSE 100	IBEX 35	PSI20	S&P 500
LBQ(10)	21,207*	19,066*	25,402**	16,668	37,002**	53,526**	29,480**
ρ_1	0,179**	0,143**	0,167**	0,110*	0,209**	0,278**	0,167**
ρ_2	-0,003**	-0,003**	0,022**	0,019*	0,090**	0,066**	0,032**
ρ_3	0,027**	0,027*	-0,013**	-0,024	-0,002**	-0,002**	0,020**

Table 7: *Ljung-Box* test and autocorrelation coefficients for weekly observations.

In order to capture the possible nonlinear serial dependence that may exist in the analysed series, we have calculated the average mutual information and the corresponding global correlation coefficient, normal mutual information and linear correlation coefficient.

Figure 2 reveals slightly higher levels of global correlation as compared to the ones presented for daily data, as well as for linear correlation. The PSI 20 index presents the highest correlation value for lag 1, but the remaining correlations tend to decay quickly, which does not happen with the FTSE 100, DAX 30 and ASE indexes, whose correlations for lagged observations do not tend to diminish with the increase of respective lags. Though, we can say that the majority of the indexes present "*short run memory*", i.e., the serial dependence tends to vanish for higher lags. We should also note that for lag 10 there is no index with a global correlation coefficient greater than 0.06.

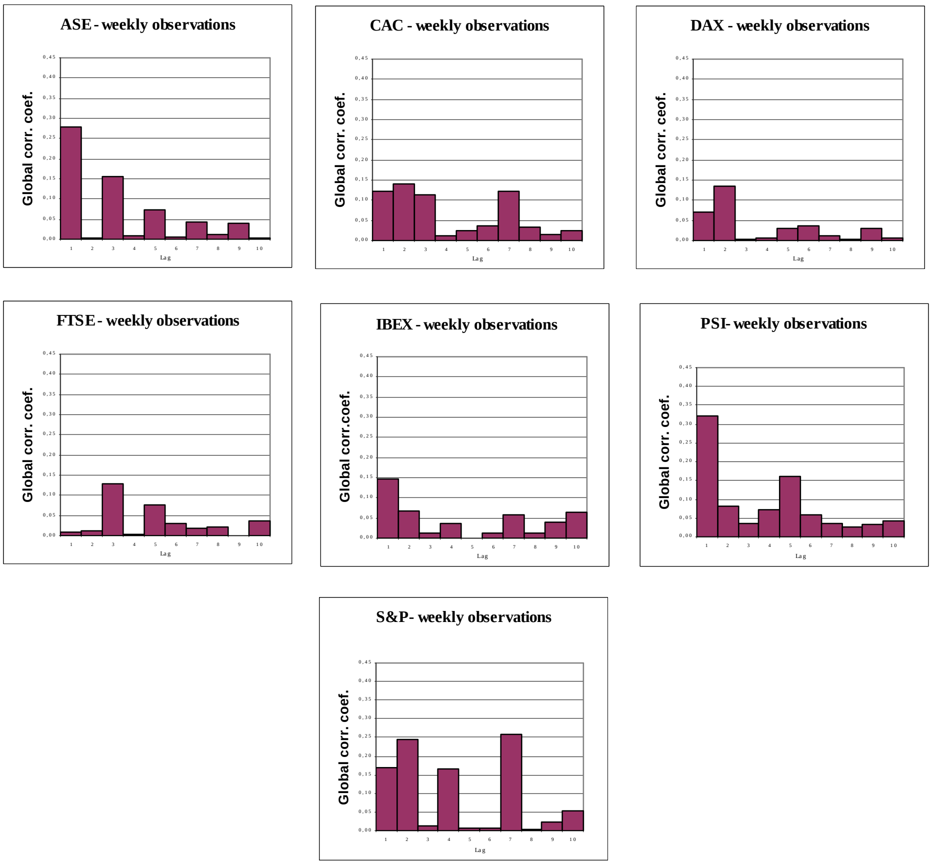

Figure 2: Global correlation coefficient for serial dependence in stock market indexes weekly returns, for lags $k = 1, ..., 10$.

Table 8 evidences the highest value for the linear correlation coefficient as well as for the global correlation coefficient, for the majority of the indexes. We have then:

$$r_{FTSE100} < r_{CAC40} < r_{S\&P500} < r_{DAX30} < r_{ASE} < r_{IBEX35} < r_{PSI20}.$$

This arrangement differs from the one for the global correlation coefficient, that is:

$$\lambda_{FTSE100} < \lambda_{DAX30} < \lambda_{CAC40} < \lambda_{IBEX35} < \lambda_{S\&P500} < \lambda_{ASE} < \lambda_{PSI20}.$$

	Mutual Inf.	*NMI*	λ	r
lag: 1				
ASE	0.04037**	0.01639	0.27851	0.17956
CAC 40	0.00752*	0.01028	0.12222	0.14268
DAX 30	0.00244	0.01424	0.06971	0.16755
FTSE 100	0.00003	0.00609	0.00775	0.11000
IBEX 35	0.01089*	0.02226	0.14676	0.20867
PSI 20	0.05489**	0.04031	0.32244	0.27833
S&P 500	0.01458**	0.01413	0.16953	0.16691

Table 8: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of weekly data relative to 1 lag.

