IS THE POLISH STOCK MARKET WEAK FORM EFFICIENT?

Pijanowski Slawomir
Poznan University of Economics

Abstract

This paper explores the definition of predictability of Warsaw Stock Index returns by using measures elaborated in Shannon-Mazur's cybernetic information theory, potentially a new approach to understand capital market informational efficiency. The main message of this research is that the use of information theory methods may shed new light on the applicability of weak-form efficiency tests and the phenomenon of return unpredictability. Cybernetic interpretation of predictability enables more refined and precise statistical interpretation in answering the question about market returns predictability and, in retrospect, may contribute to the discussion on the predictability tests of market returns.

JEL Classification: G1, G14 & D81
Key words: Cybernetic information theory, Weak form efficiency, Price predictability

1. Information theory and stock returns research

This paper attempts to examine the issue of the weak-form of market efficiency within the context of the Polish market. Two basic measures linked to the term *uncertainty* originating from information theory set, is the foundation for this inquiry, one being the amount of identifying information (entropy) and second, the number of describing information of the distribution of returns.

So far research methods on the subject rely extensively on the use of sequence tests, autocorrelation tests, and seasonality of returns within the specific period, etc. They rarely applied the information theory related to predictability of stock market returns. Some of the early examples can be found in the papers of Theil and Lenders and Fama as of 1965 using quantitative information theory developed by Shannon. Independently, during the 1960s and 1970s, in applying the theories about the science of steering (system behavior) in cybernetics, the quantitative information theory of Mazur was based on the quantitative information theory of Shannon. This paper concentrates only on the implications of differentiation in the identifying and describing information for the weak form of efficiency.

This research may add new findings to the current knowledge for the following reasons:

1) on the one hand, in the worldwide literature on Efficient Market Hypothesis (EMH) - see compilations of Fama (1970), Lo and MacKinlay (1999), Malkiel (2003), Szyszka (2003), Buczek (2005) - there is no differentiation between identifying information and describing information For the EMH hypothesis using the term "information" is intriguing since no information is used in weak form tests of this hypothesis. Filling this gap by bringing cybernetic information theory may resolve this problem of market efficiency at least in the weak form.
2) On the other hand there also exists so called non-equilibrium stream of theoretical and practical research regarding informational efficiency. It is based on entropy concepts, the status of which in the 1960s to 1980s is described by Nawrocki (1984). This stream of ideas resurfaced again in the 1990s and thereafter. Among this is the concept of *Relative Efficiency* (*RE*) by Golan, Judge and Miller (1996), Campbell (1997), Lo and Mackinlay (1999) and applied by Tambakis (2000). Relative efficiency is closely related to the topic of this paper which develops this concept further. Moreover papers of Jones, Redsun, Frye, Myers (1999, 2001, 2003) use Maxwell's Demon's idea related to physical entropy to examine market efficiency during continuous quotations.

Both approaches of relative efficiency and recent papers of Chen (2001-2004) on non-equilibrium generalized entropy theory of information present counter arguments to Grossman and Stiglitz's (1980) equilibrium based information theory, which is a paradox also described in a different way by Loistl and Veverka (2004) as based on information theory.

This paper is motivated from a desire to apply the cybernetic theory to test the weak form efficiency of price formation of the main index of Polish capital market – the Warsaw Stock Exchange (WSE) Index WIG. The cybernetic theory has yet been used to investigate weak form efficiency of market pricing processes.

The rest of the paper is organized as follows. Section 2 is a brief description of the cybernetic theory and Section 3 contains research hypotheses. The results from applying this test to selected market prices of WIG index are then presented in Section 4. The paper ends with conclusions in Section 5.

2. Information theory terms

Information is the relation between the states (elements) of the same set. These states may be of the same sort; i.e. relation between returns only within the given sample or the states may be of the same set, but of different kinds meaning the relation may be between price and turnover volume. Determination of relationships level between elements of the given set is carried out by calculation of pieces of information. How many pieces of information, then, may be contained in the exemplary set of past stock market rates of return? The following example will illustrate this point further.

Suppose that someone asks us a question: "How many pieces of information we receive with the sentence: WIG index has increased about 0.6% comparing to the yesterday's market close?" If you do not know the information theory, you could think that there is the only one information within such a sentence. But in fact you received 200 pieces of information regarding WIG index change value, based on the assumption that the change is quoted with the precision limited to 0.1 percent within the range ±10 percent. If we in turn suppose that the precision of quoting is limited to three possibilities indicating direction of index change, then it means we received three pieces of information: "WIG index increased", "it was not zero", "it did not decrease". Depending on the acceptable quoting precision, we may state that we received different amounts of information.

According to the information theory, the amount of information within any sentence may be expressed by the number of possible states of which this specific one is chosen. Thus the statement that WIG index increased about 0.6 percent contains 200 pieces of information, meaning by that no other of 199 possible states is occurring. With the use of lower level of precision (i.e. limited to 1 percent), the statement will be: "WIG index change is within the interval (0 percent to 1 percent)" or "WIG index increased about 1 percent". In this case such change contains 20 pieces of information because that is one of 20 possible states and all other of 19 choices are not occurring.

In general, we may conclude that N possible states contain N pieces of information. Possible choices may be treated as messages of specific meaning for investors while making investment decisions. However, more convenient than using the amount of N possible states one could also operate with logarithmic measure of this amount using as the base 2. It can be explained as follows.

Amount of information contained in the message (stock rate of return) chosen from the set of two possibilities is 2 can be presented as $N = 2^1$. The amount of information contained in the rate of return chosen from the set of 4 possibilities can be presented $N = 2^2$. Amount of information contained in the rate of return chosen from the set of 8 possibilities can be presented as $N = 2^3$, etc. Based on the amount of information contained in the stock return chosen from the set of N possibilities can be presented as $N = 2^H$. Applying logarithmic operation to both sides of this equation, we receive $H = \log_2 N$. As the amount of information, it was acceptable to treat H in the above formula. With the set of possibilities $N = 2$ expression H in the logarithmic measure equals $H = 1$, and it is treated as information unit called bit (*binary digit, binary information unit).* It is exactly the same information unit as it is commonly used in Information Technology science.

Recalling the previous example, the amount of information contained in stock market returns chosen from 200 possible states is 7.64 bits = $\log_2 200$. Thus the statement that WIG index changed by about 0.6 percent stores 7.64 bits of information. With an acceptance of less precise interval of return (1 percent), the possibility set equals 20, i.e. return is contained within interval (0-1 percent) and stores 4.32 bits of information. With application of only three states of possibilities (+, -, no change), the statement that "WIG index increased" stores 1.58 bit of information.

With each doubling of possible states, the amount of information increases to about 1 bit. As an example, assume that we possess the knowledge that the set of possibilities of WIG return occurrence on each session is limited to 1 percent within change limit ±10 percent, then $N = 20$. Suppose also, at this moment, that WIG return distribution is a uniform distribution and the probability of return occurrence in each of 20 intervals is equally likely. But we do not know the change in value of WIG index, because we had no access to information about it. How many questions optimally should be asked of anyone who is informed about WIG change (and is able to say only "yes" or "no") to receive this information? If we ask the questions randomly, then in the best case we would succeed in the first try and in the worst case we would have to ask 19 questions. Based on this we can conclude that we will have to ask 10 questions on average to identify the rate of return among the set of 20 possible ones with no guarantee that we will succeed.

In turn by using of information theory, we can be sure that we will succeed by asking 4 or 5 questions only. How is this possible? The following questions need to be asked:

1. "Is the rate of return negative or positive (including zero)?" The answer "yes, the rate of return is positive" eliminates half of the possibility set. This information stores 1 bit.
2. "Is this positive rate of return greater than 5 percent, or less than 5 percent?" The answer "yes, the rate of return is less than 5 percent" excludes again the half of possibility set. This information stores also 1 bit.
3. "Is the rate of return less than 3 percent or greater than 3 percent?" The answer "yes the return is lesser than 3 percent", excludes more than half the possibility set. If the answer was "yes, the rate of return is greater than 3 percent", then it would exclude less than half of possibility set. Depending on the answer, the identification of the given rate of return in the distribution may end up or may be one question longer. It is not possible to divide a set of possibilities into halves 2.5 percent because interval precision is 1 percent. This information stores more than 1 bit or less than 1 bit.
4. "Is the rate of return less than 1 percent or greater than 1 percent?" The answer "yes, the return is greater than 1 percent" excludes less than half of possibility set. So, in case of negative answer, one additional question must be asked. If the answer was: "yes, the rate of return is less than 1 percent"

the process of identification would finish, and the rate of return which was sought was within the interval 0-1 percent. It would mean that the rate of return contains in total 4 bits of information, because it was necessary to ask 4 identifying questions indispensable for identification of this rate of return from the set of 20 equally likely possibilities.

5. "Is the rate return less than 2 percent or greater than 2 percent?" The answer "yes, the return is greater than 2 percent", eliminates half of possibility set and the process of identification the rate of return in the distribution is finished. The return contains in this case in total 5 bits of information because there was necessary to ask 5 identifying questions indispensable for indication of this rate of return from the set of 20 possibilities. To sum up, in order to identify the rate of return from the set of 20 possibilities it sufficient to ask 4-5 questions. This is about a half less than average number of mentioned questions (10) what is simultaneously the interpretation of formula

$$H = \log_2 N = \log_2 20 = 4.32 \text{ bits.}$$

In that context, we can state that information theory enables optimizing of discovering the uncertainty.

