Crash Probability Occurrence and Stock Market Efficiency the Tunisian Stock Exchange Case via Shannon Entropy

SAHLI Lamia and Adel Boubaker

***Abstract*— In this study, we evaluate the relationship between efficiency and probability of the crash, thus the evolution of the daily informational efficiency is measured for Tunisian stock market index (TUNINDEX) over the period [1998 – 2008]. The efficiency, which is the issue addressed by the weak-form efficient market hypothesis, is calculated using a new method the Shannon entropy and the symbolic time series analysis. A logit model is applied in order to study the relationship between efficiency and probability of the financial crash.**

***Keywords*— Informational Efficiency, Financial Crashes, The Shannon Entropy and Logit Model**

I. INTRODUCTION

The theory of efficiency is considered as one of the most controversial theories in the financial markets theory. Despite the abundance of empirical work to test the hypothesis of efficiency, no clear conclusions seemed. This lack of consensus results is probably related to the importance of efficiency in financial theory. A market may be defined as efficient in the informational sense if the prices of the assets traded on that market instantaneously reflect all available information. The weak-form version of the Efficient Markets Hypothesis (EMH) has been subjected to years of rigorous empirical testing across national stock markets. Since Samuelson (1965) introduced the concept of fair game to financial economics, the EMH is still in controversy. Researchers' views on the EMH have changed during the past 40 years. Fama (1970) summarized that, in early researches, many articles supported the EMH; Fama (1991) himself modified his view because of a lot of anomalies reported after Fama (1970). However, the bulk of the literature focused mainly on testing whether the stock market is efficient or inefficient for the selected sample period, using statistical tests such as the serial correlation tests, runs test, variance ratio tests, unit root tests and spectral analysis. Today nobody can summarize the dispute over the EMH in a couple of lines. The battle between proponents of the EMH and advocates of behavioral finance is still ongoing and, the authors believe, will never end. Considering the possibility of finding the deterministic chaos in financial markets, or at least the fractal Brownian motion some authors have proposed the Hurst exponent as a measure of efficiency, Peters (1994-1996). Grech and Mazur [2004], Taback and cujeiro (2004, 2005, 2008, 2009) have employed the Hurst exponent to measure changes in efficiency over time. However, some authors, such as Bassler et al [2006] and Mc Cauley and others (2007) criticize this measure and argue that a Hurst exponent different from ½ (the number corresponding to a random walk process does not necessarily imply correlations of long-term as those found in fractional Brownian motion.

In the present work a new methodology suggested in order to measure the efficiency of information, commonly used in information theory. Our research methodology is distinguishable from previous studies in at least three points. First off all the symbolization of returns is applied. Secondly the Shannon entropy is calculated to measure the amount of information contained in the series, The Shannon entropy is larger the market is more efficient. Thirdly, a logit model is applied to study the relationship between efficiency and the probability of financial crash, the logit model suggests that a decrease in efficiency increases the probability of a crisis on the financial market.

The reminder of the paper is organized as follows. In the following section we will present the methodology Measurement of efficiency by the Shannon entropy and the logit model will be explained. In section 2 we will present the results for the TUNINDEX index (Tunis). Finally, we summarize the finding of this paper in the last Section.

II. RESEARCH METHODOLOGY

Risso [2008] supposes that on an efficient market the actual prices (Pt) reflect all the available information. Therefore, we assume a perfectly efficient market, the returns are unpredictable. The model most commonly used to measure the efficiency is the random walk model. This model has been intensively studied by Hurst in 1940s and later by Mandelbort in 1960s and 1970s, Grech and Masure (2004) uses the local Hurst coefficient to measure efficiency. In the present paper instead of the Hurst exponent uses the Shannon entropy.

A. Measuring the informational efficiency by the Shannon entropy

Firstly, we assume that on an efficient market the actual prices reflect all available information. Therefore, we assume that the returns in efficient market are unpredictable. The measure of efficiency is computed in two steps. First the symbolization of returns is applied in order to detect the changes over time. Secondly, the Shannon entropy used in order to measure the quantity of information contained in the series. Using the Symbolic Time Series Analyses (STSA), we can obtain richer information from a time series if we transform data series of many possible values into a symbol time series of only a few distinct values.

The problem of symbolic analysis is that there is no formal way to define the time series partitions. However, in our case, we are interested in combinations of negative and positive stock market returns. Therefore, we take (0) for negative returns and (1) for positive ones. In fact, in this case the process is a sequence of Bernoulli processes (0s and 1s).

