An econophysics approach to analyse uncertainty in financial markets: an application to the Portuguese stock market

Andreia Dionisio,[1] Rui Menezes and Diana A. Mendes[2]

[1] University of Evora, Center of Business Studies, CEFAG-UE, Largo Colegiais, 2, 7000 Evora, Portugal E-mail: andreia@uevora.pt

[2] ISCTE, Departament of Quantitative Methods, Av. Forcas Armadas, 1649-Lisboa, Portugal E-mail: rui.menezes@iscte.pt, diana.mendes@iscte.pt

Abstract. In recent years there has been a closer interrelationship between several scientific areas trying to obtain a more realistic and rich explanation of the natural and social phenomena. Among these it should be emphasized the increasing interrelationship between physics and financial theory. In this field the analysis of uncertainty, which is crucial in financial analysis, can be made using measures of physics statistics and information theory, namely the Shannon entropy. One advantage of this approach is that the entropy is a more general measure than the variance, since it accounts for higher order moments of a probability distribution function. An empirical application was made using data collected from the Portuguese Stock Market

PACS. 89.70.+c discribing text of that key – 89.90.+n discribing text of that key

1 Introduction

The application of mathematical and physics models to finance goes back to Bachelier in 1900, where it tests the hypothesis that stock prices follow a random walk. However this simple version of the model did not account for

Send offprint requests to: andreia@uevora.pt

important characteristics of price variations, such as the occurrence of crashes, nonlinear serial dependence, etc. Bachelier assumed that the price variations follow a normal distribution, constant over time, and do not pay attention to extreme events. However, the empirical evidence has shown that stock prices seldomly behave in such a way as described by Bachelier.

The stock markets are usually complex systems, because they are open systems where innumerous subsystems act and interact in a nonlinear and dynamic way, constituting an attraction for the physicists that studied the working of financial markets using different methods than those used by traditional economists.

Bonanno, Lillo and Mantegna (2001) consider that the financial markets show several levels of complexity that may occurred for being systems composed by agents that interact nonlinearly between them. These authors, among others, consider that the traditional models of asset pricing (CAPM and APT) failed because the basic assumptions of these models are not verified empirically.

The entropy is a measure of dispersion, uncertainty, disorder and diversification used in dynamic processes, in statistics and information theory, and has been increasingly adopted in financial theory [Horowitz *et al.* (1968), Philippatos *et al.* (1972), Buchen *et al.* (1996), Zellner (1996), Molgedey *et al.* (2000), Stuzer (2000), London *et al.* (2001)].

In addition to the studies mentioned above, Bouchaud, Potters and Aguilar (1997) have used entropy as an integrating measure in the process of portfolio selection based on the mean-variance model of Markowitz. This is because information is imperfect and the theoretical assumptions of portfolio selection models do not apply in the reality. The authors suggest the use of entropy with the purpose of obtaining a minimum diversification and, at the same time, an acceptable risk level to the investor. In a slightly different context, Fernholz (1999) and Samperi

(1999) analysed the entropy as a measure of diversification in financial markets. Gulko (1998) analyses market equilibrium by building a model where entropy is maximized subject to certain restrictions. He defends the "entropy pricing theory" as the main characteristic of market efficiency.

The use of entropy as a measure of uncertainty in finance appears to have many potentialities and a vast field of development, both in theoretical and empirical work. In line with the above arguments, this paper examines the ability of entropy as a measure of uncertainty in portfolio management applied to the Portuguese stock market, highlighting the fact that entropy verifies the effect of diversification.

In this article, the notion of uncertainty is related with the greater or lesser difficulty to predict the future. Generally, it is normal to relate the variance or the standard-deviation and the VaR (Value-at-Risk) as the main risk and uncertainty measures in finance. However, some authors [e.g. Soofi (1997), Maasoumi (1993), McCauley (2003)] alert for the fact that these measures can fail in specific situations as a measure of uncertainty, since they need that the probability distributions are symmetric and neglect the possibility of extreme events such as the existence of fat-tails.