The Portuguese PSI 20 index exhibits the highest levels of correlation (linear and nonlinear) and the FTSE 100 index seems to be the most "independent" from the past, which seems to indicate the effects of persistence. We should also note that S&P 500 index, whose global dependence relative to observation $(t-1)$ is stronger than the simple linear dependence, especially if we compare it with the German DAX 30 index.

Tables 9 and 10 present the measures of information theory for lags 1 and 2 and 1, 2 and 3. From Table 9 one should note that the PSI 20 index continues to lead the serial dependence at a linear level but globally the ASE index presents the highest global correlation coefficient.

	Mutual Inf.	*NMI*	λ	r
lag: 1 e 2				
ASE	0.08563**	0.02944	0.39674	0.18122
CAC 40	0.01961	0.01829	0.19612	0.14643
DAX 30	0.04497*	0.02551	0.29328	0.16913
FTSE 100	0.00066	0.01192	0.03619	0.11000
IBEX 35	0.04073*	0.03304	0.27969	0.21205
PSI 20	0.07338**	0.07877	0.36945	0.27917
S&P 500	0.05750**	0.02948	0.39959	0.16695

Table 9: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of weekly data relative to 2 lags.

The order of those measures is the following:

$$\begin{aligned} r_{FTSE100} &< r_{CAC40} < r_{S\&P500} < r_{DAX30} < r_{ASE} < r_{IBEX35} < r_{PSI20}, \\ \lambda_{FTSE100} &< \lambda_{CAC40} < \lambda_{IBEX35} < \lambda_{DAX30} < \lambda_{S\&P500} < \lambda_{PSI20} < \lambda_{ASE}. \end{aligned}$$

In Table 10 we verify that the global correlation coefficient assumes relatively high values for the majority of the indexes.

	Mutual Inf.	*NMI*	λ	r
lag: 1 , 2 e 3				
ASE	0.07106	0.04397	0.36399	0.18243
CAC 40	0.08927*	0.02656	0.40436	0.15161
DAX 30	0.01703	0.03822	0.18301	0.16951
FTSE 100	0.00400	0.01576	0.08927	0.11337
IBEX 35	0.10350*	0.05293	0.43241	0.21501
PSI 20	0.11263**	0.11805	0.44909	0.27746
S&P 500	0.13329**	0.04088	0.48374	0.16745

Table 10: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of weekly data relative to 3 lags.

The correlation coefficients arrangement evidences the changes that occur for the CAC 40 and S&P 500 indexes:

$$r_{FTSE100} < r_{CAC40} < r_{S\&P500} < r_{DAX30} < r_{ASE} < r_{IBEX35} < r_{PSI20},$$
$$\lambda_{FTSE100} < \lambda_{DAX30} < \lambda_{ASE} < \lambda_{CAC40} < \lambda_{IBEX35} < \lambda_{PSI20} < \lambda_{S\&P500}.$$

For instance, $r_{CAC40} < r_{ASE}$ and $\lambda_{ASE} < \lambda_{CAC40}$, meaning that the ASE index presents a linear correlation stronger than CAC 40 index, which does not happen at a global level. Globally, the CAC 40 index presents a higher value of dependence relative to the three first lags ($t-1$, $t-2$, $t-3$) than ASE index, for instance.

In a similar way, the series were filtered as in the daily data analysis, in order to capture exclusively nonlinear dependence. To this end, the filtered series were calculated through an $ARMA(3,0)$ process having been applied the *Ljung-Box* test, whose values are presented in Table 11.

	ASE	CAC 40	DAX 30	FTSE 100	IBEX 35	PSI20	S&P 500
LBQ(10)	7,580	9,913	11,582	10,443	11,350	14,388	14,436
ρ_1	0,004	0,005	0,002	-0,001	-0,013	0,004	-0,001
ρ_2	-0,043	-0,029	-0,003	0,010	0,053	-0,006	0,001
ρ_3	0,034	0,041	-0,002	-0,022	-0,014	-0,037	0,030

Table 11: *Ljung-Box* test for the filtered weekly series.

The linear autocorrelation coefficients are not significantly different of zero; therefore there is no evidence of linear dependence for the indexes faced to the lags. We also applied a BDS test (see *Appendix B*) which leads us to reject the null hypothesis of independence for almost all indexes, except for the FTSE 100 and IBEX 35, which present conflict results. In order to clarify the possible existence of nonlinear dependence in the filtered series, we computed the measures of information theory (Table 12).

	Mutual Inf.	*NMI*	λ	r
lag: 1				
ASE	0.00183	0.00001	0.06039	0.00472
CAC 40	0.00101	0.00001	0.04482	0.00510
DAX 30	0.01105*	0.00000	0.14786	0.00200
FTSE 100	0.00289	0.00000	0.07595	0.00100
IBEX 35	0.00000	0.00009	0.00195	0.01353
PSI 20	0.01961**	0.00001	0.19612	0.00387
S&P 500	0.00828*	0.00000	0.12816	0.00000

Table 12: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient for filtered daily data.