Thus we define a time series of size T defined as (r_1, r_2, r_3, ..., r_T), being r_t the asset returns at time t, for t = 1, 2,3 .... T.

$$r_t = \log P_t - \log P_{t-1} \qquad (1)$$

Then we proceed to transform the time series in a symbolic one, according to eq (2).

$$\begin{cases} r_t < 0 & s_t = 0 \\ r_t \geq 0 & s_t = 1 \end{cases} \qquad (2)$$

Thus it is obtained the following symbolic time series (s_1, s_2, s_3,….., s_T),it is the same time series, as sequences of 0s and 1s, representing the decreases and increases in prices, respectively.

First, we implement the Shannon entropy (H) as our measure of efficiency. This measure making its maximum value at 1 when the process is completely random, and its minimum at 0, when there is a completely certain event. The theoretical expression for *H* is the Eq. (3).

$$H_t = \left(\frac{-1}{\log_2(n)}\right)\sum p_t \log_2 p_t \qquad (3)$$

Where n: total number of sequences

Pi: the probability of sequence i = 1, 2, ..., n

It must be noted that the entropy reaches its maximum value 1 when for the n sequences or events are equally probable, $p_1 = p_2 = p_3 = ... = p_n$. However, the entropy reaching its minimum value 0 when one event cumulates all the probability. It means that there is a completely certain event. For example, if we computed our measure for only two events, assuming that P is the probability of having a decrease in prices (symbol 0) and (1-P) is the probability of having an increase in prices (symbol 1), according to the Shannon entropy the efficiency will be given by Eq. (4).

$$H_t = \frac{-1}{\log_2(2)}\left(p\log_2 p + (1-p)\log_2(1-p)\right) \qquad (4)$$

We should mention that H is a concave function, its maximum value is obtained for p=1/2 and its minimum value is obtained for p = 1 and p = 0. In practice, the probability of decrease and increase in price is the frequency of each event throughout the period divided by the total number of events in the same period. Since we are interested in the evolution of efficiency over time, we take a sub-period V <T and move it through time. Then, the entropy is computed for each time window from moment 1to T. Different window sizes are considered, 50, 100, 250 days. Grech and Mazur (2004), (using the Hurst exponent) suggest that the window should not be so large in order to capture the locality. We choose the window that adopts the best R^2 in the logit model which we employed later to investigate the relationship between efficiency and the probability of financial crises.

B. The logit model

Logit and probit models are two families' models for binary endogenous variables. Suppose there is a variable Y that takes one of the two values, 0 and 1. In the present case, the financial crash (1) and no-crash (0). Let Y * a latent variable such that:

$$y_t^* = \alpha + \beta H_t + \varepsilon_t \qquad (5)$$

Where H is the measurement of efficiency and ε_t follows is called an extreme value distribution, see McFadden [1984]. We do not observe y* there, but rather y, which takes on values of 0 or 1 according to the following rule

$$y_t = \begin{cases} 1 & \text{if } y_t^* > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

In our study, the variable y takes the value (1) when the empirical distribution of index returns cumulates the 1% (the negative tail), and (0) otherwise. According to Johnston and DiNardo (1997), the logit model is given by:

$$p(y_t = 1) = \frac{\exp(\alpha + \beta H_t)}{1 - \exp(\alpha + \beta H_t)} \qquad (7)$$

Eq. (7) indicates that the probability of financial crash in one day, p (y_t = 1) depends on the efficiency level H.

The formulation of the model (7) ensures that the predicted probabilities are between 0 and 1. However, the calculation of the probability of change is not as simple as it was in a linear model. The derivative of the probability of crash with respect to the efficiency H varies with H as regards the efficiency varies with H.

$$\frac{\partial E(y)}{\partial H} = \frac{\exp(\alpha + \beta H)}{(1 - \exp(\alpha + \beta H))^2}\beta \qquad (8)$$

The logit model can be expressed as the odds-ratio as in Eq. (9) which is usually more intuitive.

$$\frac{p(y_t=1)}{1-p(y_t=1}=\exp(\alpha+\beta H_t) \qquad (9)$$

$\exp(\alpha+\beta H_t)$ is the effect of the independent variable (our measure of efficiency) in the odds-ratio.

III. EMPIRICAL RESULTS

In this study, we employed daily Tunisian Stock Exchange index time series over the period 1998 to 2008.

Data was extracted from the Tunisian stock Exchange Web site.

A. Presentation of stock market in Tunis

Tunis Stock Exchange, officially called Stock Exchange of Tunis (BVMT) since November, 15, 1995, is a stock based in Tunis. It is responsible for management, security and the promotion of Tunisian stock market securities and their shareholders are companies on the stock market intermediation. Its main index is the TUNINDEX. The creation of the Tunisian Stock Exchange dates back to 1969. The role of stock market in the financing of the economy of Tunisia has been limited or even insignificant because of the predominance of the banks.

