Uncertainty analysis in financial markets: can entropy be a solution?

Andreia Dionísio*, Rui Menezes** and Diana A. Mendes**
*Universidade de Evora
Dep. Gestão de Empresas, Largo Colegiais, 2, 7000 Evora, Portugal
E-mail: andreia@uevora.pt
**ISCTE
Departamento de Metodos Quantitativos, Av. Forcas Armadas, 1649-Lisboa, Portugal
E-mail: rui.menezes@iscte.pt, diana.mendes@iscte.pt

Version: August, 2005

The investor is not only interested in the expected return, he is also concerned with the risk assumed by the investment and the uncertainty. One of the concepts more used to measure the risk and the uncertainty is the variance and/or the standard-deviation. We pose the follow question: *is the standard-deviation a good measure of risk and uncertainty? Which are the potentialities of the entropy in this context? Can entropy present some advantages as a measure of uncertainty and simultaneously verify some base assumptions of the portfolio management, namely the effect of diversification?*

Key Words: Uncertainty, risk, entropy, mutual information, stock market.

1. INTRODUCTION

This paper examines the adequacy of entropy as a measure of uncertainty in portfolio management and its behaviour is compared with the most traditional risk measure used in finance: the variance.

The notion of "risk" and "uncertainty" in economics and the distinction between these concepts was preconised by Knight (1921). According to this author, risk and uncertainty are associated with the imperfect knowledge, but there is a conceptual difference between them. In Knight's interpretation, "risk" refers to situations where the decision-maker can assign mathematical probabilities to the randomness which he is faced with. In contrast, Knight's "uncertainty" refers to situations when randomness cannot be expressed in terms of specific mathematical probabilities. This idea was reinforced by Keynes (1937), who expressed the following sentence:

"*By 'uncertain' knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty...The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence...About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.*" [Keynes (1937)]

In this paper, the notion of uncertainty is related with the more or less difficulty to predict the future. It is quite common to relate the variance or the standard-deviation and the VaR (Value-at-Risk) as the main measures of risk and uncertainty in finance. However, some authors [e.g. Soofi (1997)] alert for the fact that these measures may fail in some specific situations as a measure of uncertainty, since they require that the underlying probability distribution is symmetric and neglect the possibility of extreme events such as the existence of fat-tails.

The main goal of this paper is to assess the ability of entropy as a measure of uncertainty on portfolio management. In this way, we must examine whether entropy is sensitive to the effect of diversification. The risk of a portfolio can be splitted into specific risk and systematic risk (not diversifiable). Using the concept of entropy we can provide similar information, since the mutual information between two random variables is comparable to the systematic risk.

The rest of the paper is organized as follows. In Section 2 we present the theoretical background about entropy and its mathematical properties. Section 3 presents a comparative analysis between the entropy and the standard-deviation as measures of uncertainty in the stock market. In Section 4 we analyse the relationship between the Portuguese stock index PSI 20 and the individual stocks used in our study through the CAPM in order to isolate the systematic risk and the specific risk, and we compare the results with some measures of the information theory, namely entropy and the mutual information. Finally, in Section 5 we present the main conclusions of this paper.

2. THEORETICAL BACKGROUND

Suppose that we have a set of possible events whose probabilities of occurrence are $p_1, ..., p_n$ and μ is a measure of uncertainty. According to Shannon (1948), a good measure of uncertainty $\mu = \mu(p_1, ..., p_n)$ should satisfy the following properties:

1. μ should be continuous in $p_i, i = 1, ..., n$;
2. If $p_i = 1/n$, then μ should be a monotonic increasing function of n;
3. μ is maximized in a uniform probability distribution context;
4. μ should be additive;
5. μ should be the weighted sum of the individual values of μ (see Figure 1)

We require, as in this special case that

$$\mu\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) = \mu\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}\mu\left(\frac{2}{3}, \frac{1}{3}\right). \tag{1}$$

According to Shannon (1948) a measure that satisfies these properties is the entropy which is defined as $H(X) = -\sum_i p_i \log p_i$. When the random variable has a continuous distribution, and $p_X(x)$ is the density function of the random variable X, the entropy is given by

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

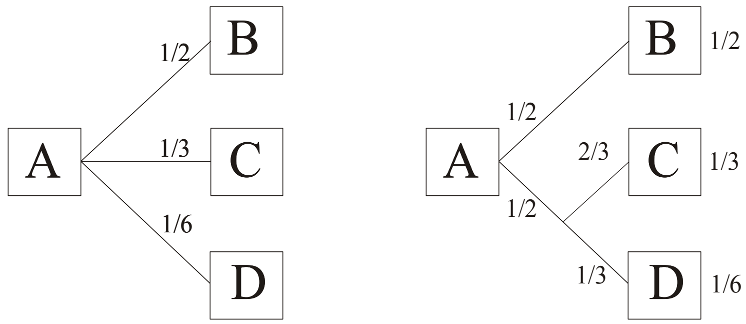

FIG. 1 Decomposition of a choice from three possibilities. Source: Shannon, C. E. (1948). A Mathematical Theory of Communication, Bell Systems Tech. (1 e 2), 27: 379-423, 623-656.