In Table 12 the linear correlation coefficient presents very low values, as well as the mutual information based on the normal distribution (IMN). There is also a sharp decline in the global correlation coefficients. Although, the values of the global correlation coefficient are much higher than those presented by the linear correlation coefficient, indicating the possible presence of nonlinear dependence in the analyzed time series. The arrangement of the measures shows us the existence of many discrepancies. The linear correlation coefficient can be disposed in the following way:

$$r_{S\&P500} < r_{FTSE100} < r_{DAX30} < r_{PSI20} < r_{ASE} < r_{CAC40} < r_{IBEX35},$$

while the global correlation coefficient presents the following arrangement:

$$\lambda_{IBEX35} < \lambda_{CAC40} < \lambda_{ASE} < \lambda_{FTSE100} < \lambda_{S\&P500} < \lambda_{DAX30} < \lambda_{PSI20}.$$

The IBEX 35 index is the one presenting the major levels of linear correlation and also the minor value for global correlation, which leads us to conclude that there is no nonlinear dependence, while the S&P 500 index presents a practically null linear correlation but a high global correlation with a value of 0.12816. The FTSE 100 index does not lead us to a similar conclusion because $\lambda = 0.07595$.

Although the global correlation coefficient presents high values of dependence faced to lags 1, 2 and 3, as Tables 8,9 and 10 confirm, there is a high probability that this global correlation is supported by the linear correlation, since the values for the global correlation coefficient of the filtered series are not too high. That does not happen in the DAX 30, PSI 20 and S&P 500 indexes, which evidence the presence of a nonlinear dependence. This fact can not be explained by market characteristics (namely dimension and liquidity) of the indexes, respectively, since they are indexes from very different markets.

3.3 Monthly data analysis

In the monthly data analysis we considered the averages monthly rate of return of the indexes under study, which gave about 119 observations. In this analysis the sample is relatively small, especially for estimation of entropy and mutual information.

Darbellay and Vadja (1999) alert for the possibility of underestimating the mutual information when we use small samples (less than 500 observations). In this context, the same authors advise to use a higher value for the β parameter used to evaluate if the subpartition must or not be realized. The authors [Darbellay *et al.*, (1999)] consider $\beta = 4$, being β the number of subpartitions in the space.

Once more, we applied a *Ljung-Box* test to evaluate linear dependence in the data concerning to its lagged values (Table 13). The test results reveal the existence of serial linear dependence in a significant way just for the PSI 20 index, and in a weaker way, to ASE and DAX 30 indexes. In this case, can we expect that the indexes are independent and identically distributed? Obviously that such conclusion will only be true if and only if the nonlinear dependence is not statistically significant.

	ASE	CAC 40	DAX 30	FTSE 100	IBEX 35	PSI20	S&P 500
LBQ(10)	17,429	6,7866	9,2902	15,745	14,848	17,634	19,652
ρ_1	0,232*	0,155	0,229*	0,149	0,208*	0,324**	0,200*
ρ_2	0,009*	0,022	0,073*	0,029	-0,117*	0,019**	0,024
ρ_3	0,043	0,034	-0,026	0,065	-0,023	-0,009**	0,087

Table 13: *Ljung-Box* test and autocorrelation coefficients for monthly observations.

>From Figure 3 we can see that there is a serial dependence of the monthly observations. In a general way, the global correlation tends to decrease for time lag increases, such as verified for the daily and weekly observations. ASE, DAX 30 and PSI 20 indexes present the highest global correlations value, that, in a certain way, does not refute our previous conclusions.

We calculated the average mutual information, the normal mutual information, the coefficient of linear correlation (r) and the coefficient of global correlation (λ) for lags considered in a joint way (Table 14). From Table 14 we can see that the PSI 20 index presents the highest levels of linear and nonlinear correlation, and FTSE 100 index presents the least linear correlation.

The correlation coefficients arrangement is the following::

$$\begin{aligned} r_{FTSE100} &< r_{CAC40} < r_{S\&P500} < r_{IBEX35} < r_{DAX30} < r_{ASE} < r_{PSI20} \\ \lambda_{IBEX35} &< \lambda_{FTSE100} < \lambda_{CAC40} < \lambda_{S\&P500} < \lambda_{ASE} < \lambda_{DAX30} < \lambda_{PSI20}. \end{aligned}$$