2 Theoretical background

Suppose that we have a set of possible events whose probabilities of occurrence are $p_1, p_2, ..., p_n$ and H is a measure of uncertainty. According to Shannon (1948), a good

measure of uncertainty $H = H(p_1, ..., p_n)$ should satisfy the following properties:

1. H should be continuous in $p_i, i = 1, ..., n$;

2. If $p_i = 1/n$, then H should be a monotonic increasing function of n;

3. H is maximized in a uniform probability distribution context;

4. H should be additive;

5. H should be the weighted sum of the individual values of H

According to Shannon (1948) one measure that satisfies all 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

$$H(X) = -\int p_X(x) \log p_X(x) dx. \tag{1}$$

The properties of the entropy of continuous and discrete distributions are mainly alike. In particular we have [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 $H(X, Y) \leq H(X) + H(Y)$ 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 $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$, since $H(X) + H(Y) \geq H(X, Y)$, then $H(Y) \geq H(Y|X)$ 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

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

Arafat, Skubic and Keegan (2003) consider that a measure of uncertainty should attend to the following properties: (i) *Symmetry*, that is $H(X) = H(1 - X)$; 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.

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).

---

[1] The histogram-based estimators are divided in two groups: equidistant cells and equiprobable cells, i.e. marginal equi-quantisation [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 (1998)].

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 entropy and variance, because entropy is a concave function allows its use as an uncertainty function.

3 Entropy and diversification effect: an example

Historically, the variance has had a fundamental role in the analysis of risk and uncertainty. However, according to Maasoumi (1993), entropy can be an alternative measure of dispersion and in addition 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 "utility" of using $p_X(x)$ in place of 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 measures reflect concentration but their respective metrics of concentration are different. Unlike the variance that measures concentration only around

the mean, the entropy measures diffuseness of the density irrespective of the location of concentration.

In terms of mathematical properties, entropy $[H(X)]$ is non-negative in the discrete case. In the discrete case, $H(X)$ is invariant under one-to-one transformations of X, but the variance is not. For the continuous case, neither the entropy nor the variance are invariant under one-to-one transformations. The entropy of a continuous random variable X takes values in $]-\infty, +\infty[$ [Shannon (1948)].

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. A Legendre series expansion reveals that entropy may be related to higher-order moments of a distribution which, unlike the variance, could offer a much better characterization of $p_X(x)$ since it uses more information about the probability distribution than the variance.

In this paper we examine the sensitivity of entropy to the effect of diversification. The risk of a portfolio can be splitted into specific risk and systematic risk, that is not diversifiable. Using entropy we can obtain a similar type of information, since $H(X) = I(X,Y) + H(X|Y)$, where $I(.)$ is the mutual information between X and Y and may be comparable with the systematic risk and $H(.|.)$ is the

conditional entropy that can be comparable with the specific risk. We must emphasize that the measures of information theory are not directly comparable to the analysis of variance in metric terms.

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 $E\left[\sigma\left(X\right)\right] \geq \sigma\left[\left(EX\right)\right].$[2] 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 is a concave function and has a maximum for most of the probability distributions, and this fact leads us to think 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 the probability itself and the property $H\left(X,Y\right) \leq H\left(X\right) + H\left(Y\right)$ can bring some hope in this way.

In this paper we perform a similar analysis to that presented by Elton and Gruber (1995). These authors showed that 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 $1/N$, being N the number of assets in the portfolio. We use daily closing prices of 23 stocks rated on the Portuguese stock market (*Euronext Lisbon*), spanning from 28/06/1995 to 30/12/2002, which

corresponds to 1856 observations *per* stock, in order to compute the rates of return. The statistical analysis of these time series revealed that we must reject the null that the empirical distributions are normal, since they show high levels of kurtosis and skewness.

In order to compare the behaviour of entropy with the standard-deviation in a coherent way, we use the normal entropy (equation 2), since the normal entropy is a function of the standard-deviation.

Our results (see Figure 1) 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 respects the condition of subadditivity suggested by Reesor and McLeish (2002), where $H[\theta X] + H[(1-\theta)Y] \geq H[\theta X + (1-\theta)Y]$, being $\theta = 1/N$.

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.