Arafat, Skubic and Keegan (2003) consider that a measure of uncertainty should attend the following properties: (i) *Symmetry*, that is $H(X) = H(-X)$; and (ii) *Valuation*: $H(X \cup Y) + H(X \cap Y) = H(X) + H(Y)$. These authors discuss combined methods of uncertainty and conclude that entropy can be a good measure of uncertainty.

If we have two arguments $X \in \overrightarrow{X}$ and $Y \in \overrightarrow{Y}$ and $p_{X,Y}(x,y)$ is the probability of the joint occurrence, the joint and conditional entropies are given by

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

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

$$H(X|Y) = -\int\int p_{X,Y}(x,y)\log\frac{p_{X,Y}(x,y)}{p_Y(y)}dxdy. \tag{5}$$

The entropies of continuous distributions have most of the properties of the discrete case. In particular we have the following [Shannon (1948); Kraskov *et al.* (2004)]:

(a) If X is limited to a certain volume v in its space, then $H(X)$ is a maximum and is equal to $\log v$ when $p_X(x)$ is constant, $1/v$, in the volume;

(b) For any two variables X and Y, we have

$$H(X,Y) \leq H(X) + H(Y) \tag{6}$$

where the equality holds if (and only if) X and Y are statistically independent, i.e. $p_{X,Y}(x,y) = p_X(x)p_Y(y)$;

(c) The joint entropy can be given by

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y) \tag{7}$$

and

$$H(X) + H(Y) \geq H(X,Y) \tag{8}$$

meaning that $H(Y) \geq H(Y|X)$ and $H(X) \geq H(X|Y)$.

The assumption that the data and the residuals follow a normal distribution is very common in portfolio management and regression analysis. Thus, the equation used to estimate parametrically the entropy of a normal distribution, $NH(X)$, is

$$\begin{aligned} NH(X) &= \int p_X(x) \log \sqrt{2\pi}\sigma dx + \int p_X(x) \frac{(x-\bar{x})^2}{2\sigma^2} dx \\ &= \log\left(\sqrt{2\pi e}\sigma\right). \end{aligned} \tag{9}$$

where σ is the standard-deviation and let $H(X)$ and $H(Y)$ be the entropies of the random variables $X \in \overrightarrow{X}$ and $Y \in \overrightarrow{Y}$, $H(X,Y)$, is the joint entropy and $H(Y|X)$ and $H(X|Y)$ are the conditional entropies, then the mutual information can be defined as follows [Shannon (1948)]

$$\begin{aligned} I(X,Y) &= H(Y) - H(Y|X) \\ &= H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y) \\ &= \int\int p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)\, p_Y(y)} dxdy. \end{aligned} \tag{10}$$

The mutual information is a nonnegative measure [Kullback (1968)], being equal to zero if and only if X and Y are statistically independent. In this way, the mutual information between two random variables X and Y can be regarded as a measure of dependence between these variables, or even better, it can be regarded as a measure of the statistical correlation between X and Y. However, we can not say that X is causing Y or vice-versa.

In order to normalize the measure we need to define a measure that can be directly comparable with the linear correlation coefficient. In equation (10), we have $0 \leq I(X,Y) \leq +\infty$, which hampers eventual comparisons between different samples. However, we can easily compare the mutual information with the covariance, since both vary between 0 and $+\infty$. Some authors, namely Granger and Lin (1994), Darbellay (1998) and Soofi (1997) have used a standard measure for the mutual information, referred to as the global correlation coefficient, defined by

$$\lambda(X,Y) = \sqrt{1 - e^{-2I(X,Y)}}. \tag{11}$$

This measure varies between 0 and 1 being thus directly comparable with the linear correlation coefficient.

The function $\lambda(X,Y)$ captures the overall linear and nonlinear dependence between X and Y. This measure can be regarded as a measure of predictability based on an empirical probability distribution, although it does not depend on any particular model of predictability. In this particular case, the above mentioned properties assume the following form:

(i) $\lambda(X,Y) = 0$, if and only if X contains no information on Y;

(ii) $\lambda(X,Y) = 1$, if there is a perfect relationship between the vectors X and Y. This is the farthest case of determinism;

(iii) When modelling the input-output pair (X,Y), by any model with input X and output $U = f(X)$, where f is a function of X, the predictability of Y by U cannot exceed the predictability of Y by X, i.e., $\lambda(X,Y) \geq \lambda(U,Y)$.

One of the difficulties to estimate the mutual information on the basis of empirical data lies on the fact that the underlying *pdf* is unknown. To overcome this problem, there are essentially three different methods to estimate mutual information: histogram-based estimators, kernel-based estimators and parametric methods.[1] In order to minimize the bias that may occur, we will use the marginal equiquantization estimation process, proposed by Darbellay (1998).[2]

The introduction of entropy as a measure of uncertainty in finance goes back to Philippatos and Wilson (1972), which present a comparative analysis between the behaviour of the standard-deviation and the entropy on portfolio management. These authors conclude that entropy is more general and has some advantages facing to the standard-deviation. According to Lawrence (1999) the two main measures of uncertainty are the entropy and the variance, because the entropy is a concave function allows its use as an uncertainty function.