Besides the arrangement of the global and linear correlation coefficients, the

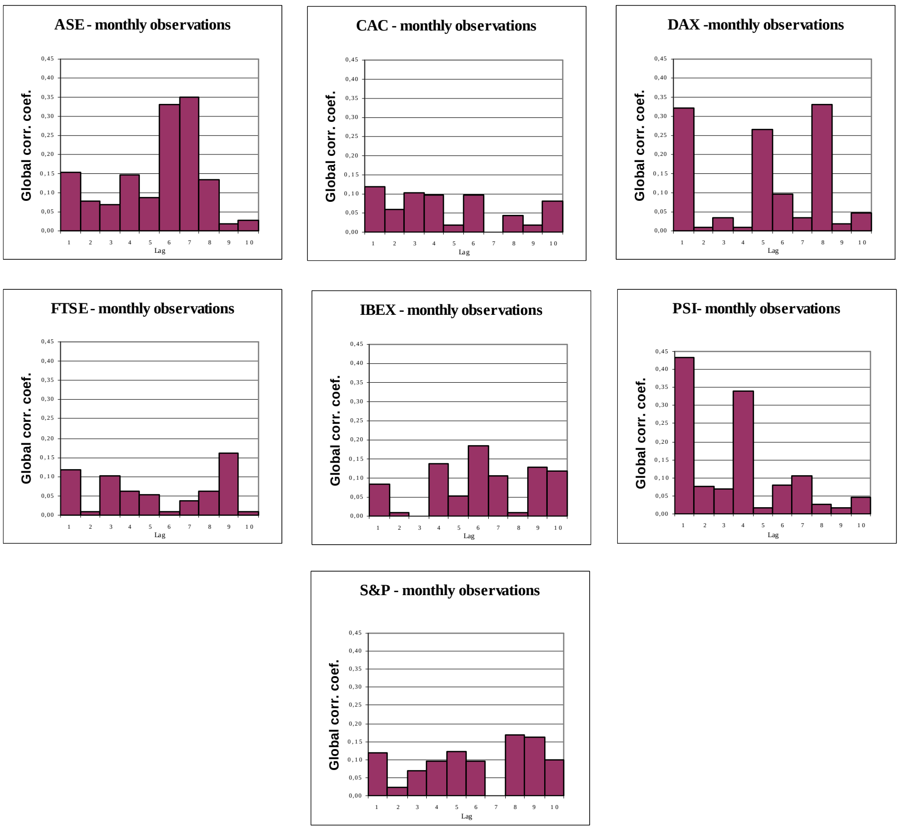

Figure 3: Global correlation coefficient for serial dependence in the stock market indexes monthly returns, for lags $k = 1, ..., 10$.

	$Mutual\ Inf.$	NMI	λ	r
lag: 1				
ASE	0.01168	0.02829	0.15195	0.23456
CAC 40	0.00705	0.01226	0.11833	0.15566
DAX 30	0.05494*	0.02715	0.32258	0.22988
FTSE 100	0.00705	0.01158	0.11833	0.15128
IBEX 35	0.00360	0.02215	0.0864	0.20816
PSI 20	0.10401**	0.05749	0.43338	0.32958
S&P 500	0.00705	0.02051	0.11836	0.20047

Table 14: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of monthly data relative to 1 lag.

analysis of Table 14 allows us to verify that for the majority of the indexes the normal mutual information presents highest values than those of the empirical mutual information. This situation is illustrated in the Figure 4.

These results seems to refute the theory, since the empirical mutual information will always be higher than the normal mutual information. An identical conclusion can be taken for the comparative analysis between the linear and global correlation coefficients. We presumed that such discrepancies are related with the small number of observations (119 observations) which leads to the undervaluation of the mutual information. >From the joint analysis of Tables 14, 15 and 16 we verify that PSI 20 index has maintained in a relatively constant way the values for the linear correlation coefficient as for the global correlation coefficient. Only in the case on which we analyze the global dependence relative to the first three lags, the empirical mutual information presents higher values than normal mutual information, as well as the global correlation coefficient presents much higher values than linear correlation coefficient. In addition to the enumerated property ($IM \geq IMN$), the additivity property is not also respected by the DAX 30 index, since $IM_1 = 0.05494$; $IM_{1,2} = 0.16112$ and $IM_{1,2,3} = 0.05101$.

We filtered the time series in study in order to eliminate the linear dependence, by the following processes $ARMA(1,0)$ for ASE, DAX 30, IBEX 35 and S&P 500 indexes and $ARMA(3,0)$ for PSI 20 index. We didn't apply any of those processes to the CAC 40 and FTSE 100 indexes, since none of this presented a statistically significant linear autocorrelation. *Ljung-Box* test was applied to the filtered time series, which confirms the inexistence of any type of serial linear dependence (Table 17).

The BDS test results lead us not to reject the null hypothesis of independence for the majority of the indexes, except the DAX 30 index which to distances of 0.5σ, 1σ and 1.5σ consider that we must reject the null hypothesis with a significance level of 1%. The information theory measures, applied to the filtered time series, point to the non existence of nonlinear dependence (Table 18), in accordance with BDS test results.