The above examples illustrate the difficulty of identifying the rate of return in the distribution under the assumption that the given rate of return has occurred. In case of future returns which has not occurred the amount of bits of information may be the reflection of prediction difficulty level. Investor during process of prediction must eliminate first the most general premises: positive, negative or no change in the rate of return - excluding in this way -most of all possible states. If he or she is convinced that the premises support positive change, then he/she must determine more detailed premises: "what are pros and cons for the change greater than 5 percent and less than 5 percent?" After being convinced that the change below 5 percent is more reasonable, then he/ she has to search for more detailed premises: "what are the premises for the change below or above 2 percent?" While being convinced of the change more than 2 percent, then it remains only the choice of premise with the most detailed level (1 percent): "are there premises for the change greater or lesser than 1 percent?"

The number of possibilities is 20, so there is 1 chance against 20 that investor will predict correctly the rate of return. Probability of successful prediction of occurrence of the rate of return in one of 20 intervals is then 1/20 = 0.05. Alternatively with almost the certain probability 0.95 we may conclude that investor is not able to predict accurately the rate of return due to the possibility of making mistake in 19 against 20 cases possible. As a result, we may say that it is practically not possible to predict the rate of return with precision. If investor was to predict the rate of return occurrence with accuracy extended to 0.1 percent, his chance is like 1/200 = 0.005. So with such accuracy level 0.1 percent it is impossible the successful prediction even of one trading session change value, not speaking about the period longer than one day.

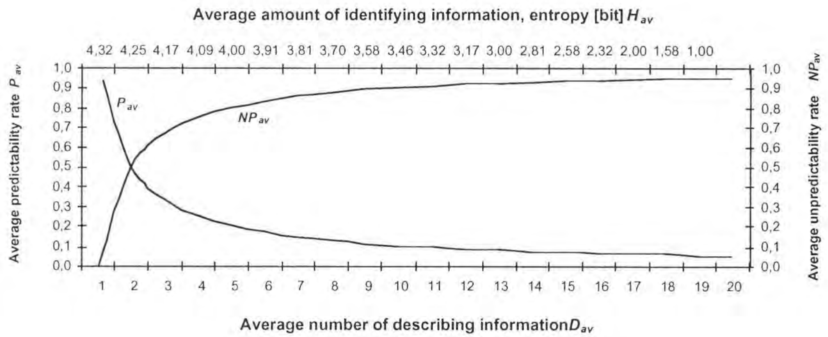
Relation of the one to the possibility set (1/3 or 1/20 or 1/200) will be called the *predictability rate* $P = 1/N$, whereas the *unpredictability rate* is the relation of the difference of maximal possibility set and number one to the maximal set of possibilities (2/3 or 19/20 or 199/200), $NP = (N-1)/N = 1-P$. Suppose now that the possibility set N is decreasing in time from the maximal set of possibilities i.e. 20 to the minimal set of possibility which is one. The question we would like to answer is: "how would either predictability rate P change or unpredictability rate NP change if possibility set was decreasing by value equal 1?" Results are presented in the Table 1 and Figure 1.

Table 1: Changes in predictability and unpredictability rate together with changes of possibility set N and the amount of information H reflected by the rate of return

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
NP = (N-1)/N	0	0,50	0,67	0,75	0,80	0,83	0,86	0,88	0,89	0,90	0,91	0,92	0,92	0,93	0,93	0,94	0,94	0,94	0,95	0,95
P = 1/N	1,00	0,50	0,33	0,25	0,20	0,17	0,14	0,13	0,11	0,10	0,09	0,08	0,08	0,07	0,07	0,06	0,06	0,06	0,05	0,05
H	0	1,00	1,58	2,00	2,32	2,58	2,81	3,00	3,17	3,32	3,46	3,58	3,70	3,81	3,91	4,00	4,09	4,17	4,25	4,32

Source: author's own calculations.

Figure 1: Dependency of predictability and unpredictability rate on the number of the possibility set N and amount of information H



For the predictability rate to achieve the value $P = 0.1$ the possibility set N has to be increased up to 10. Thus the choice from the half (50 percent, $N = 10$) of the total possibility set (20) means decrease the predictability rate only to $P = 0.1$ and unpredictability rate equal to $NP = 0.9$. Further increase about half of possibility set (from 10 to 20) causes only little 10 percent decrease of total predictability P being 0, and NP equal to 1. Similar considerations may be carried out concerning the unpredictability rate. These are the following implications of this dependency:

1. Changes of predictability rate are non-linear. Change of possibility set about the same value (i.e. 1) starting from different level of possibility set indicates the different changes in predictability rate i.e. the change in possibility

set about 1 from the level $N = 2$ to $N = 1$ causes the increase of predictability rate from $P = 50$ percent to $P = 100$ percent, and the change of possibility set also about 1, but from the possibility level $N = 10$ to $N = 9$ indicates increase of predictability rate only from $P = 8$ percent to $P = 9$ percent.

2. The most intensive downfall of predictability rate P is by the increase of possibility set N from 1 to 2. That means also that the downfall of unpredictability rate is the most by the decrease of possibility set N from 20 to 19.
3. If the possibility set N equals to 1, then there is occurring the state of full predictability and simultaneously no choice of other possibility. If the possibility set N is equal to 20 there is occurring the state of full unpredictability and simultaneously the maximal possibility of choice.
4. In order to assure the very high predictability or unpredictability rate equaling respectively to 10 percent or 90 percent, it is sufficient to enable the choice 1 of 10 possibilities, so half of total possibility set $N = 20$.

Uncertainty level can be expressed also by other measures than the above mentioned predictability P or unpredictability NP. We may ask for example "how the entropy H and number of possibilities N relate to its maximum or minimum values?" We assumed earlier that probabilities of return occurrence in each of 20 intervals are equal giving us the results: $N = 20$ and $\log_2 20 = 4.32$ bits. If we accept that N and H are maximum values in this case, then in order to observe changes in relation to their maximal values, the only possibility is to decrease them. Assume again that number of possibilities N changes in time scale. If in one interval it was not possible for any occurrences and the other 19 probabilities were still equally likely, the number of possibilities would be $N = 19$ and the number of information would be $\log_2 19 = 4.25$ bits. We have then, apart from measures H, D, P, and NP, additional four possibilities emerge to express uncertainty: $C_N = 19/20 = 0.95$ or $C_H = 4.25$ bits/4.32bits $= 0.984$; $R_N = 1 - C_N = 0.05$ or $R_H = 1 - C_H = 0.016$.

What does it mean in this case? Relations marked with letters C and R will be called compression and redundancy respectively, both related also to possibilities as number N and entropy H respectively. In general, compression reflects the extent current returns results from possibilities of free choices of investors (whether deliberate or accidental it is not discussed here due to high-level ("black-box") analysis of this process). In turn redundancy means to what extent current rate of return results from the past structure (relationships) of returns distribution.

Compression $C_N = 95\%$ means that 95 percent (i.e. 19 intervals) of all possibilities (20 intervals) of rate of return occurrences are important for the process of returns generation. As a result the process (number of possibilities) of returns generation can be compressed to the level of 95 percent its maximal value (from

20 to 19). Thus, total uncertainty of return process generation resulting from existence of 20 intervals may be reduced about 5 percent (i.e. 1). The percentage of reduced intervals will be called the redundancy of possibility set R_N. Redundancy R_N = 5% means that 5 percent of all uncertainty of the process (1 interval) is not important for the process of rate of returns generation, therefore this 1 interval may be treated as redundant.

In our example, according to its assumption, all 20 intervals were important for rate of return generation (uniform distribution), therefore no compression to the lower level was possible (C_N = 20/20 = 1 or 100%)) and no redundancy of possibility set were possible (R_N = 1 - 1 = 0 or 0%).

In turn, compression C_H = 98.4% means that 98.4 percent (i.e. 4.25 bits) of total possible number of questions or premises (4.32 bits) is important and optimal for identifying or predicting the given return in the given interval of distribution. As a result, the process of return generation can be compressed only a little (0.07bit) to the level of 98.4 percent, its maximal value (from 4.32 bits to 4.25 bits). Similarly the percentage of reduced questions or premises will be called the redundancy of information R_H. Redundancy R_H = 1.6% means that 1.6 percent of all uncertainty process (0.07 bit, so "0.07 part of question" - in fact no question) is not important for the process of rate of return generation, therefore in fact no redundancy in number of questions or premises occurs here. Thus in this case still the 4 of 5 questions or premises must be set to identify the return in the distribution.

As for our example, according to its assumption, all 4-5 questions or premises were important to identify or predict rate of return in the given interval of the uniform distribution, therefore no compression to the lower level was possible (C_H= 4.32bits/4.32bits = 1 or 100%) and no redundancy of questions or premises were possible (R_H = 1 - 1 = 0 or 0%).

Further decrease in the number of possibilities *N* by value equal 1 causes the respective changes of compression and redundancy: C_D, C_H, R_D, and R_H, which are presented in Table 2, figures 2 and 3.