3. ENTROPY AND VARIANCE: COMPARATIVE ANALYSIS

The variance is a measure of dispersion and its simplicity remains its major attraction. Historically, the variance has had a primordial role in the analysis of uncertainty and risk. However, according to Maasoumi (1993), entropy can be an alternative measure of dispersion and Soofi (1997) considers that the interpretation of the variance as a measure of uncertainty must be done with some precaution.

The entropy is a measure of disparity of the density $p_X(x)$ from the uniform distribution. It measures uncertainty in the sense of the "utility" of using $p_X(x)$ rather than the uniform distribution. The variance measures an average of distances of outcomes of the probability distribution from the mean. According to Ebrahimi, Maasoumi and Soofi (1999), both these measures reflect concentration but their metrics for concentration are different. Unlike the variance which measures concentration only around the mean, the entropy measures diffuseness of the density irrespective of the location of concentration.

Ebrahimi, Maasoumi and Soofi (1999) examined the role of variance and entropy in ordering distributions and random prospects, and conclude that there is no universal relationship between these measures in terms of ordering distributions. These authors found that, under certain conditions, the order of the variance and entropy is similar when continuous variables are transformed and show (using a Legendre series expansion) that the entropy depends on many more parameters of a distribution than the variance. The Legendre series expansion reveals that entropy may be related to high-order moments of a distribution which, unlike the variance, could offer a much closer characterization of $p_X(x)$ since it uses much more information about the probability distribution than the variance [see Ebrahimi *et al.* (1999)].

Maasoumi and Racine (2002) argue that in the case that the empirical probability distribution is not perfectly known, the entropy constitutes an alternative

[1] According to Moddemeijer (1999), histogram-based estimators are divided in two groups: equidistant cells and equiprobable cells, i.e. marginal equiquantisation [see e.g.Darbellay (1998)]. The second approach presents some advantages, since it allows for a better adequacy to the data and maximizes mutual information [Darbellay (199a)]. 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 and bias caused by the finite histogram.

[2] A good description of the estimation method based on the marginal equiquantization can be found in Darbellay (1998).

measure for the uncertainty, predictability and goodness-of-fit. Such statement is justified by the fact that entropy is a function of many moments of the distribution, so it is more general than the traditional methods based on the variance. In this context, McCauley (2003) defends that entropy represents the disorder and uncertainty of a stock market or a particular stock, since the entropy has the ability to capture the complexity of the systems, without requiring rigid assumptions that can bias the results obtained.

It is important to note some properties of the variance (and standard-deviation) and entropy as measures of uncertainty. The standard-deviation is a convex function, which according to the Jensen inequality $E\left[\sigma\left(X\right)\right] \geq \sigma\left[\left(EX\right)\right]$.[3] This property allows the variance and the standard-deviation to be used as risk measures in stock portfolios, since they take into account the effect of diversification.

The entropy, on the other hand, is a concave function and has a maximum for most of the probability distributions, and this fact leads us to presume that entropy will not satisfy the effect of diversification. However, we must note that entropy is not a function of the values of the variables but of the probability itself and the property $H\left(X,Y\right) \leq H\left(X\right) + H\left(Y\right)$[4] can bring some hope in this field.

We use the daily closing prices of 23 stocks rated on the Portuguese stock market (*Euronext Lisbon*), spanning the period from 28/06/1995 to 30/12/2002, which corresponds to 1858 observations *per* stock, in order to compute the rates of return. In this paper we use the stock index PSI 20 as the market benchmark or the proxy, since it is the stock index that better represents the Euronext Lisbon. The statistical analysis of these time series (rates of return) revealed that we must reject that the empirical distributions are normal, since they show high levels of kurtosis and skewness.

Firstly, we perform a comparative analysis between the entropy and the standard-deviation for each of the stocks in our data set and for the stock index PSI 20. The entropy was estimated using the equation (2), measured in *nats*.

Table 1 presents results obtained for the empirical entropy (H), normal entropy (NH) and standard-deviation (σ). We must emphasize the apparent positive correlation between the entropy and the standard-deviation, revealing that possibly the two measures assume a similar behaviour.

In order to provide a correct comparison between the entropy and the standard-deviation, we calculate the logarithm of the standard-deviation, $\ln\left(\sigma\right)$.[5] The more the empirical distribution is close to the Gaussian curve, the relationship between the entropy and $\ln\left(\sigma\right)$ is stronger and close to linear. Figure 2 presents the results of this relationship.

As we can see from Figure 2, there is a strong and positive relationship between the entropy and the $\ln\left(\sigma\right)$, although we note some deviations. The null hypothesis of the Jarque-Bera test that the rate of returns of each stock is normally distributed was rejected for all stocks and for the stock index PSI 20. According to the statistical analysis made, we can see that the stocks in Figure 2 show a higher deviation from the regression line, namely Salvador Caetano, Barbosa & Almeida, Banco Totta and Cin; are the ones that exhibit the higher levels of skewness and kurtosis. The stocks Portucel, PT e Sonae show smaller levels of skewness and kurtosis than the others, and the probability distribution is more close to the Gaussian curve.

[3] The equality occurs when the linear correlation coefficient between the variables is 1.

[4] The equality remains valid when the variables are statistically independent.