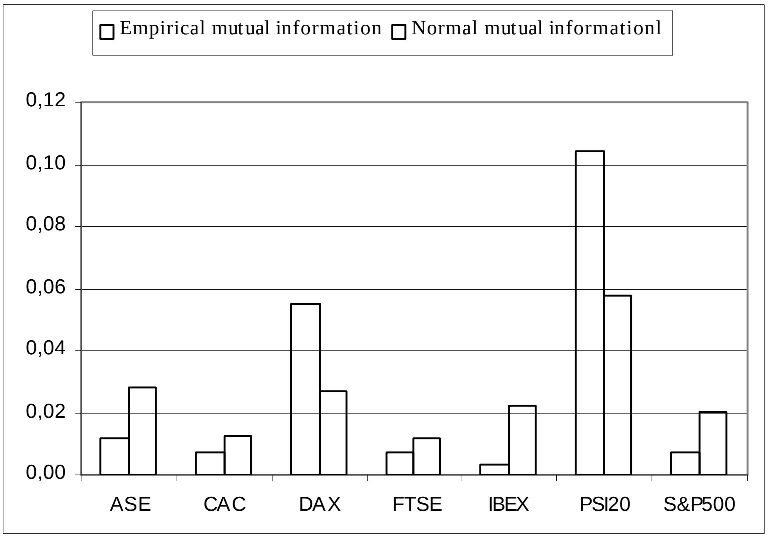

Figure 4: Empirical mutual information and normal mutual information to monthly data relative to lag 1.

We must underline that the filtered time series related to the ASE, PSI 20 and S&P 500 indexes, present positive values for the statistics, revealing the possible presence of nonlinear dependence, although not statistically significant.

4 Conclusions

On the basis of the presented results in this paper, we may conclude that the mutual information and the global correlation coefficient are efficient measures for testing and evaluating serial dependence, since they capture not only the linear dependence, but also the nonlinear dependence and this allows us to study with some detail nonlinear systems.

The estimation of mutual information for empirical data, with unknown theoretical probability distribution, must be based on one of the two enunciated algorithms in order to obtain an adequate space partition. Only a partition that allows for a uniform distribution (*algorithm A*) or local independence (*algorithm B*) can provide a correct computation of the mutual information.

According to Maasoumi and Racine (2002) and Granger and Lin (1994), the mutual information can also be used as a measure of predictability, through the calculation of the global correlation coefficient (λ), allowing for the construction of prediction models.

The empirical evidence presented by Granger and Lin (1994), Granger and Maasoumi (2000), Granger, Maasoumi and Racine (2002), Maasoumi and Racine

	Mutual Inf.	*NMI*	λ	r
lag: 1 e 2				
ASE	0.02815	0.05020	0.23396	0.24808
CAC 40	0.01976	0.03498	0.19687	0.15775
DAX 30	0.16112*	0.07454	0.52486	0.22906
FTSE 100	0.01681	0.02284	0.18185	0.16616
IBEX 35	0.01774	0.07775	0.18668	0.26678
PSI 20	0.10557	0.09052	0.43627	0.34529
S&P 500	0.01806	0.02005	0.18836	0.20073

Table 15: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of monthly data relative to 2 lags.

	Mutual Inf.	*NMI*	λ	r
lag: 1, 2 e 3				
ASE	0.05103	0.07805	0.31149	0.25184
CAC 40	0.06249	0.05451	0.34275	0.15289
DAX 30	0.05101	0.13315	0.31142	0.23370
FTSE 100	0.04197	0.03562	0.28376	0.17108
IBEX 35	0.06167	0.12165	0.34064	0.27088
PSI 20	0.09175	0.15803	0.40946	0.34683
S&P 500	0.04162	0.03768	0.28261	0.22220

Table 16: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of monthly data relative to 3 lags.

(2002), Bernhard and Darbellay (1999), Darbellay (1998a, 1998b, 1998c) and Darbellay and Wuertz (2000) demonstrates that these measures can capture the global relation established between the relevant variables, being this a linear or nonlinear relationship.

The results presented in this paper are in some way consistent with the results presented by the above mentioned authors, since we made a comparison between some stock market indexes, not only for serial linear dependence but also for serial nonlinear dependence.

Generally, we can say that there are certain indexes which present nonlinear dependence signals, namely the DAX 30, PSI 20 and S&P 500 indexes, for any periodicity of observations, leading us to conclude that the serial scale does not affect the conclusions obtained, except for the monthly data which did not allow us to make many conclusions, since there was a small number of observations for this type of analysis. In spite of some indexes evidence significant serial nonlinear dependence, that does not mean that these markets are not efficient.

The presence of serial dependence must provide some strategy to generate

	ASE	DAX 30	IBEX 35	PSI20	S&P 500
LBQ(10)	8,872	2,168	9,059	2,641	9,681
ρ_1	0,001	0,001	0,008	0,003	0,003
ρ_2	-0,016	0,011	-0,008	-0,001	-0,003
ρ_3	0,067	-0,045	0,027	0,034	0,086

Table 17: *Ljung-Box* test for the filtered monthly series.

	$Mutual\ Inf.$	NMI	λ	r
lag: 1				
ASE	0.00238	0.00000	0.06891	0.00100
DAX 30	0.00000	0.00000	0.00000	0.00100
IBEX 35	0.00000	0.00003	0.00000	0.00768
PSI 20	0.00004	0.00000	0.00877	0.00316
S&P 500	0.00033	0.00000	0.02571	0.00316

Table 18: Average mutual information, global correlation coefficient (λ), normal mutual information and linear correlation (r) coefficient of filtered monthly data.

systematic profits, for that market can be considered not efficient. In this case, the investors must found which function is behind the global and nonlinear dependence to use that to take decisions and make profits. Even if the investors found the function, that would not be meaning of direct profit, because transaction costs could simply eliminate those profits.