What emerged now is linearity and nonlinearity of changes in these relative measures of uncertainty dependent on the compression and redundancy possibility set *N* or amount of information *H*. When all 8 derived measures of uncertainty are being compared, we observe that for the given possibility set (for

Table 2: Changes in compression and redundancy together with changes of possibility set *N* and the amount of information *H* reflected by the rate of return

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	0	1	1.58	2	2.32	2.58	2.81	3	3.17	3.32	3.46	3.58	3.7	3.81	3.91	4	4.09	4.17	4.25	4.32
C_D	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
C_H	0.00	0.23	0.37	0.46	0.54	0.60	0.65	0.69	0.73	0.77	0.80	0.83	0.86	0.88	0.91	0.93	0.95	0.97	0.98	1
R_D	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.45	0.40	0.35	0.30	0.25	0.20	0.15	0.10	0.05	0
R_H	1.00	0.77	0.63	0.54	0.46	0.40	0.35	0.31	0.27	0.23	0.20	0.17	0.14	0.12	0.09	0.07	0.05	0.03	0.02	0

Figure 2: Dependency of compression and redundancy on the number of the possibility set N

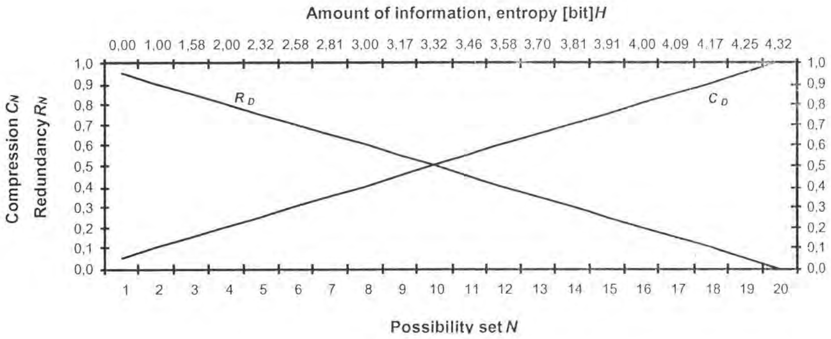
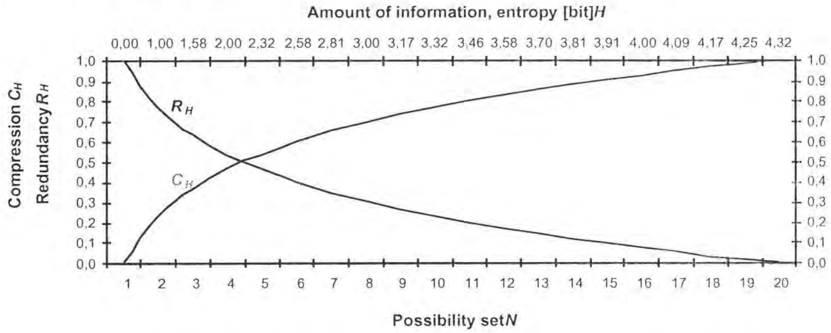


Figure 3: Dependency of compression and redundancy on the amount of information H



example $N = 2$) redundancy of possibility set R_N is greater than redundancy of amount of information R_H, which in turn is greater than predictability P ($R_N > R_H > P$). Analogically the same relation is between unpredictability NP, compression of amount of information C_H, and compression of possibility set C_N ($NP > C_H > C_N$). In view of this, there are at least three different levels of uncertainty for the given situation, and it is important to indicate about which measure the discussion is about, to avoid misunderstandings regarding, for example, trueness of specific level of uncertainty.

Moreover, only predictability P and unpredictability NP, compression C_H and redundancy R_H of amount of information change nonlinearly with increase or decrease of uncertainty measured by possibility set N. Only compression C_N and redundancy R_N of possibility set are changing linearly with increase or decrease of possibilities set N.

Predictability and unpredictability, redundancy and compression of possibility set are in turn not defined in extremes: in full certainty or full uncertainty. It results directly from the definition of information which exists only if there are at least two elements and one relation between them. In the case of only one possibility $N = 1$ and $H = \log_2 1 = 0$ bits, so the state of full predictability $P = 1$ contains no amount of information expressed in bits. Consequently compression ($C_H = H/H_{max}$) and redundancy of amount of information ($R_H = 1 - C_H$) cannot exist as $H = 0$, i.e. everything is known before such event occurs. But compression C_N of possibility set cannot be zero as 1 possibility at least must remain and represent fully predictable occurrence. Therefore redundancy R_N cannot be full (100%) because compression C_N cannot be zero. If there would be no possibility ($N = 0$), then no occurrence exists for prediction, and no uncertainty level measurement makes sense. It is now clear that the amount of information, i.e. identifying information measured in bits cannot be the only one type of information.

Similarly predictability P cannot be equal to zero as zero predictability does not make sense, because it would imply no occurrence to be predicted ($P = 0$ when $P = 0/N$) or situation where one of infinite number of possibilities is predicted ($P = 0$ when $1/\bullet$). Consequently unpredictability NP cannot reach maximal level 1 (100%) as it requires also one possibility at least to exist, to be unpredictable. Otherwise, it would cause necessity of division by zero: $(N-1)/N$. We also are not able to apply common entropy formula to measure predictability P because for full predictability $H = 0$, situation like $P = 1/H$ is not possible. We need then possibility set N to determine full predictability, instead of entropy; because N must be equal to at least 1 for full predictability which indicates 100 percent predictability.

The above consideration has been carried out with the assumption that return distribution is the uniform one. Thus, the terms: number of information H (expressed in bits) and possibility set N were sufficient. In the case of empirical rate of return distributions there is a necessity to apply average measures of possibility set and average number of information. The average real possibility set which reflects differentiation of distribution is calculated by the measure of average number of describing information D_{av}. The average real amount of information reflecting the distribution is expressed by the measure of average amount of identifying information called also entropy H_{av}. The average number of describing information D_{av} is analogous to the set of equally probable possibilities N, and the average amount of identifying information H_{av} is analogous to number of information H, in the above example. Formal definitions of describing and identifying information are the following.

Describing information (absolute measure) – is an information among the least possible number of pieces of information indispensable to define any message (i.e. rate of return) in the given information set (return quoting possibilities, samples). Average number of observations can be expressed as the geometric mean of the average set of possibilities occurring in the given interval of the distribution. The formula is as follows:

$$D_{av} = \left(\frac{n}{n_a}\right)^{\frac{n_a}{n}} \left(\frac{n}{n_b}\right)^{\frac{n_b}{n}} \cdots \left(\frac{n}{n_m}\right)^{\frac{n_m}{n}} = \left(\frac{1}{p_a}\right)^{p_a} \left(\frac{1}{p_b}\right)^{p_b} \cdots \left(\frac{1}{p_m}\right)^{p_m} \qquad (1)$$

where,

p_a : probability of rate of return occurrence in the given interval,
m : the number of intervals of the return distribution i.e. (0.1%-1%), (1.1%-2%),
$n_{a,}\ n_{b,} \ldots n_m$: absolute frequency of rate of return occurrence in the defined interval, and
$a, b, \ldots, m$; n: return sample size.

The average number of described observations represents the real set of possibilities resulting from the rate of return distribution. During calculation of the number of observations, all returns within the given interval will be treated as one class of returns and treated as they were the same. Geometric mean is applied here for the reason of relatively high differentiation of population of returns. Relative number (compression) of observations is the relation of average number of describing information to the maximally possible number of describing information: $C_D = D_{av}/D$. C_D is analogous to C_N, in the above example. This is in fact is the compression level of possibility set including its distribution. Redundancy of describing information is expressed $R_D = 1 - C_D$. C_D is analogous to C_N, and R_N is analogous to R_D in the above example.

Identifying information (absolute entropy): this relates to information among the least possible number of pieces of information indispensable to define the state chosen (differentiated) from the respective information set.

$$H_{av} = \sum_{a=1}^{m} \frac{n_a}{n} \log_2 \frac{n}{n_a} = \sum_{a=1}^{m} \frac{1}{p_a} \log_2 p_a = -\sum_{a=1}^{m} p_a \log_2 p_a \text{ , } m \neq n,\ D \neq n \qquad (2)$$

where,

p_a : probability of rate of return occurrence in the given interval,
m : number of defined intervals of return distribution i.e. (0.1%-1%, 1.1%-2%),
$n_{a,}\ n_{b,} \ldots n_m$: absolute frequency of rate of return occurrence in the given interval $a, b, \ldots, m$, and
n : return sample size.

If the number of intervals equals to returns' sample size then $m = n$, to $n_a = 1$, $n_b = 1$, $n_m = 1$, $D = n$, $D_{av} = D$, $H_{av} = H$, and formula (2) is reduced to:

$$H_{av} = \sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{n}{1} = \sum_{i=1}^{n} \frac{1}{n} \log_2 n = \log_2 n = \log_2 D = H \qquad (3)$$

Relative number (compression) of identifying information (relative entropy) refers to the relation of the average amount of identifying information to the maximally possible amount of identifying information: $C_H = H_{av}/H$. Redundancy of identifying information means that $R_H = 1 - C_H$. These relative measures of identifying information remain the same as earlier ones in the example.