[5] Since $NH\left(X\right) = \ln\sqrt{2\pi e} + \ln\sigma$.

Stocks	H *(nats)*	σ (%)	NH *(nats)*
Barb. & Almeida	2.2530	4.5397	2.9318
BANIF	1.8217	1.9327	2.0779
BCP	1.7240	1.6257	1.9049
B. Totta	1.7792	1.9756	2.0998
BES	1.5883	1.4894	1.8173
Caima	2.2077	3.7044	2.7285
CIN	2.0697	3.1189	2.5564
Corticeira Amorim	1.8458	1.8449	2.0314
Estoril-Sol	2.0356	2.6695	2.4008
Fisipe	2.0733	3.1787	2.5754
Inapa	1.6769	1.6787	1.9370
Mod. & Continente	1.8880	1.8904	2.0557
Jer. Martins	2.0514	2.1816	2.1990
Mota-Engil	1.7870	1.9649	2.0944
Pap. Fernandes	2.4249	3.4922	2.6695
Portucel	1.9199	1.8392	2.0283
PT	2.1190	2.1473	2.1831
Salv. Caetano	2.6430	8.2134	3.5247
Soares Costa	2.1051	2.4329	2.3080
Somague	2.0646	2.3727	2.2830
Sonae	2.0540	2.0782	2.1504
Sonae-Ind.	1.9585	2.0813	2.1519
Tertir	2.2261	2.9266	2.4928
PSI 20	1.5059	1.1942	1.5964

TABLE 1
Entropy (H), Standard-deviation (σ) and normal entropy (NH) for all the stocks and the index PSI 20.

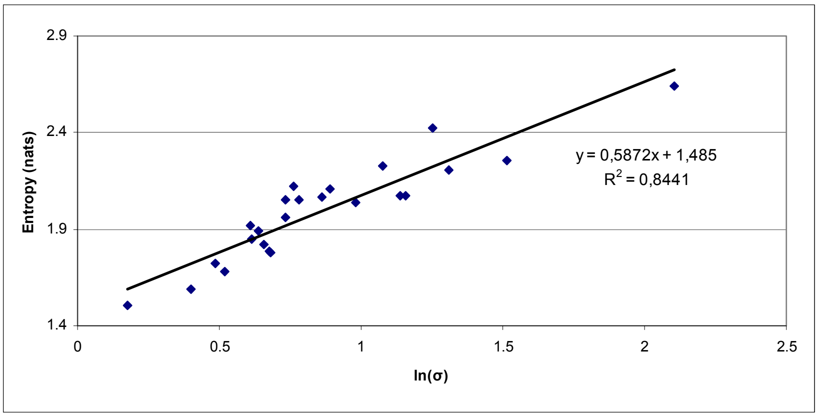

FIG. 2 Entropy *versus* $\ln(\sigma)$.

We perform a comparative analysis between the empirical entropy and the normal entropy for each stock (Figure 3).

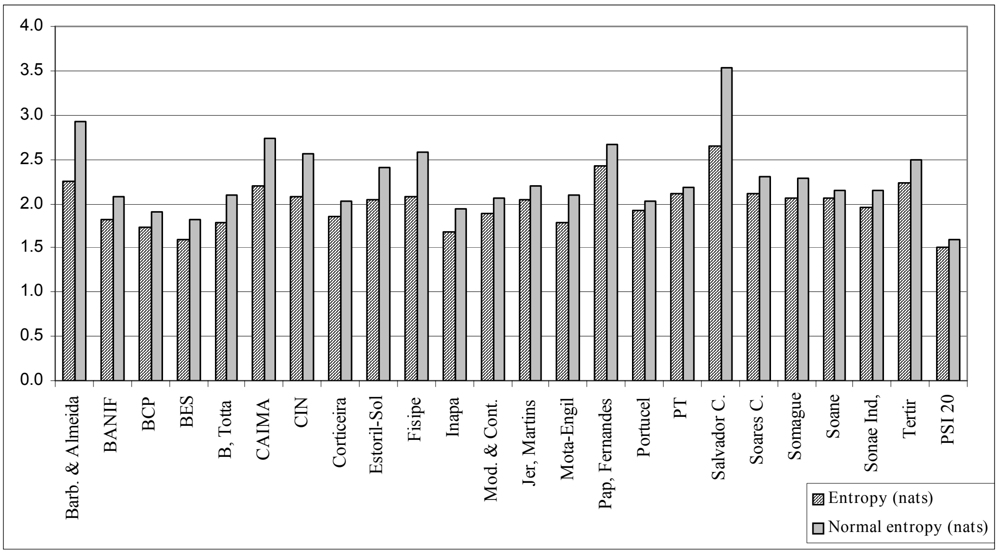

FIG. 3 Entropy (H) and normal entropy (NH) for all the stocks and the stock index PSI 20.