The IBEX 35, CAC 40 and FTSE 100 indexes present small significant levels of nonlinear dependence in all serial scales analyzed. The ASE index presented the highest discrepancy on the serial nonlinear dependence levels. For the daily data this was one of the indexes that presented the highest levels of nonlinear and global dependence, while for the weekly periodicity the values of the mutual information were nor statistically significant for the filtered time series.

The weekly observations present a slightly higher levels of dependence than the daily observations. That might be explained by the fact that those observations are weekly and monthly averages and not isolated values of a day in a week or month, and also because the daily movement can contain some bias resulting from slow dilutions of information in the markets. In what refers to monthly analysis, the results were inconclusive because the number of observations was not enough to apply mutual information in a conclusive way.

In spite of this, much of the global correlation in the weekly observations is supported by the linear correlation observed, which seems to indicate that periodicity is not the cause of the major or minor level of nonlinear dependence. Moreover, it is also apparent that the structure of the statistical dependence in the global level, as well as its degree, are not related to any specific market characteristics.

In the end, we can say that the main advantage of the application of the mutual information in financial time series is the fact that this measure captures the global serial dependence (linear and nonlinear) without any request about some theoretical probability distribution or specific model of dependency. Even if this dependence can not to refute the efficient market hypothesis, it is important to the investor to know that the rate of returns are not independent and identically distributed.

References

[1] Bernhard, H. and Darbellay, G. (1999). Performance Analysis of the Mutual Information Function for Nonlinear and Linear Processing, Acts: *IEEE International Conference on Acoustics, Speeche and Signal Processing (USA),3*, 1297-1300.

[2] Bonanno, G., Lillo, F. and Mantegna, R. (2001). Levels of Complexity in Financial Markets, *Physica A*, 299, 16-27.

[3] Curto, J.D. (2003). Distribuições de Pareto Estável: Aplicação aos Índices PSI 20, DAX 30 e DJIA; in *Temas em Métodos Quantitativos 3,* E. Reis e M. Hill eds., Edições Sílabo.

[4] Darbellay, G. and Vadja, I. (1999). Estimation of the Information by an Adaptative Partitioning of the Observation Space, *IEEE Transactions on Information Theory*, 45, May, 1315-1321.

[5] Darbellay, G. (1998*a*). An Adaptative Histogram Estimator for the Mutual Information, *UTIA Research Report, n.º 1889*, Acad. Sc., Prague.

[6] Darbellay, G. (1998*b*). Predictability: an Information-Theoretic Perspective, *Signal Analysis and Prediction* ed. A. Procházka, J. Uhlír, P.J.W. Rayner e N.G. Kingsbury; Birkhauser, Boston, 249-262.

[7] Darbellay, G. (1998*c*). Statistical Dependences in R^d: an Informatio-Theoretic Approach, Actas *3rd IEEE European Workshop on Computer-intensive Methods in Control and data Processing*, (Prague), 83-88.

[8] Darbellay, G. (1999). An Estimator of the Mutual Information Based on a Criterion for Independence, *Computational Statistics and Data Analysis, 32, 1-17.*

[9] Darbellay, G. and Wuertz, D. (2000). The Entropy as a Tool for Analysing Statistical Dependence's in Financial Time Series, *Physica A,* 287, 429-439.

[10] Darbellay, G. and Franek, J. software: http://siprint.utia.cas.cz/timeseries/.

[11] Díaz, A., Grau-Carles, P and Mangas, L. (2002). Nonlinearities in the Exchange Rate Returns and Volatility; *Physica A*, 316, 469-482.

[12] Dobrushin, R. (1959). General Formulation of Shannon's Main Theorem in Information Theory; *Uspekhi Mat. Nauk*, 14, 3-104.

[13] Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, 25, 5, May, 383-417.

[14] Fama, E. (1991). Efficient Capital Markets: II, *Journal of Finance*, 46, 5, December, 1575-1617.

[15] Fieldman, D. (1998). *A Brief Introduction to: Information Theory, Excess Entropy and Computational Mechanics*; preprint, University of California.

[16] Garner, W. and Mcgill, W. (1956). The Relation Between Information and Variance Analysis, *Psychometrica,* 219-228.

[17] Granger, C. and Lin, J. (1994). Using the Mutual Information Coefficiente to Identify Lags in Nonlinear Models, *Journal of Time Series Analysis*, 15, 4, 371-384.

[18] Granger, C. and Maasoumi, E. (2000). A Dependence Metric for Nonlinear Time Series; preprint in http://ideas.repec.org/p/ecm/wc2000/0421.html.

[19] Granger, C, Maasoumi, E. and Racine, J. (2002). A Dependence Metric for Possibly Nonlinear Processes, UCSD *Working Paper.*

[20] Hsieh, D. (1989). Testing for Nonlinear Dependence in Foreign Exchange rates: 1974-1983, *Journal of Business*, 62, 339-368.

[21] Klan, P. and Darbellay, G. (1997). An Information-Theoretic Adaptative Method for Time Series Forecasting, *Neural Network*, 2, 227-238.