In order to choose the one rate of return from the defined set, there is need for a criterion of choice and also the possibility of compliance verification with this criterion. Choice criterion may be the time (definition of the moment of chosen rate of return occurrence), space (definition of place of rate of return occurrence), sequence (definition of rate of return number), etc. With regard to the given return equal to n, we may identify: return occurrence in the defined sequence between 1 and n; return occurrence within the given period; return occurrence in the defined rate of return distribution interval.

In this paper the last criterion will be examined. Let us formulate now the average measures of predictability and non-predictability analogous to the mentioned predictability and unpredictability rates. Average predictability rate identifies the probability indicated by the relation of one to average number of describing information of the defined return distribution: D_{av}. $P_{av} = 1/D_{av}$. Average unpredictability rate refers to the probability indicated by a relation of the difference between average number of describing information and one to number of average describing information of the of the defined return distribution. $NP_{av} = (D_{av} - 1)/D_{av}$. The relations between all measures remain also the same, i.e.: $R_D > R_H > P_{av}$ and $NP_{av} > C_H > C_D$.

Distribution uncertainty region is the difference between maximal and minimal average amounts of entropy and the difference between the maximal and minimal average numbers of describing information in the defined period: $H_{max} - H_{min}$, $D_{max} - D_{min}$. Based on the above parameters of average uncertainty, empirical uncertainty of WIG index distribution will be examined.

To sum up the discussion in this section, this consideration fulfilled its goal. It indicated the reason for differentiation of identifying and describing information, based on entropy and possibility set, respectively.

3. Research hypotheses

Based on the definitions presented in previous section empirical research for predictability can be conducted. The main issue is to decide which level of predictability could be treated as border for informational efficiency and thus which for non-efficiency of capital market in the weak form. It was decided that average predictability P_{av} less than 50% or average unpredictability NP_{av} greater than 50% will be treated as the beginning of the weak form efficiency frontier. On the contrary, predictability P_{av} greater than 50% or average unpredictability NP_{av} less than 50% will be treated is non-efficiency of market in the weak form, The reason is as follows: if we have the chance less than 1 to 2 of successful prediction or of making the error during prediction, then successful prediction with the given precision rate is almost sure and occurs more often than unsuccessful prediction. However, we must be aware that 50% prediction P_{av} is not equal 50% probability of loss or win (for example, in case the distribution is asymmetric).

For the purpose of this research the example from previous section has been extended to about three additional intervals: "change greater than 10%", "change less than -10%", "no change") as there were several empirical returns exceeding 10% or -10%, and that gives 23 intervals ($D_{av} = N = 23$).

Using again predictability scale, it is possible to assign 50 percent average predictability or non-predictability to its corresponding values R_D, R_H, or C_D, C_H what is shown in Table 3.

Calculations presented in Table 3 will enable us to present graphically the hypothesis of this research. Hypothesis will be true, which means "market is efficient" if values of P_{av}, NP_{av}, R_D, R_H, C_D, C_H are outside rectangular area marked in Figure.4 and Figure 5 as the letter "H". The border values separating efficient market from non-efficient market are marked in bold in Table 3. Due to such definition of efficiency we can see that there is, only a little space for non-efficiency (within

Table 3: Changes in predictability and unpredictability rate for the respective values of *H, D,* C_H, C_D, R_H, R_D.

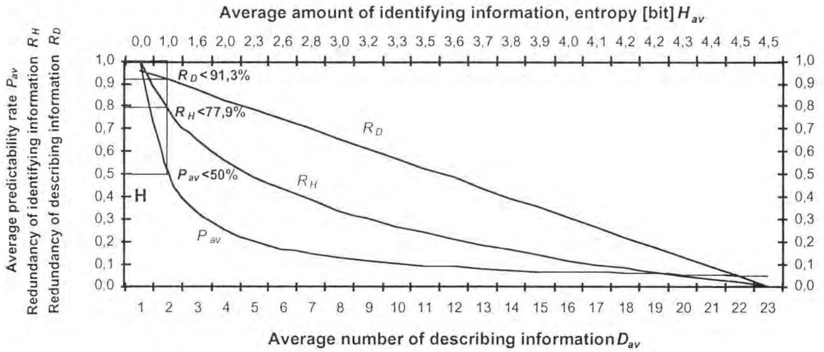
D_{av}=N	H_{av}	P_{av}	NP_{av}	C_H	R_H	C_D	R_D
1	-	100.0%	0.0%	0.0%	100.0%	4.3%	95.7%
2	**1.00**	**50.0%**	**50.0%**	**22.1%**	**77.9%**	**8.7%**	**91.3%**
3	1.58	33.3%	66.7%	35.0%	65.0%	13.0%	87.0%
4	2.00	25.0%	75.0%	44.2%	55.8%	17.4%	82.6%
5	2.32	20.0%	80.0%	51.3%	48.7%	21.7%	78.3%
6	2.58	16.7%	83.3%	57.1%	42.9%	26.1%	73.9%
7	2.81	14.3%	85.7%	62.1%	37.9%	30.4%	69.6%
8	3.00	12.5%	87.5%	66.3%	33.7%	34.8%	65.2%
9	3.17	11.1%	88.9%	70.1%	29.9%	39.1%	60.9%
10	3.32	10.0%	90.0%	73.4%	26.6%	43.5%	56.5%
11	3.46	9.1%	90.9%	76.5%	23.5%	47.8%	52.2%
12	3.58	8.3%	91.7%	79.3%	20.7%	52.2%	47.8%
13	3.70	7.7%	92.3%	81.8%	18.2%	56.5%	43.5%
14	3.81	7.1%	92.9%	84.2%	15.8%	60.9%	39.1%
15	3.91	6.7%	93.3%	86.4%	13.6%	65.2%	34.8%
16	4.00	6.3%	93.8%	88.4%	11.6%	69.6%	30.4%
17	4.09	5.9%	94.1%	90.4%	9.6%	73.9%	26.1%
18	4.17	5.6%	94.4%	92.2%	7.8%	78.3%	21.7%
19	4.25	5.3%	94.7%	93.9%	6.1%	82.6%	17.4%
20	4.32	5.0%	95.0%	95.5%	4.5%	87.0%	13.0%
21	4.39	4.8%	95.2%	97.1%	2.9%	91.3%	8.7%
22	4.46	4.5%	95.5%	98.6%	1.4%	95.7%	4.3%
23	4.52	4.3%	95.7%	100.0%	0.0%	100.0%	0.0%

rectangular area marked as "H") and a lot of space for efficiency. The specific hypotheses are as follows:

1a) Polish market is efficient in the meaning of average predictability rate P_{av} (precision 1%) being less than 50% (P_{av}<50%) with redundancy of identifying and describing information non-exceeding respectively: $R_H < 77.9\%$ and $R_D < 91.3\%$ which is shown in Figure 4.

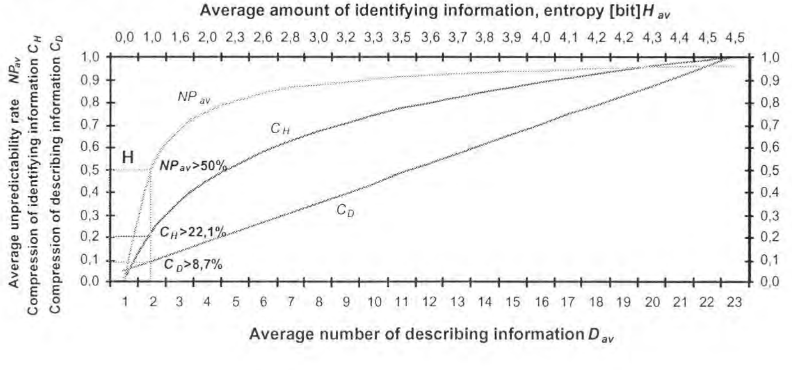
1b) Polish market is efficient in the meaning of average unpredictability rate NP_{av} (precision 1%) being grater than 50 percent (NP_{av}>50%) with compression of identifying and describing information exceeding respectively: $C_H > 22.1\%$ and $C_D > 8.70\%$ what is shown in Figure 5.

Figure 4: Graphical illustration of hypothesis 1a



Despite the fact that such formulation of hypotheses privileges positive verification of weak market hypothesis and supports EMH followers, it provides additional important observation resulting from the fact that uncertainty level is not changing proportionally with predictability change. Consequently, we have

Figure 5: Graphical illustration of hypothesis 1b



to accept the opinion of opponents of EMH regarding price changes is not being so much random because even low level of uncertainty gives high level of unpredictability. This will be explained in detail in Section 4.

4. Methodology, findings and analysis

The subject of this research is one-session WIG index rates of return distribution with 3,002 observations forming the samples size (i.e. over the period 16-04-1991 to 04-02-2005). Rates of return have been calculated using logarithmic value described by the formula (4) (dividends were not included, due to non-regularity of paying them on WSE):

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{4}$$

The price P_t used in the calculation of rate of returns, is the day's close value of WIG index on the given trading session. Rate of return distribution is built with interval precision 1 percent (broadness of interval) and set of possibilities mainly within the range ±10 percent plus two extreme intervals exceeding this range). If in the given interval no sample has been observed then in identifying information (entropy) formula number "zero" was put and for describing information formula number "1" was put. Such constructed distribution is shown in Figure 6. The calculation of uncertainty parameters is presented in the tables 4-6.