The normal entropy is always higher than the empirical entropy, which indicate that the uncertainty of these stocks and index is smaller than that we would observe if they were normally distributed. Thus, we can admit the evidence of some predictability of the rate of returns, or at least that it is higher than the one assumed by the financial theory. The main differences between the normal entropy and the empirical entropy are found in the stocks Barbosa & Almeida, Caima, Cin and Salvador Caetano which are the stocks that exhibit the highest levels of kurtosis, skewness, autocorrelation and heteroskedasticity. On the other way, the normal entropy and the empirical entropy are almost equal for the stocks Portucel, PT and Sonae. As we can see, the empirical entropy is sensitive to high-order moments and provide a more general information about the stock and its probability distribution.

In order to analyse if the effect of diversification is assumed by the entropy, we perform a similar analysis to that presented by Elton and Gruber (1995).[6] For randomly selected portfolios, our results (see Table 2 and Figure 4) show that the entropy and the standard-deviation tend to decrease when we include one more asset in the portfolio. This fact allows us to conclude that entropy is sensitive to the effect of diversification. These results can be explained by the fact that when the number of assets in the portfolio increases, the number of possible states of the system (portfolio) declines progressively and the uncertainty about that portfolio tends to fall. Besides, we verify that the entropy respect the condition of subadditivity suggested by Reesor and McLeish (2002), where $H[\theta X] + H[(1-\theta)Y] \geq H[\theta X +$

[6] These authors demonstrated that the diversification is a factor of minimization of the specific risk (measured by the standard-deviation). They made a random selection of the assets to compose portfolios and the only premise is the fact that the proportion invested in each asset is $\frac{1}{N}$, being N the number of assets in the portfolio.

$(1-\theta)Y]$, being $\theta = \frac{1}{N}$.

Number of stocks	*Porfolio*	H *(nats)*	σ (%)	NH *(nats)*
2	A	1,7427	1,5023	1,8260
3	B	1,8868	1,9551	2,0894
4	C	1,6587	1,3675	1,7319
5	D	1,5753	1,2256	1,6223
6	E	1,6362	1,8447	2,0312
7	F	1,4268	1,1424	1,5520
8	G	1,3940	1,1397	1,5497
9	H	1,4010	1,3502	1,7192
10	I	1,2761	0,9627	1,3809
11	J	1,4099	1,2504	1,6424
12	K	1,3387	1,1735	1,5789
13	L	1,2854	1,0209	1,4397
14	M	1,3239	1,1597	1,5671
15	N	1,3739	1,1623	1,5694
16	O	1,3150	1,0808	1,4967
17	P	1,2570	1,0413	1,4594
18	Q	1,2525	1,0268	1,4454
19	R	1,2469	1,0198	1,4385
20	S	1,2389	0,9874	1,4062
21	T	1,2362	0,9844	1,4033
22	U	1,2459	0,9801	1,3989
23	V	1,2292	0,9762	1,3949

TABLE 2
Entropy (H), standard-deviation (σ) and normal entropy (NH) of the randomly selected portfolios, where $\theta = 1/N$.

We must highlight the fact that, in this example, the normal entropy assumes always higher values than the empirical entropy. This means that the predictability level of each portfolio is higher than the one assumed by the normal distribution.

We can conclude that entropy observes the effect of diversification and is a more general uncertainty measure than the variance, since it uses much more information about the probability distribution. However, the use of the entropy in risk analysis and portfolio selection needs some care because it does not take into account the actual values of the variables and this fact can compromise its inclusion in a utility function.

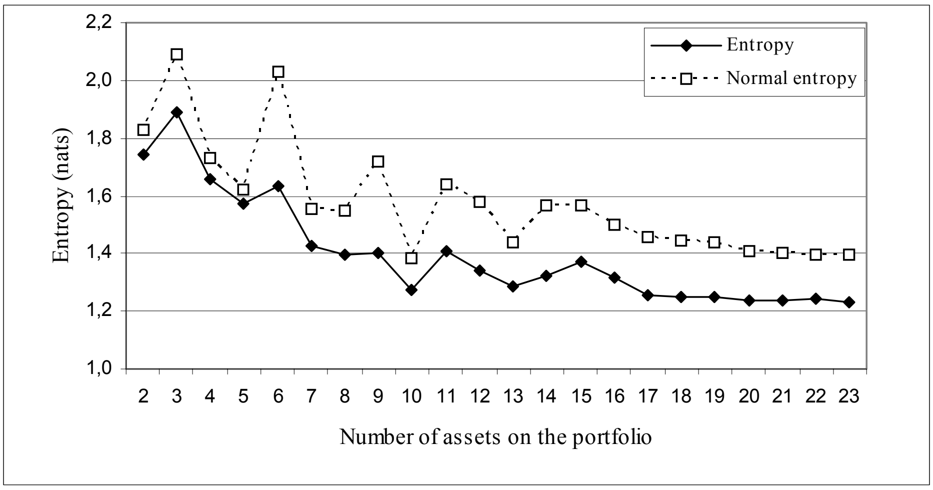

FIG. 4 Comparative analysis of the empirical entropy (H) and the normal entropy (NH) for portfolios randomly selected. Entropy is measured in *nats* because we use natural logarithms.

4. DEPENDENCY ANALYSIS BETWEEN THE STOCKS AND THE PSI 20 STOCK INDEX

In the context of portfolio management it is more important to evaluate diversified portfolios than individual assets. For this reason, it is normal not to give much importance to the specific risk once this is susceptible of minimization (and in theory can be null) by an effective diversification process. The systematic risk is one of the main objectives of the financial analysis attention.