[22] Kraskov, A., Stögbauer, H. and Grassberger, P. (2003). Estimating Mutual Information, preprint in http://www.arxiv:cond-mat/0305641.

[23] Kullback, S. (1968). *Information Theory and Statistics,* Dover, New York.

[24] Maasoumi, E. (1993). A Compendium to Information Theory in Economics and Econometrics, *Econometric Reviews*, 12 (2), 137-181.

[25] Maasoumi, E. and Racine, J. (2002). Entropy and Predictability of Stock Market Returns, *Journal of Econometrics*, 107, 291-312.

[26] Mantegna, R., Palágyi, Z. and Stanley, H. (1999). Application of Statistical Mechanics to Finance, *Physica A*, 274, 216-221.

[27] Moddemeijer, R. (1999). A Statistic to Estimate the Variance of the Histogram-Based Mutual Information Estimator on Dependent Pairs of Observations, *Signal Processing*, 75, 51-63.

[28] Paninski, L. (2003). Estimation of Entropy and Mutual Information, *Neural Computation*, 15, 1191-1253.

[29] Pesaran, M. and Timmermann, A. (1995). Predictability of Stock Returns: Robustness and Economic Significance, *Journal of Finance*, 50, 1201-1228.

[30] Peters, E. (1996). *Chaos and Order in the Capital Markets*, Second Edition; Wiley Finance Edition, New York.

[31] Perez A. (1957). Notion Généralisées d'Incertitude, d'Entropie et d'Information du Point de Vue de la Théorie des Martingales, *Transactions of the 1st Prague Conference in Information Theory, Statistical Decision Functions, Random Processes*, 183-208.

[32] Pompe, B (1998). Ranking Entropy Estimation in NonLinear Time Series Ananlysis, preprint in: H. Kantz, J. Kurths and G. Mayer-Kress eds., *Non-Linear Analysis of Physiological Data*; Springer, Berlin.

[33] Qi, M. (1999). Nonlinear predictability of Stock Returns using Financial and Economic Variables, *Journal of Business and Economic Statistics*, 17, 4, 419-429.

[34] Shannon, C. (1948). A Mathematical Theory of Communication ($1and2$), *Bell Systems Tech*; 27; 379-423 e 623-656.

[35] Soares, J. (1994). *Preços de Acções na Bolsa de Lisboa: Análise, Previsão e Regras de Compra e Venda*, PhD Thesis, IST, Lisbon.

[36] Soofi, E. (1997). Information Theoretic Regression Methods, Thomas Fomby and R. Carter Hill eds., *Advances in Econometrics - Applying Maximum Entropy to Econometric Problems*; Vol. 12.

[37] Stengos, T. (1995). Nonparametric Forecasts of Gold rate of Return, in A.K. Barnett e M. Salmon eds., *Nonlinear Dynamics in Economics*, Cambridge University Press.

[38] Tambakis, D. (2000). On the Informational Content of Asset Prices, preprint in http://ideas.repec.org/p/sce/scecf0/101.html

[39] Urbach, R. (2000). *Footprints of Chaos in the Markets - Analysing non-linear time series in financial markets and other real systems*, Prentice Hall, London.

5 Appendix A

Critical values tables for testing serial independence through mutual information fot $N(0,1)$ data. 5000 replications were computed.

N=100				*N=200*			
	Percentiles				*Percentiles*		
Lag	*90*	*95*	*99*	*Lag*	*90*	*95*	*99*
1	0.0185	0.0323	0.0679	1	0.0092	0.0214	0.0361
2	0.1029	0.1232	0.1933	2	0.0561	0.0701	0.1080
3	0.1059	0.1260	0.1722	3	0.0591	0.0918	0.1318

N=500

	Percentiles		
Lag	*90*	*95*	*99*
1	0.0037	0.0070	0.0144
2	0.0222	0.0369	0.0501
3	0.06799	0.0788	0.1128

N=1000

	Percentiles		
Lag	*90*	*95*	*99*
1	0.0019	0.0045	0.0071
2	0.0133	0.0191	0.0311
3	0.0340	0.0399	0.0568

N=1500

	Percentiles		
Lag	*90*	*95*	*99*
1	0.0013	0.0026	0.0045
2	0.0101	0.0133	0.0224
3	0.0222	0.0267	0.0369

N=2000

	Percentiles		
Lag	*90*	*95*	*99*
1	0.0009	0.0019	0.0033
2	0.0061	0.0094	0.0147
3	0.0169	0.0203	0.0278

N=2500

	Percentiles		
Lag	*90*	*95*	*99*
1	0.0008	0.0015	0.0030
2	0.0054	0.0078	0.0129
3	0.0134	0.0171	0.0251