Since 2003, the Tunis Stock Exchange was in a raise trend and one year to the other Tunindex, which is the index capitalization, reached levels more and more raised. Certainly that the number of listed companies also increased but the levels of valuation increased even more. The offer became more varied and investors now have more choices. The year 2008 was despite of the release of the financial crisis at the international level. Just before, Tunindex was up 30, 37 % (09/09/2008), but then its performance crumbled away to finish really up 10, 32 %. This increase was the second best performance in the world among the three reported. Without repeating the reasons for the resistance of the Tunisian stock market to such external shocks, but it gave a good breath for 2009.

B. Descriptive statistics

Table 1 and Table 2 provide descriptive statistics, mean, standard deviation, skweness, kurtosis and the value of statistical normality test of Jarque-Bera. Remember that for a normal law, the skweness is zero and the kurtosis is 3, the law being characterized by its symmetry with respect to the average and the low probability of extreme points. Under the null hypothesis of normality, the JB statistic follows a Law Chi-square with two degrees of freedom.

Table 1: The descriptive statistics of returns

Mean	S.D	Skweness	Kurtosis	Jarque-Bera
0.000000	0.000000	0,020044	8,960596	3370945

Statistics provided by Eviews 6.0

As it is shown in table 1 the null normality hypothesis is rejected. Firstly, we see that the coefficient of Kurtosis is very high; it's much higher than 3 .This phenomenon of excess of kurtosis is confirmed by the leptokurtique characteristic of the stock market return series. Secondly, the coefficient of Skweness is different from 0 (theoretical value of the coefficient of Skweness for the normal distribution), we note that the coefficient is positive Skweness for the series of returns of the index TUNINDEX. This indicates that the distribution of series is spread to the right. This illustrates the presence of asymmetry, which may be indicating of nonlinearity.

C. The results of the logit model

The logit model was estimated using the efficiency measure as an independent variable, taking different order or combination of days (1,2,3,4) and different time windows (100,220,350,430 days). 16 models were estimated in total. The maximum pseudo-R 2 is shown in bold characters and highlights the logit model specification which fits the best.

Table 2: R2 index TUNINDEX (Tunis)

	100	220	350	430
1	0,0057	0,0006	0,004	0,000
2	0,0114	0,0024	0,0001	0,0012
3	0,0120	0,0027	0,0001	0,0008
4	**0,0151**	0,0013	0,0005	0,0005

We observe that the window (v = 100, 4d) corresponds to the highest R2. Therefore it is the window that best explains the relationship of the logit model; the exogenous variable H explains the endogenous variable y.

IV. RESULTS AND COMMENT

The results of the logit model for the window (v = 100, 4d) are presented in Table 3 and Fig. 1:

Table 3: The logit model estimating results for the index TUNINDEX, for windows (v = 100, 4d)
Log likelihood = -255.59273

Crash prob[c]	coefficient	Standard error	t	p-value>t
Entropy β	-11.42838	3.942558	-2.90	0.004
Constant α	6.862849	3.68616	1.86	0.063
Crash prob[d]	Odds-ratio	Standard error	t	p-value>t
Entropy	0.0000109	.0000429	-2.90	0.004[b]

The results are obtained by the econometric software STATA.

a: indicates that the model is significant at the 5%.

b: indicates that the coefficients are significant at the 5%.

c: the estimation of Eq. (7).

d: the model expressed in odds-ratio Eq. (8).

We should mention that the probability of a crash depends negatively on our measure of informational efficiency H. When the informational efficiency decreases, the probability of having a crisis increases.

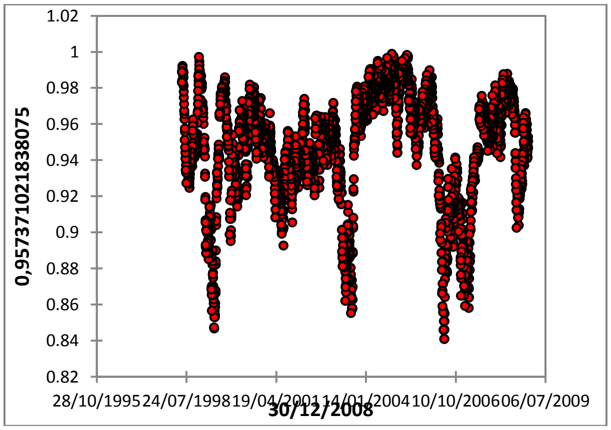

Fig. 1: the daily changes of efficiency for the index TUNINDEX (Tunis)

Fig. 1 reports the evolution of the efficiency of the TUNINDEX index (taking v = 100 and an order of 4 day). We note that there are two beams of light inefficiency between December 2007 and February 2008 and between July and October 2008.A local minimum is obtained in December 2007, August and September 2008. After a study of the national and international economy for the period studied we find that the events caused by those two beams of inefficiency are:

The first beam of inefficiency is the result of the current world economic situation which marked by a new surge in oil prices which have exceeded $ 98 a barrel and an unprecedented rise in prices of staple food products. Inflation remained under control despite the effects of price increases on certain products of developing economies, the direction of the dollar continues to decline and disruptions in international finance. These last ones have been exceeded before the end of the year; have registered a return whose effects have resulted in a crash felt in all global markets in early 2008.