Usually the systematic risk is measured by the *Beta* of the CAPM. In this context, it is assumed that the rate of returns of some asset is equal to the sum of two components: the risk free rate of return (R_f) and the compensation for the risk $\{[E(R_m) - (R_f)]\beta_i\}$. In this second component, the coefficient *Beta* (β) assume special importance, since it measures the sensibility of the rate of returns of the asset (or portfolio) face to the risk premium, and it measures the systematic risk. In order to exemplify that the *Beta* can be a measure of the systematic risk, we must remember that the variance of an asset $\left(\sigma_i^2\right)$ (or portfolio) can be decomposed in two components: the sum of squares of the regression and the residual sum of squares, that is

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{ei}^2, \tag{12}$$

where σ_m^2 is the variance of the independent variable, in this case is the variance of the market benchmark and σ_{ei}^2 is the residual part of the variance, that can be minimized through a diversification process.

In order to estimate the Beta of the CAPM we use the Market Model, that can be given by

$$R_{it} = R_{ft} + [R_{mt} - R_{ft}]\beta_i + \varepsilon_{it}. \tag{13}$$

The Market Model is estimated by OLS method. However, statistical tests and the evidence show that the residuals are not a white noise.[7]

The main goal of this section is to evaluate the level of dependence between each one of the stocks and the stock index. The measures of information theory, namely the entropy, the conditional entropy and the mutual information are used to evaluate the (in)dependence between the stocks and the stock index PSI 20. If the residuals are white noise, then the global correlation coefficient (λ) based on the mutual information will be equal to the linear correlation coefficient (R), and the Beta can be considered a good measure of the systematic risk. In case of existence of nonlinearities and irregularities in the behaviour of the residuals resulting from the estimation process of the Market Model, the simple linear model cannot be enough to capture the existence of a global relationship between the stocks and the stock index PSI 20. In this case, the mutual information and the global correlation coefficient can be potential sources of information for the investor.[8]

According to the investor's point of view, a stock with a high Beta can be meaning of higher risk, since it is understood how the stock reacts face to variations in the benchmark market. On the other hand, if that correlation (linear and non-linear) is really strong, it is natural that the uncertainty about the behaviour of that stock decreases, therefore the knowledge of the behaviour of the stock index will increase the possibility of a higher level of predictability. In this context, exist some "rivalry" between the concepts of risk and uncertainty, since the smallest uncertainty results from the strong correlation between the stock and the stock index and, at the same time, it is understood as a larger systematic risk assumed by the investor.

Through the measures of the information theory used in this paper, it is possible to distinguish the global uncertainty from the "residual" uncertainty, according to the properties of the entropy; the entropy of a title (or any other variable) can be decomposed in the following way

$$H(X) = I(X,Y) + H(X|Y).$$

If we consider that X represents a stock and Y is the stock index PSI, then

$$H(X) = I(X,PSI) + H(X|PSI). \tag{14}$$

The equation (14) can be comparable, in terms of its behaviour, to the equation (12), where the first term refers to the level of association or dependence between the asset (stock or portfolio) and the proxy (here the proxy is the index PSI 20), and the second term is the variation of that asset (or stock, or portfolio) that is independent from the proxy used.

We estimated the *Beta* (β) for each stock and the linear correlation coefficient (R) resulting from the Market Model. In addition, we also calculated the systematic risk, $\beta_i^2\sigma_m^2$, and the specific risk, σ_{ei}^2. In order to provide some comparisons between the two approaches, we calculate the mutual information between each stock and

[7] Bartholdy and Peare (2004) highlight the fact that the estimation method of the CAPM, the type of proxy used and the periodicity of the data are not consensual on the academic field [vide e.g. Chen (2003)]. In order to estimate the Beta of the CAPM we used OLS and also Two Stage Least Squares (TSLS), Three Stage Least Squares (3SLS) and Generalized Method of Moments (GMM). However, we will just present the OLS results since we did not find significant differences through those methods. We also used weekly data, but the results were not significantly different from the daily data.

[8] Several studies alert to he fact that the Beta is not enough to measure the systematic risk [e.g Fama *et al.* (1993); Fama *et al.* (1996)].

the stock index PSI 20, $[I(X, PSI)]$, the conditional entropy $[H(X|PSI)]$, and the global correlation coefficient, λ.

As we can see in the Figures 5 and 6, there exists a positive relationship between the systematic risk and the mutual information, and between the specific risk and the conditional entropy.