6 Appendix B

Index	m	*E*			
		0.5σ	1σ	1.5σ	2σ
ASE	2	9.410**	9.480**	9.098**	8.292**
	3	12.087**	11.752**	10.880**	9.615**
	4	14.609**	13.910**	12.742**	11.689**
	5	17.963**	15.629**	13.845**	12.402**
CAC 40	2	4.200**	4.974**	6.707**	8.595**
	3	5.940**	7.384**	9.753**	11.887**
	4	7.325**	9.067**	11.570**	14.092**
	5	8.398**	10.247**	12.704**	15.327**
Dax 30	2	8.692**	10.104**	10.848**	11.257**
	3	12.440**	14.008**	14.825**	14.950**
	4	15.183**	16.710**	17.327**	17.180**
	5	18.322**	19.712**	19.267**	19.079**
FTSE 100	2	5.998**	7.177**	8.297**	8.866**
	3	8.371**	9.722**	11.018**	11.845**
	4	10.120**	11.771**	13.140**	14.028**
	5	12.052**	13.488**	14.818**	15.844**
IBEX 35	2	4.578**	5.520**	6.229**	7.073**
	3	6.866**	7.163**	8.564**	9.318**
	4	9.271**	10.165**	10.750**	11.279**
	5	11.213**	12.103**	12.587**	13.036**
PSI 20	2	9.410**	9.480**	9.098**	8.292**
	3	12.087**	11.752**	10.880**	9.615**
	4	14.609**	13.910**	12.742**	11.389**
	5	17.963**	15.629**	13.845**	12.402**
S&P 500	2	8.639**	8.698**	9.230**	10.098**
	3	13.308**	13.232**	13.040**	13.197**
	4	17.491**	16.272**	15.161**	14.843**
	5	23.349**	20.011**	17.536**	16.594**

Table 19: BDS test for the filtered daily data. m is the embedding dimension and E is the distance between points measured in terms of number of standard deviations of raw data.

Index	m	E			
		0.5σ	1σ	1.5σ	2σ
ASE	2	5.284**	5.038**	4.740**	3.761**
	3	6.472**	6.371**	6.212**	5.162**
	4	7.961**	7.750**	6.968**	5.609**
	5	9.339**	8.904**	7.563**	5.971**
CAC 40	2	4.517**	4.421**	4.543**	4.505**
	3	4.447**	4.369**	4.678**	4.655**
	4	5.574**	4.657**	4.844**	4.808**
	5	6.051**	4.622**	4.866**	4.879**
Dax 30	2	3.804**	4.399**	4.983**	5.170**
	3	6.879**	6.538**	6.678**	6.345**
	4	11.249**	8.672**	8.162**	7.154**
	5	15.005**	10.257**	9.147**	7.736**
FTSE 100	2	1.491	1.749	2.292*	2.953**
	3	1.386	2.271*	2.731*	3.454**
	4	1.780	3.020*	3.460**	4.040**
	5	2.962**	3.621**	4.162**	4.606**
IBEX 35	2	1.426	1.322	1.546	1.547
	3	1.650	1.748	1.911	1.920
	4	1.841	2.585**	2.697**	2.569*
	5	1.857	2.952**	2.843**	2.541*
PSI 20	2	5.143**	4.875**	4.060**	3.496**
	3	5.776**	5.781**	4.965**	4.354**
	4	6.596**	6.916**	5.961**	5.013**
	5	7.331**	7.937**	6.645**	5.315**
S&P 500	2	4.672**	3.837**	2.848**	2.429**
	3	6.434**	4.919**	3.547**	3.174**
	4	8.553**	6.607**	4.778**	4.066**
	5	10.645**	7.920**	5.605**	4.664**

Table 20: BDS test for the filtered weekly data. m is the embedding dimension and E is the distance between points measured in terms of number of standard deviations of raw data.

Index	m	E			
		0.5σ	1σ	1.5σ	2σ
ASE	2	2.593**	0.576	0.753	0.178
	3	1.856	1.419	1.175	0.381
	4	0.593	1.798	1.268	0.386
	5	-0.388	1.994*	1.480	0.283
CAC 40	2	0.270	1.548	0.867	0.743
	3	-0.503	0.863	0.101	-0.069
	4	-0.879	1.204	0.618	0.876
	5	-1.647	1.107	0.184	-0.118
Dax 30	2	3.939**	4.226**	3.061**	1.531
	3	4.776**	3.511**	2.164**	0.175
	4	8.332**	4.666**	2.706**	0.925
	5	11.711**	6.090**	2.342*	-0.265
FTSE 100	2	1.807	0.429	0.695	0.641
	3	1.567	0.685	1.006	0.510
	4	1.435	1.666	1.811	1.089
	5	2.725**	2.101*	1.625	0.334
IBEX 35	2	2.381*	2.504*	2.426*	2.203*
	3	3.505**	1.369	1.522	0.430
	4	3.529**	1.937	2.147*	0.853
	5	4.486**	2.147*	1.833	0.486
PSI 20	2	0.548	0.291	1.125	1.101
	3	0.443	-0.011	0.453	-0.433
	4	1.117	0.478	0.873	0.182
	5	0.793	1.828	1.977*	1.374
S&P 500	2	1.289	0.866	0.321	0.048
	3	2.165*	1.618	0.849	-0.072
	4	3.706**	3.081**	1.842	0.849
	5	4.916**	3.289**	1.737	0.765

Table 21: BDS test for the filtered monthly data. m is the embedding dimension and E is the distance between points measured in terms of number of standard deviations of raw data.