Entropy value for each interval has been calculated as follows: -0.004_$\log_2$0.004 = 0.032. Total entropy value H_{av} = 3.02 bit is a sum of entropy of the given interval. Describing information for each interval is a result of raising of the reciprocal relative frequency to the power of relative frequency, i.e. $250^{0.004}$ = 1.022. Total value of describing information D_{av} = 8.23 [number of units] is a product of describing information in each interval. Results has been confirmed by the calculation of formula $\log_2 D_{av} = \log_2 8.12 = 3.02$ bits.

Figure 6: WIG index returns distribution of 3,002 trading prices

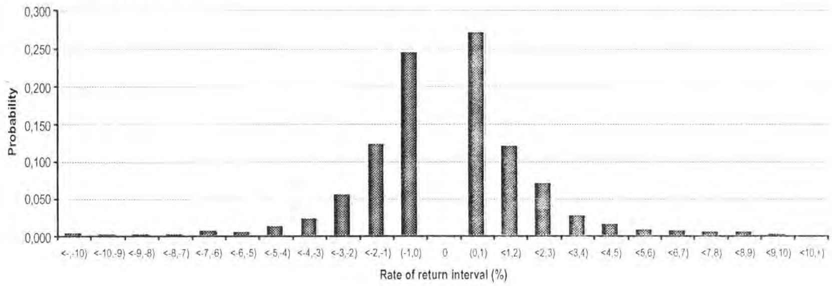


Table 4: Entropy and describing information of empirical distribution of WIG returns

Rate of return interval	Frequency (A)	Relative frequency (B)	Entropy $C=-B\log_2 B$	Inversion of relative frequency (D = 1/B)	Describing information ($E=D^B$)
⟨-,-10%)	12	0.0040	0.032	250	1.022
⟨-10%,-9%)	7	0.0023	0.020	429	1.014
⟨-9%,-8%)	3	0.0010	0.010	1 001	1.007
⟨-8%,-7%)	5	0.0017	0.015	600	1.011
⟨-7%,-6%)	17	0.0057	0.042	177	1.030
⟨6%,-5%)	15	0.0050	0.038	200	1.027
⟨-5%,-4%)	36	0.0120	0.077	83	1.054
⟨-4%,-3%)	65	0.0217	0.120	46	1.087
⟨-3%,-2%)	165	0.0550	0.230	18	1.173
⟨-2%,-1%)	369	0.1229	0.372	8	1.294
⟨-1%,0%)	732	0.2438	0.496	4	1.411
0%	1	0.0003	0.004	3 002	1.003
(0%,1%)	810	0.2698	0.510	4	1.424
⟨1%,2%)	356	0.1186	0.365	8	1.288
⟨2%,3%)	208	0.0693	0.267	14	1.203
⟨3%,4%)	79	0.0263	0.138	38	1.100
⟨4%,5%)	46	0.0153	0.092	65	1.066
⟨5%,6%)	23	0.0077	0.054	131	1.038
⟨6%,7%)	17	0.0057	0.042	177	1.030
⟨7%,8%)	14	0.0047	0.036	214	1.025
⟨8%,9%)	13	0.0043	0.034	231	1.024
⟨9%,10%)	8	0.0027	0.023	375	1.016
$D_{max} = 23$ $H_{max} = 4.52$ bit	Population: 3002		Sum: $H_{av} = 3.02$		Product: $D_{av} = 8.12$

The calculations indicate that WIG rates of return for the sample of 3,002 prices contain (or reflects) on average 3.02 bits of identifying information (entropy value) and describing information of number 8.12. This means that it is indispensable to ask on average 3 questions in order to identify the given return in the 8-9 intervals of the above distribution. Identification of the single return in the given interval of the above distribution indicates that, simultaneously, the returns are not occurring in the remaining 7-8 most probable intervals of this distribution ($2^{Hav} - 1 = 2^{3.02}-1$). Those 3 questions to be asked are the following:

1. "Is the rate of return is positive or negative?" If the answer is: "positive", then 5 more probable intervals and 6 less probable ones of negative returns are excluded from further identification.
2. "Is this positive rate of return in the interval (0 percent-2 percent) occurs in the interval ·2 percent-10 percent)?" This question is formulated so, for the reason that the most probable occurrence of the positive rate of return is within intervals (0 percent-2 percent). If the answer is: "yes, it is in the interval (0 percent-2 percent)", then excluded are 8 intervals of positive returns from further identification.

Table 5 : Predictability, compression, redundancy of empirical distribution of WIG returns

Relative entropy, compression of identifying information	Compression of describing information	Average predictability rate	Average unpredictability rate	Redundancy of identifying information	Redundancy of describing information
$C_H = H_{av}/H_{max}$	$C_D = D_{av}/D_{max}$	$P=1/D_{av}$	$NP_{av}=1-P$	$R_H=1-C_H$	$R_D = 1-C_D$
C_H=3.02/4.52=66.8%	C_D= 8.12/23=35.3%	P_{av} = 1/ 8.12 =12.3%	NP_{av} = 1 -12.3% = 87.7%	R_H=1- 66.8%=33.2%	R_D=1-35.3%= 64.7%

Table 6: Predictability, compression, redundancy changes in time of empirical distribution of WIG returns

Maximal empirical entropy and number of describing information	Maximal empirical compression of describing information	Minimal empirical average predictability rate	Maximal empirical average unpredictability rate	Uncertainty region of identifying information	Uncertainty region of describing information
$H_{emp\ max}$ $D_{emp\ max}$	$C_{D\ emp\ max}$	$P_{emp\ min}$	$NP_{emp\ max}$	$H_{emp\ max} - H_{av}$	$D_{emp\ max} - D_{av}$
4.07 bit 16.76	= 16.76 / 23 =72.9%	= 1/ 16.76 = 6.0%	NP = 1- 6% = 94%	4.07-3.02 = 1.05 bit	16.76 - 8.12 = 8.64

3. "Is the rate of return occurring in interval (0 percent-1 percent) or ·1 percent-2 percent)?" If the answer is: "yes, it is in the interval (0 percent-1 percent)", then there is excluded one of two most probable intervals of the rate of return occurrence from further identification and the process ends. In this case the return contains 3 bits of information. If the answer was that the return occurs in the intervals ·2 percent-9 percent) or ·2 percent-10 percent), it would be indispensable to ask additional one or two identifying questions. In such a case, the rate of return would contain 4-5 bits of information. But we may conclude from the distribution observed that such occurrences are relatively rare.

Compression of the observations of WIG index rate of return empirical distribution contains on average basis C_D = (8.12)/23 = 0.353 = 35.3 percent (redundancy R_D = 64.7 percent) of the maximal uncertainty of its uniform distribution for which the rate of average unpredictability is 87.7 percent = (8.12-1)/8.12, or average predictability rate equals to 1/(8.12) = 12.3 percent meaning they explain on average 12.3 percent of the current returns. Seven chances against 8 will be an error while identifying of WIG index rate of return with 1 percent precision. Thus, other factors than past returns explain current WIG changes. The results are shown in Figure 7.

Redundancy R_D means that the current process of rate of return generation explains in 64.7 percent past rate of returns giving 12.3 percent of average predictability. Current possibility set reflects in 64.7 percent past possibilities set (number of intervals with non-empty samples). Thus the other factors than past returns explain in 35.3 percent current WIG changes. This indicates directly that it is not possible to find full explanation of current returns in the past return distribution: this is equivalent to the interpretation of coefficients in the serial correlations tests used in EMH. Consequently, "full reflection of past prices in current prices" is neither necessary nor sufficient condition for weak efficiency form. As a result, weak efficiency may be even treated as an assumption: based on past returns only it is hardly or not possible to predict with 1 percent precision current rate of returns. But such new definition of weak efficiency in the meaning of predictability is independent based on full reflection of past prices in currents ones. The point is that "full reflection" refers to the causes of uncertainty or unpredictability and causes of correlations which cannot be determined on the level of "black-box" analysis which is in fact any test regarding weak form of efficiency.

This evidence shows that the high rate of unpredictability does not require high level of uncertainty, i.e. namely low level of relationships between returns. For WIG index rate of return distribution even 35.3 percent of its maximal possible uncertainty guarantees a very low average rate of unpredictability 12.3 percent. This is a very significant conclusion as it clarifies the academic and practical controversies about the rate of return character being predictable or as to whether returns follow random a walk. Differentiating two measures of uncertainty, average predictability rate and compression (redundancy) of rate of returns affords a statement that the question about predictability rate and random walk of return is not adequate and not precise. In fact such question contains at least the question about the average rate of return predictability rate and the average return compression (redundancy) level.