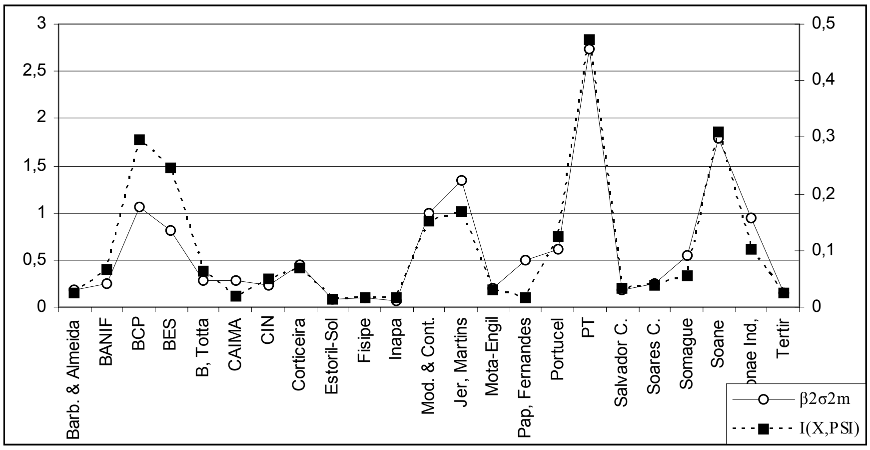

FIG. 5 Comparative analysis between the systematic risk, $\beta_i^2 \sigma_m^2$, and the mutual information I.

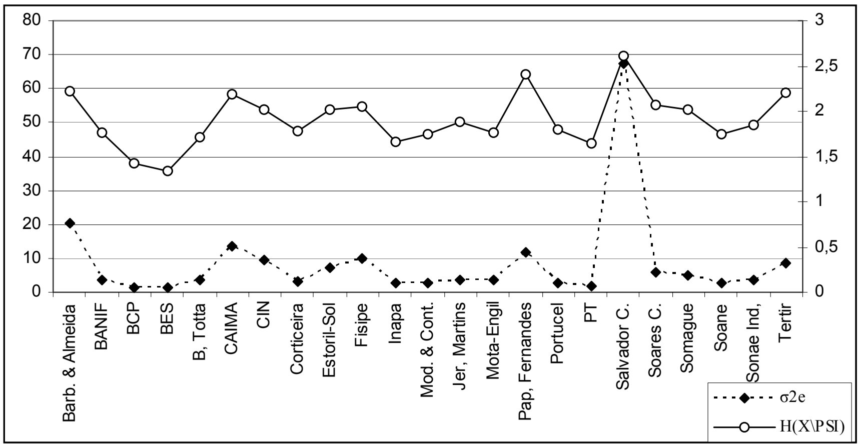

FIG. 6 Comparative analysis between the specific risk, σ_{ei}^2, and the conditional entropy, $H(X|PSI)$.

In spite of the apparent positive and strong relationship between the variance analysis and the measures of the theory of the information it is convenient to proceed to a comparison among measures whose values can be directly comparable. In this context, we compare the global correlation coefficient (λ) and the linear correlation coefficient (λ) (see Figure 7). If the relationship between the variables can be faithfully represented by a linear model and the residual are white noise, then λ and R should provide similar figures.

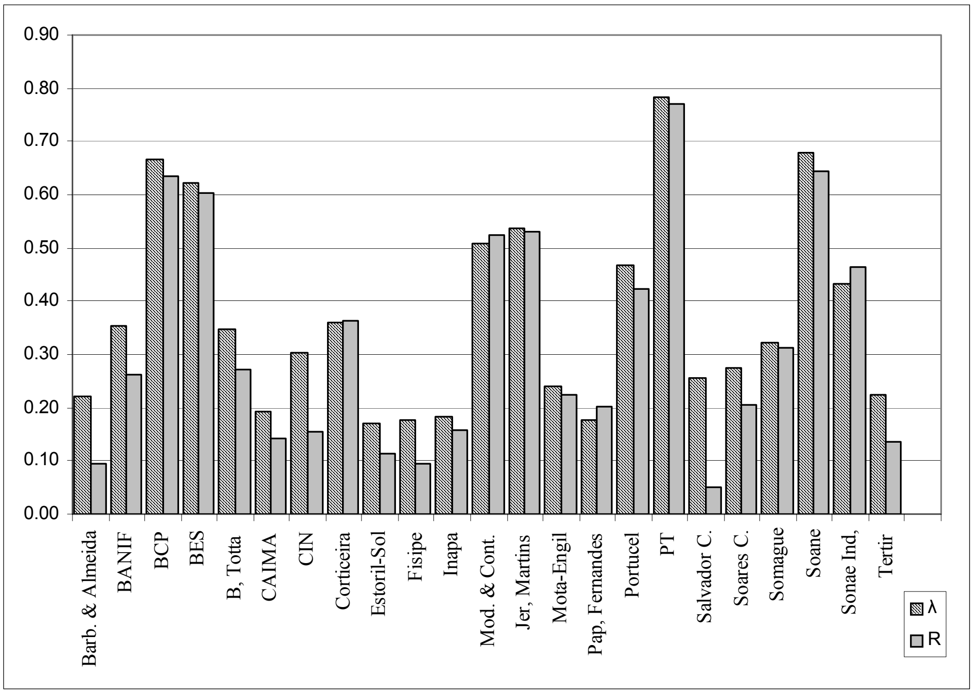

FIG. 7 Global correlation coefficient (λ) and linear correlation coefficient (R) between each of the stocks and the stock index PSI 20.