From this preliminary analysis, we can conclude that entropy observes the effect of diversification and is a more general uncertainty measure than the variance, since it
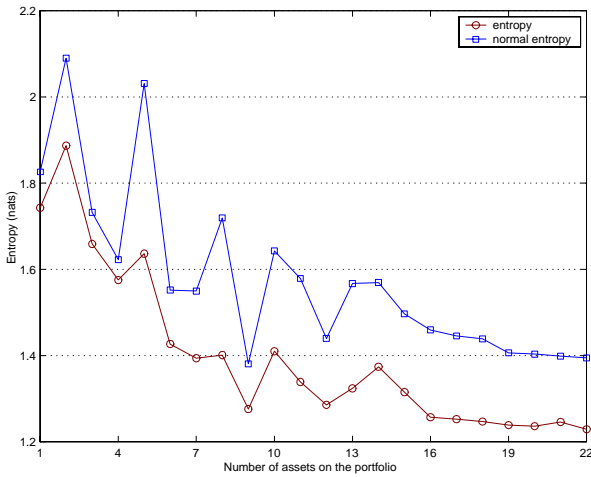
---

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

Fig. 1. 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.

uses much more information about the probability distribution.

4 Final remarks

This paper analyses the use of entropy as a measure of uncertainty in portfolio management. This can be a complementary way to the traditional mean-variance models, whose assumptions are typically quite restrictive. Our approach takes into account the higher-order moments of the empirical probability distributions, rather than just the variance that only uses the second moment.

The results suggest that entropy is sensitive to the effect of diversification and is apparently a more general measure of uncertainty than the variance.

References

1. Arafat, S; M. Skubic and K. Keegan. Combined Uncertainty Model for Best Wavelet Selection, Proccedings of the IEEE 2003 International Conference on Fuzzy Systems, 2003 (edited by St. Louis, MO).

2. Bonanno, G., F. Lillo and R. Mantegna. Physica A, 299, 16-27, (2001)

3. Bouchaud, J.P., M. Potters and J.P. Aguilar. Missing Information and Asset Allocation, preprint in http://xxx.lanl.gov/cond-mat/9707042, (1997).

4. Buchen, P.W. and M. Kelly. Journal of Financial and Quantitative Analysis, **31**, 1, (1996).

5. Darbellay, G. Predictability: an Information-Theoretic Perspective, Signal Analysis and Prediction, A. Procházka, J. Uhlír, P.J.W. Rayner and N.G. Kingsbury, Birkhauser eds., Boston , 249-262, (1998).

6. Ebrahimi, N., E. Maasoumi and E. Soofi. Journal of Econometrics, **90**, 2, 317-336, (1999).

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

8. Fernholz, R. Journal of Mathematical Economics, **31**, 3, 393-417, (1999).

9. Gulko, B. *The Entropy Pricing Theory - Market Beliefs, Valuation, Incomplete Markets, Asset Pricing*, PhD dissertation, UMI – Dissertations Services, 1998.

10. Horowitz, A. and H. Horowitz. Journal of Industrial Economics, 196-211, (1968).

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

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

13. London, M.D., A. K. Evans and M.J. Turner. Quantitative Finance, **1**, 414-426, (2001).

14. Maasoumi, E. Econometric Reviews, **12**, 2 , 137-181, (1993)

15. Molgedey, L. and W. Ebeling. European Physics Journal B, **15**, 733-737, (2000).

16. McCauley, J. Physica A, **329**, 199-212, (2003)

17. Philippatos, G. and C. Wilson. Applied Economics, **4**, 209-220, (1972).

18. Reesor, R. and D. McLeish. Risk, Entropy and the Transformations of Distributions, preprint in Working Paper 2002-11, Bank of Canada, (2002).

19. Samperi, D. Entropy and Statistical Model Selection for Asset Pricing and Risk Management, Working Paper SSRN, http://papers.ssrn.com (1999)

20. Shannon, C. E. Bell Systems Tech., **27**, 379-423, 623-656, (1948).

21. Soofi, E. 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, 1997).

22. Stuzer, M. Entropy, **2** , 70-77, (2000).

23. Zellner, A. Journal of Econometrics, **75** , 51-68, (1996).