We now present the analysis on the WIG index rate of returns distribution uncertainty changes in time scale. Changes in average entropy, describing information and predictability of WIG index rate of return occurrence in the given interval, has been calculated for the period between 23 of April 1991 and 4 of February 2005: i.e. for the population from 1 to 3,002 increasing by 1. Changes

Figure 7: Relationships between uncertainty parameters (H_{av}, D_{av}), the average rate of predictability and compression of WIG index return

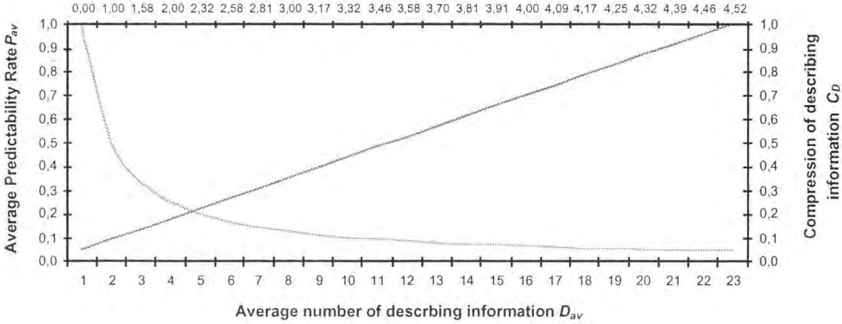
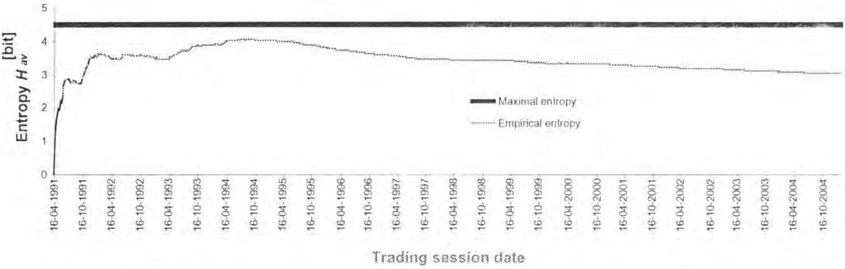


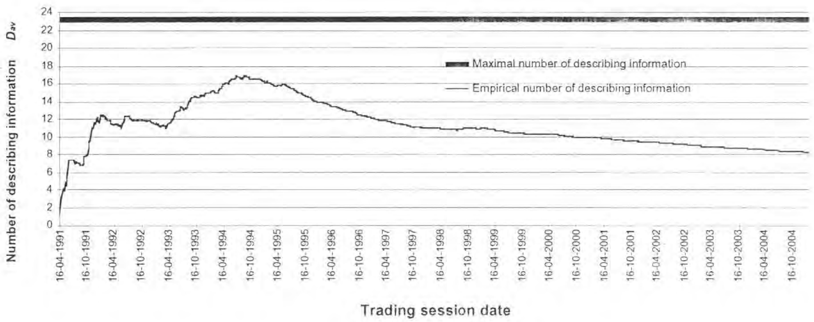
Figure 8: Changes in average entropy H_{av}: WIG returns



in uncertainty parameters are shown in the Figures 8-10. Figure 9 is a graph of changes in entropy of one trading session WIG rate of return distribution together with population increase and in comparison with the maximal entropy of WIG uniform distribution. Figure 10 presents the changes in the number of describing information of WIG index returns in comparison with the maximal possible number of describing information for uniform distribution. Figure 11 illustrates the changes of average predictability rate for the presented in figures 9-10 – the level of entropy and describing information corresponding with the given sample's number.

Maximal empirical entropy of WIG index distribution with interval precision of 1% has reached the level of 4.07 bits and 16.76 describing information and it occurred in the trading session No. 406 as of date 14-09-1994. What is interesting, it happened almost 90 trading sessions after very violent downfall of stock prices which started on trading session No. 318, and two weeks before the date 20-09-1994 of implementation of each-day quotations on WSE. Since trading session 406 to the date 04-02-2005, the level of uncertainty has fallen doubly from 4.07 to 3.02 bits and from 16.76 to 8.12 describing information. This is an important conclusion against common opinion that market efficiency in the weak form is rather growing than decreasing, in general. From Figure 9 and Figure 10, it is visible that the only important difference between identifying and describing

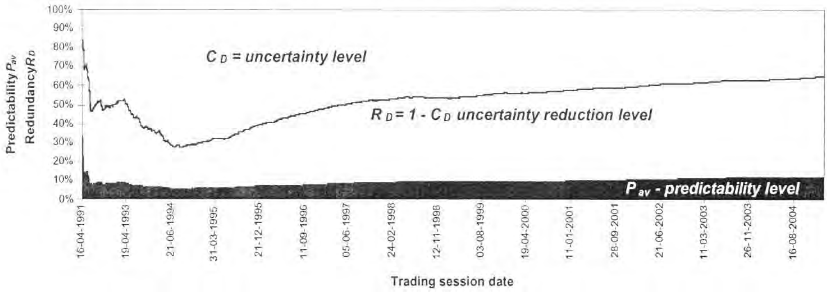
Figure 9: Changes in number of describing information D_{av} WIG returns



information is $H_{av} = \log_2 D_{av}$.

Simultaneously, within the last 11 years almost (1994-2005), the average predictability rate of the WIG returns occurrence in the given interval of the distribution has increased from the level of 6 percent to 12.3 percent. Analogically redundancy of average possibility set R_D increased from 27.1% to 64.7 percent

Figure 10: Changes in average predictability rate P_{av} of WIG index in time.



meaning that 7 intervals have been minimally redundant (23-16) in WIG distribution in 1994 and its number increased to 15 redundant intervals (23-8) in 2005, which were not important for generation of WIG returns. Current WIG returns (in 2005) result is twice more than in the past structure of returns in 1994. In turn, the compression of average possibility set (C_D) with relation to its maximal number has decreased from 72.9% (16/23) to 35.3% (8/23) meaning that current WIG returns result is twice less from free choice of investors, expressed by number of possible intervals. Despite significant – doubling of decrease of uncertainty level from 16.76 to 8.12 (slightly over 1 bit) - it still gives a low average rate of predictability. It especially reflects non-linear dependence between probability, entropy and describing information.

Distribution uncertainty region within the years 1994-2005 reached 1.05 bit and 8.64 of identifying and describing pieces of information, respectively. For the reason that predictability P_{av} is not expressed in logarithmic measure the compression and redundancy of identifying information C_H and R_H cannot be directly compared to it. Uncertainty measured by C_H has also changed doubly, i.e. decreased from 90% (4 of 4-5 questions) in 1994 to 66.8% (3 of 4-5 questions) in 2005 meaning the number of questions or premises throughout almost 11 years have been decreased by 1. Consequently R_H has also increased from 10 to 33.2 percent meaning no question or premise to identify return in the WIG distribution was redundant with relation to its maximal value in 1994 and 1 question was redundant in 2005 for identification of the given return in distribution.

Parameters C_H and R_H are measure of another picture of uncertainty related to number of questions and premises (number of details level), but they do not measure prediction errors (NP_{av}) or successfulness (P_{av}) resulting from average possibility set D_{av} nor does the measure indicate the possibility set D_{av} to its maximal value D_{max}. In view of this, opponents of weak form efficiency think rather in terms C_D than C_H while saying market is "not so much random". Therefore, it seems to be adequate focusing in further analysis on uncertainty level measured by C_D being compared to predictability P_{av}. It is shown in Figure 12.

Interpretation of the average predictability rate is relative in the following meaning: this average predictability rate will be P = 12.3% if the uncertainty level (H_{av} = 3.02 bits D_{av} = 8.12) remains on the same level from the past on the next trading session. If uncertainty level increases in time then average predictability rate of P_{av} = 12.3% will be incredible in the future due to insufficiency (scarcity) of current information with relation to the future ones. If entropy and describing information levels decrease in time then average predictability rate P_{av} = 12.3% will also be incredible for future for the reason that current information is redundant with relation to the future ones. As of date 04-02-2005 the picture of past returns credibility in case of WIG index is such that the past returns contain redundant information because the entropy of WIG index distribution is consequently decreasing. This means that there is still an increase of relationships level between returns. If the trend of WIG entropy changes and will be increasing then it will be the sign of generation of new relationships between returns and simultaneously the symptom of decreasing role of historical dependencies.

For the informational efficiency the redundancy or scarcity of past returns has the implication that investors are not able to react adequately to previous information about the returns because they are equivocal so there cannot exist only rational expectations about future changes. In fact, it is not possible under the situation – on such level of analysis - that past prices information can be ever effectively consumed, or reflected in current prices, as we are not able to explore causes of changes, and therefore, again, "full reflection of past prices in the current ones" - as a proposal of such a cause - is neither required nor sufficient condition for unpredictability of stock returns, and hence for weak form of efficiency.

Formation of entropy and describing information in time enables also the identification of population sample size for a given distribution such that a return distribution may be treated as a credible basis to the future prediction using the past. We take for this study, for the Polish capital market as reflected generally by WIG index, the maximal empirical entropy of WIG distribution occurred on 14-09-1994. Thus starting with a sample of trading session 406, the distribution may be treated as credible register of changes in the past.