As we can see from Figure 7, there are stocks whose relationship with the stock index PSI 20 exhibit strong discrepancies when analyzed from a global or purely linear perspective. We must highlight the stocks Barbosa & Almeida, Banif, Caima, CIN and Salvador Caetano that present the most significant differences between λ and R. In order to find the possible causes of such differences several tests were accomplished for the residuals produced by the estimation of the linear regression model (Market Model), namely the Ljung-Box test, the Jarque-Bera test, the Engle test and the stability tests *CUSUM* and *CUSUM-Q*. The results of those tests indicate that are precisely the residuals resulting from the application of the Market Model among the stocks Barbosa & Almeida, Banif, Caima, CIN and Salvador Caetano and the stock index PSI 20, that show the highest evidence of autocorrelation, non-normality, heteroskedasticity and non-stability, and this seems to be an indicator that the linear analysis is possibly not enough to evaluate risk and uncertainty.

On the other hand, the residuals of the stocks BCP, BES, Corticeira Amorim, Modelo & Continente, Jerónimo Martins, PT, Somague and Sonae exhibit the closest values of λ and R. If we analyze the behaviour of those residuals we see that are precisely the ones that present smaller values of linear autocorrelation, as well as the values of the asymmetry and kurtosis are closer to the normal distribution, and there is no evidence of structural breaks. Such results can be related with the fact that these stocks present a high level of liquidity and consequently, it is higher the possibility that their efficiency levels are higher than that of the stocks which liquidity and transaction volumes are relatively weak.

In general, these results indicate that in most cases the use of a linear model to evaluate the relationship between the stock and the stock index PSI 20 can be not enough, since there is evidence of nonlinearities on the data and the residuals resulting from the Market Model are not white noise. The measures of the information theory used in this section, namely the entropy, the conditional entropy and the mutual information show a strong relationship with the analysis of variance and have the capacity to capture some particularities of the variables and the residuals that a linear model practically omits. It is verified that the entropy and their variants has the ability to capture (individually or globally) the existence of asymmetry, kurtosis, autocorrelation, heteroskedasticity and the existence of structural breaks.

5. CONCLUSIONS

From the present analysis, we can conclude that entropy observes the effect of diversification and is a more general measure of uncertainty than the variance, since it uses much more information about the probability distribution. The mutual information and the conditional entropy show a good performance in comparison with the systematic risk and the specific risk estimated through the Market Model.

Nevertheless, the use of the entropy in risk analysis and portfolio selection needs some care because it does not take into account the actual values of the variables and this fact can compromise its inclusion in the context of a utility function.

REFERENCES

[1] Arafat, S; M. Skubic and K. Keegan, 2003. Combined Uncertainty Model for Best Wavelet Selection, Proccedings of the *IEEE 2003 International Conference on Fuzzy Systems*, May 2003, St. Louis, MO.

[2] Bartholdy, J. and P. Peare, 2004. Estimation of Expected Return: CAPM vs Fama and French, *Working Paper Series N.^{o} 176*, Centre for Analytical Finance, University of Aarhus - Aarhus School of Business.

[3] Chen, M. 2003. Risk and Return: CAPM and CCAPM, *The Quarterly Review of Economics and Finance*, 43, 369-393.

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

[5] Ebrahimi, N., E. Maasoumi and E. Soofi, 1999. Ordering Univariate Distributions by Entropy and Variance, *Journal of Econometrics*, 90, 2, 317-336.

[6] Elton, E. J. and M. J. Gruber, 1995. *Modern Portfolio Theory and Investment Analysis* (Jonh Wiley & Sons, 2^a Ed, New York).

[7] Fama, E. and K. French, 1993. Common Risk Factors in the Returns on Bonds and Stocks, *Journal of Financial Economics*, 33, 3-56.

[8] Fama, E. and K. French, 1996. The CAPM is Wanted, Dead or Alive, *Journal of Finance*, 51, 5, 1947-1958.

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

[10] Keynes, J. 1937. Alternative Theories of the Rate of Interest, *Economic Journal.*

[11] Knight, F. 1921. *Risk, Uncertainty, and Profit*, in Hart Schaffner and Marx; Boston ed.: *Houghton Mifflin Company* (The Riverside Press, Cambridge).

[12] Kraskov, A., H. Stögbauer, R. Andrzejak and P. Grassberger, 2004. Hierarchical Clustering Based on Mutual Information, preprint http://www.arxiv:q-bio.QM/0311039.

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

[14] Lawrence, D. 1999. *The Economic Value of Information* (Springer, New York).

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

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

[17] McCauley, J. 2003. Thermodynamic Analogies in Economics and Finance: Instability of Markets, *Physica A*, 329, 199-212.

[18] 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.

[19] Philippatos, G. and C. Wilson, 1972. Entropy, Market Risk and the Selection of Efficient Portfolios, *Applied Economics*, 4, 209-220.

[20] Reesor, R. and D. McLeish, 2002. Risk, Entropy and the Transformations of Distributions, preprint in *Working Paper 2002-11*, Bank of Canada.

[21] Shannon, C. E. 1948. A Mathematical Theory of Communication, *Bell Systems Tech.*, 27: 379-423, 623-656.

[22] Soofi, E. 1997. Information Theoretic Regression Methods, Fomby, T. and R. Carter Hill ed: *Advances in Econometrics - Applying Maximum Entropy to Econometric Problems*, vol. 12 (Jai Press Inc., London).