The second beam of inefficiency is the result of the outbreak of the international financial crash and the succession of financial events a worldwide. Our stock exchange slowed despite appeals for calm and assurance from the Government on the good performance of our financial system, including banking. Moreover, during this period the place has undergone a wave of financial communications from listed companies to reassure investors about the soundness of their financial situations. However, the mass exit of foreign investment funds has led to a phase of panic and falling prices. The launch by the state of the two funds of MDT 50 each, unfortunately, has not made its effect, and the stock market continued to lose off the beginning of the year ending in a performance of 10.65%.

V. CONCLUSION

In this paper, a new approach was proposed in order to measure the efficiency of information, commonly used in information theory. Under this methodology the degree of efficiency is done in two stages, first, the symbolization of returns is applied, thereafter, the Shannon entropy is computed to measure the amount of information contained in the series, the larger the Shannon entropy is, the more efficient market is. Once the efficiency is measured, a logit model is applied to study the relationship between efficiency and the likelihood of financial crash, the logit model suggests that a decrease in efficiency produces an increase in the probability of a crisis on the financial market.

The study reveals that the Tunisian financial market tends towards tends efficiency without entirely meeting, so we see that the market goes through periods of inefficiency because of the crisis affecting the international financial market is causing the Tunisian beams of inefficiencies. Moreover our results confirm the negative relationship between efficiency and the likelihood of financial crises, but the market will find the way back to efficiency.

REFERENCES

Cajueiro, D. O. and Tabak, B. M. [2004], “The Hurst’s exponent over time: testing the assertion that emerging markets are becoming more efficient”; Physica A; Pages 336, 521; (2004).

Cajueiro, D. O. and Tabak, B. M. [2005], “Testing for timevarying long-range dependence in volatility for emerging markets”; Physica A, 346, Pages 577-588 (2005).

Cajueiro, D. O. and Tabak, B. M. [2008], “Testing for time-varying long-range dependence in real estate equity returns”; Chaos, Solitons & Fractals, Volume 38, Issue 1, October 2008, Pages 293-307

Cajueiro, D. O. and Tabak, B. M. [2009], “Testing for long-range dependence in the Brazilian term structure of interest rates”; Chaos, Solitons & Fractals, Volume 40, Issue 4, 30 May 2009, Pages 1559-1573.

Fama E.F. [1965], “the behavior of stock market prices”, journal of Business, vol.38, n°1, Pages 34-105.

Grech and Z. Mazur [2004], "Can one make any crash prediction in finance using the local Hurst exponent idea? Phys. A336, Pages 133-145.

Jensen-[1978], "Some anomalous evidence regarding market efficiency", Journal of Finance. Econ. Vol.6, Pages 95-101.

Johnston J, DiNardo J. [1997], "Econometric methods," fourth ed.McGraw Hill.

McCauley J, Gunaratne and Bassler G K, [2007] "Hurst exponents, Markov processes, and fractional Browiam motion", Phy. A 379 No. 1, Pages 1-9.

McFadden - [1984], "Econometric analysis of qualitative choice models. In: Giliches and Intriligator (Eds.), Handbook of Econometrics, North-Holland (Chapter 24).

Peters. E(1994), "Fractal Market Analysis", First Edition ed.wiley finance.

Peters. E (1996), "Chaos and order in the capital markets" second ed. Wiley Finance Edition.

Philippe Gillet, [1999], efficience de marché financier, Edition Economica

Risso W. [2008], "The Informational Efficiency and the Financial Crashes," Research in International Business and Finance, Vol. 22, Pages 396-408

Robert Gillet, [1991]," efficience informationnelle du marché boursier: definitions, tests empiriques et interpretation cohérente des resultants", cahiers Economiques de Bruxcelles, N°132 4^{éme} trimestre 1991.

Singal [2004], "Beyond the random walk: a guide to stock market anomalies and low-risk investing. Oxford university press.

Tabak B. M. and Cajueiro D. O., [2006], "Assessing inefficiency in euro bilateral exchange rates", Physica A: Statistical Mechanics and its Applications, Volume 367, 15 July 2006, Pages 319-327

Tabak B. M. and Cajueiro D. O.; [2007], "Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility"; Energy Economics, Volume 29, Issue 1, January 2007, Pages 28-36.