Summing up for the reason that entropy and describing information is increasing, stable, or decreasing in time, there may be two sorts of incredible returns data from the past and one credible data from the past. Past returns may be invalid for future returns prediction as they reflect (contain) insufficient (too few) information about future quotations – i.e. the current entropy of the return distribution increases in relation to the preceding one. Past returns may be also irrelevant for future prediction as they reflect redundancy (too much) information (of which part may be out-of-date, redundant, contradictive, misleading) and in that case the current entropy is decreasing in relation to the preceding one. Past returns are valid for prediction of the next ones when they contain the same (similar) information about past and present returns (quotations). In this context, we may state that an increase in the number of samples does not always cause an increase in the average predictability rate of return occurrence, so the Law of Great Numbers may be perturbed especially in the periods when the entropy noticeably increases or decreases. For that reason parameter estimations of various statistical models are difficult and they are valid within certain period only. By this means we become familiar by presentating entropy and describing information in the context of predictability. Changes in time of WIG index distribution of the average predictability rate in the context of predictability scale is presented in

Figure 12: Predictability and compression of WIG index returns in time

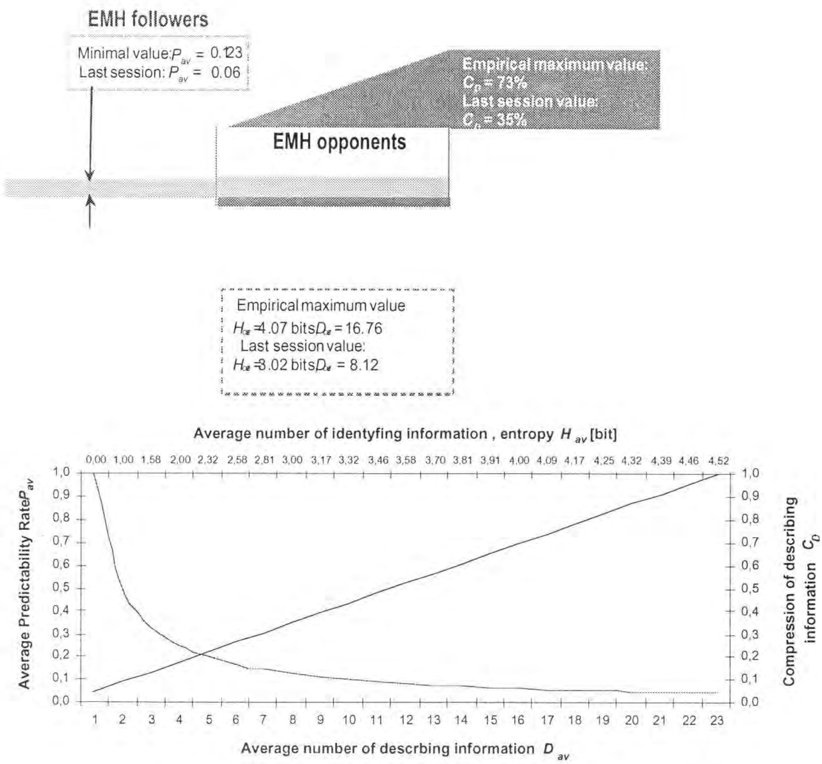


Figure 11.

Figure 11 confirms also that WIG predictability and uncertainty level are still far from the border of market non-efficiency ($P_{av} > 50\%$), but in time scale they are approaching to it. This verifies positively the research hypotheses: market is efficient in the meaning of predictability P_{av}.

Entropy, describing information and random walk of returns can be judged now. Entropy is a function of probability of a random variable representing the uncertainty in the system. It represents the amount of information (questions and premises) indispensable for full identification of return in the distribution. Describing information is a function of the average number of important intervals in the distribution from which this random variable is generated. In this context WIG index return distribution uncertainty level has never reached its maximal possible level of uncertainty $H = 4.52$ bits and $D_{max} = 23$, but maximally reached 4.07 bits and 16.76 describing information whereby it contributed 73 percent of maximal uncertainty.

Thus, we may state that random walk model (in its most rigorous form) should be consistent with maximal uncertainty level meaning that the maximal level of disorder and the lack of relationships means maximal uncertainty. The same features but expressed in different terms characterize also the process of random walk. Thus, the measures H_{av} and D_{av} support the assumption regarding random walk of returns but seems to be radical when compared to empirical behavior in the capital market. The character of price changes must be at a minimum level correlated and dependent. In turn, variability of distribution entropy and describing information together with sample size increase support the idea that return distribution is neither identical and nor stable.

The question about predictability of rate of returns should read like this: "to what extent the rate of returns are predictable or not predictable for the respective uncertainty (entropy and describing information) level of the distribution?" The question about the relationships level between the returns should be asked in the context of predictability, which imposes the level of accuracy of the relationships to be measured during the research. Therefore, it will be the following: "does relationship level between the returns characterize the high or low level of uncertainty assuming that prediction precision is for example 1 percent?" As a result, another question emerges additionally: "does the relationship level between the returns characterize the high or low level of the random walk?" As a random walk model does not suggest any graduation of random walk process like: low, medium, high but rather the statement "yes" or "no", it cannot cover the whole range of changes in uncertainty. As it was somewhat convincingly proved above, even low level of uncertainty ($C_D = 35.3$ percent) can still generate high level of unpredictability ($NP_{av} = 87.7$ percent), therefore we need other than random walk model, another model or formula that will cover changes in uncertainty and predictability from full predictability to full unpredictability.

Moreover, at such a level of conceptualization, we are not able to determine the causes of uncertainty, deliberate, rational or accidental, not rational, random decisions of investors, and other factors of qualitative nature. Therefore it is more appropriate not to apply the term "random" for the character of return changes, as it implies the one of possible cause of uncertainty, but not the only one, and as it was proved than randomness is not required for high unpredictability. More accurate is the case then for using the term uncertainty or unpredictability for description of this phenomenon. Many possibilities of rational decisions of investors occurring simultaneously with many possibilities of non-rational decisions of investors are at least two complementary causes of unpredictability of stock price changes, apart from other institutional, economic factors, etc.

The above evidence has significant implications for understanding the present state of various outstanding issues of informational efficiency. One of them is mentioned earlier as relative efficiency *RE* as the measure of uncertainty of price change character. As it is based on identifying information and is similar to compression of identifying information (relative entropy C_H), it cannot then indicate on predictability P_{av} using describing information D_{av}. This suggests why differentiation between identifying and describing information is significant.

We may conclude also based on the predictability scale (Figure 12), that there can be no less than medium level of uncertainty of the return distribution (relatively high level of correlations, relationships between returns) despite the fact that the average predictability level will still remain low. The increase of relationship level between returns does not mean linear and proportional increase of predictability level. With relationship level between returns reaching 50 percent of its maximal possible level, the average predictability rate is only 10 percent. To reach the average predictability rate of 50 percent, it will be necessary to increase the level of relationship between returns to 90 percent, and simultaneously decrease the maximal possible uncertainty to 10 percent.

In that regard, it seems that the definition of *random walk* is the most appropriate to express the relationship of high level of average uncertainty (H_{av} i D_{av}) and very low level of average predictability rate of returns in this sample of WIG prices based on assumption that such high level of uncertainty would become stable in time. In the case of WIG index return distributions, it may be the interval between 0 percent-5 percent of full predictability (probability 0-0.05) for 95 percent level of maximal possible uncertainty became greater than 4.32 bits up to 4.52 bits and the number of describing information ranged from 20-23. In view of the above results, it is rather hardly probable that WIG reaches ever such uncertainty and unpredictability level.

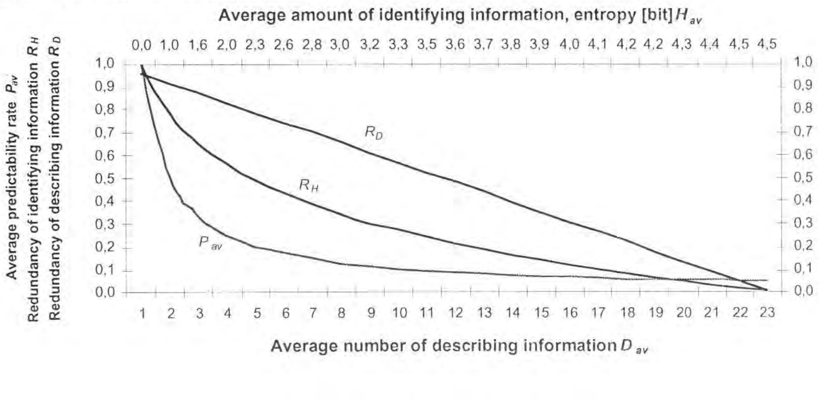
For the predictability rate of WIG index to achieve the value P_{av} = 10% the possibility set $N=D$ has to be increased further from 5 to 10. Thus the choice from the 43% (10/23 = C_D) of the total possibility set means decrease the predictability rate only from P_{av} = 20% to P_{av} = 10% and unpredictability rate equal to NP_{av} =90%. Further increase of possibility set grater than 43% (from 10 to 23) causes only little 10% decrease of total predictability P_{av} being close to 0%, and *NP* equal to 100%. Similar considerations may be carried out concerning the unpredictability rate.

The average predictability of returns P_{av} can be classified as follows as in Figure 12:

a) Very low predictability, P_{av} < 5%, C_D > 87.0% of maximal uncertainty, 4.52 bits >H_{av}> 4.32 bits, 23 >D_{av}> 20.
b) Low predictability, P_{av} = 5%-10%, C_D > 43.5% of maximal uncertainty, 4.32 bits > H_{av} > 3.32 bits, 20 > D_{av} >10.
c) Low-Medium predictability, C_D > 21.7% of maximal uncertainty, P_{av} = 10%-20%, 3.32 bits > H_{av} > 2.32 bits,10 > D_{av} > 5.
d) Medium predictability, C_D > 8.7% of maximal uncertainty, P_{av} = 20-50%, 2.32 bits > H_{av} >1 bit, 5> D_{av} > 2.
e) High predictability, P_{av} = 50%-95%, C_D > 4.8% of maximal uncertainty, 1 bit > H_{av} > 0.1 bit, 2 > D_{av} > 1.1.
f) Very High predictability, P_{av} > 95%, C_D < 4.8% of maximal uncertainty, H_{av} < 0.1 bit, D_{av} < 1.1.

In that context, initial classification of predictability levels for the respective levels of distribution uncertainty could be applied precisely to each publicly listed security and published in daily press or other media. This is a practical approach possible. Practical implementation of such parameters would enable investors to have a significant increase of awareness concerning capital market uncertainty related to each change of price or market indices. It is also possible to examine the entropy or describing information for various kinds of returns: weekly, monthly, quarterly, etc. and for more or less detailed intervals of the return distribution. It is also possible to measure entropy of trends: horizontal, growing, and decreasing. Finally apart from examination of uncertainty parameters by accumulation of samples in time (*n*+1) they can be measured for the defined constant period *m*

Figure 12: Proposal of predictability grades for the given uncertainty level



moved in time scale ($n+m$-1).

5. Conclusions

Based on the results of this research, we are of the opinion that we have an answer to the question: "are stock market returns predictable or does it follow random walk?" While asking about the predictability rate, we ask in fact about the relation of one event to the average number of describing information of the above mentioned WIG distribution: $P_{av} = 1/D_{av} = (D_{av} - 1)/D_{av} = 1 - NP_{av} = 12.3\%$. While asking about the level of random walk, we in fact mean compression of describing information of WIG returns – the relation of average uncertainty of the return distribution to its maximal possible value: $C_D = D_{av}/D = 35.3\%$ or redundancy 64.7% = 1 - C_D. As a result, we inquired not only one parameter, but in fact, about two independent parameters of uncertainty. In this context, the controversy between both followers and opponents of informational market efficiency may be resolved (see again Figure 12), because the first group means in fact the weak form of efficiency (i.e. Fama) being the near to low rate of predictability P = 12.3 percent, while the second one means the compression of returns C_D = 35.3 percent (the relation of number of describing information to its maximal possible level (redundancy 64.7 percent)). Compression reflects the fact that the correlation level is not proportional to the predictability level; therefore it must be normal for such situation that returns are not predictable and still can have correlations. This confirms the intuitive opinion that market cannot be "so much random" as the random walking process suggests, and it really is not (C_D = 35.3 percent of maximal uncertainty).

Additionally WIG contains 3.04 bits of identifying information, which means that, on average, we need 3 or 4 identifying questions (pieces of information) or premises to identify the individual return in the past distribution. WIG returns distribution contains also 8.12 pieces of describing information meaning that, on average, the rate of return is generated from 8 intervals of distribution and not from 23 of all possible ones. Taking into account only current perception of efficiency, we have to state that market is weak form efficient or is not, so the answer may be both: "yes" and "no", dependent on the terms we will treat efficiency - compression or predictability - as they are integral, but different measures (different functions) and pictures of uncertainty. As a result they cannot be hidden under the same term of "efficiency", otherwise internal self-contradiction occurs.

We observed that uncertainty and unpredictability have several measures, which mean something different. In fact due to the measures derived in this paper and based on information theory of Shannon-Mazur (entropy, identifying information compression, redundancy), we gained a broadened perspective on uncertainty phenomenon apart from existing ones using autocorrelation, other tests, etc.

1. It has been proved that it is very important to define - during discussion and research on market weak form of efficiency - what is meant by efficiency, as

the final results are very relative and are dependent on: a) prediction precision (0,1%, 1%, 10% etc.); b) object of prediction: daily return value divided into intervals, or daily consecutive rate of return change types (++, —, -+, +-), etc.; c) measure used: average predictability *P* or unpredictability *NP*, compression *C* or redundancy *R*, correlation , etc.; d) type of information: identifying of describing information *H* and *D* used for uncertainty measures of information itself and for uncertainty level measures based on them, i.e. compression and redundancy.

2. This research shows the evidence – in opposition to existing knowledge – that efficiency is not only increasing but may also decrease in time. For WIG index efficiency decreased doubly.
3. The factors mentioned above directly are responsible for misunderstanding and controversies relating to efficiency in weak form and partially with informational efficiency in general. Both opponents and followers of efficiency hypothesis have to define what they mean by efficiency otherwise such discussion will be unproductive. The measures derived from combined Shannon-Mazur information theory should enable better communication between both sides participating in the discussion about market efficiency.

All this was possible to prove owing to precise mathematical definitions of uncertainty, specifically, absolute and relative measures of identifying and describing information in order to resolve this forty-year old controversy on price changes and stock returns character within the range of its predictability level. For this we need the science of cybernetics.

Acknowledgement: This paper is one of eight best papers selected after blind review process at the 3rd International Banking and Finance Conference, Szczecin, Poland, 2005.

Author statement: Slawomir Piotr Pijanowski s current research is on cybernetic interpretation of all forms of market efficiency and dual-hypothesis problem. The author wishes to thank participants of the Conference for helpful comments. The editors of the Journal made significant changes to the original version of the paper, for which I record my sincere appreciation. I also owe my thanks to Merouane Lakehal-Ayat for motivation to continue the research and Mohamed Ariff for valuable feedback during submission process. The errors are our responsibility.

References

Buczek S.B., (2005). Efektywno__ informacyjna rynków akcji. Teoria a rzeczywisto__. [*Informational Efficiency of Stock Markets. Theory and Reality*]. Szko_a G_ówna Handlowa w Warszawie – Oficyna Wydawnicza, Wydanie II poprawione, Warszawa 2005.

Campbell J.Y., Lo A.W., MacKinlay A.C., (1997). *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Chen J., (2004). Generalized Entropy Theory of Information and Market Patterns, *Corporate Finance Review* 9 (No. 3): 21-32.

Cozzolino J.M., Zahner M.J., (1973). The Maximum-Entropy Distribution of the Future Market Price of a Stock. *Operations Research pp.* 1200-1211.

Fama E.F., (1965). Tomorrow on the NewYork Stock Exchange. *Journal of Business*. 38 (No. 3): 285-299.

Fama E.F., (1970). Efficient Capital Markets, A Review of Theory and Empirical Work. *The Journal of Finance 45(*No. 2): 383-417.

Golan A., Judge G., Miller D., (1996). *Maximum Entropy Econometrics: Robust Estimation With Limited Data*, New York: Wiley & Sons.

Jones R.D., Redsun S.G., Frye R.E, Myers K.D.: *The Maxwell Demon and Market Efficiency*, Complexica Report 031115, http://internet.cybermesa.com/~rfrye/complexica/tech.htm.

Grossman S. J., Stiglitz J.E., (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70: 393-408.

Lo A.W., MacKinlay A.C., (1999). *A Non-Random Walk Down Wall Street*, Princeton, Princeton University Press.

Loistl O., Veverka A., (2004). *Entropy and Capital Market Efficiency, Presentation on 19th Workshop of the Austrian Working Group on Banking and Finance*, Vienna University of Economics and Business Administration Vienna, Austria November 26-27, 2004.

Malkiel B.G., (2003). *A Random Walk Down Wall Street*, W.W. Norton & Company.

Mazur M., (1969). *Cybernetyka a zarz_dzanie* [*Cybernetics and Management*]. Ministerstwo Spraw Wewn_trznych, Biblioteka zagadnie_ literatury wspó_czesnej, Warszawa.

Mazur M., (1970). *Jako_ciowa teoria informacji* [*Qualitative Theory of Information*] Wydawnictwa Naukowo-Techniczne, Warszawa.

Nawrocki D., (1984). Entropy, Bifurcation and Dynamic Market Disequilibrium, *The Financial Review* 26. 6, May 1984.

Shannon C.E., (1948). A mathematical theory of communication, *Bell Systems Technical Journal* 27 (No. 3-4): 379-423 & 623-656.

Nicolis G., .Prigogine. I., (1997). *Self-Organization in Nonequilibrium Systems.* John Wiley and Sons.

Szyszka A., (2003). *Efektywno__ gie_dy papierów warto_ciowych w Warszawie na tle rynków dojrza_ych* [*Efficiency of Warsaw Stock Exchange in the light of matured capital markets*].Wyd. Akademii Ekonomicznej w Poznaniu, Poznan_.

Tambakis D.N., (2000). On the Informational Content of Asset Prices, Working Paper.

Theil H., Lenders C.T., (1965). Tomorrow on the Amsterdam Stock Exchange, *Journal of Business* 38 (No. 3): 277-284.