Jacek Batóg
Department of Statistics and Econometrics
University of Szczecin
64, Mickiewicza Street, 71-101 Szczecin, Poland
batog@univ.szczecin.pl
http://www.us.szc.pl/batog

APPLICATION OF CERTAIN STATISTICAL MEASURES OF SIMILARITY IN THE ANALYSIS OF INFLUENCE OF MARKET STRUCTURE ON DIVIDEND POLICY

Abstract

Decisions connected with dividend policy belong to the most important financial decisions made by enterprises. A payment of a dividend raises the share of debt in total liabilities, which leads to a decrease in the cost of capital and increase in the market value of equity. It has also a significant impact on the firm's likeness and decrease in tax rates. In the paper the author tried to identify determinants of the decision whether to pay dividend. The methodology is based on some heterogeneity, correlation and distance measures, which are used in classification problems. The recurring procedure applied uses a specific function as a criterion of heterogeneity of received subsets. Using statistical analysis of correlation between dependent and potential independent variables we are able to chose two main determinants of dividend payment: market value of the company and type of it's activity. All the employed measures pointed out that the most important one is the latter.

Key words: dividend policy, market structure, statistical measure, classification.

Introduction

In the paper an attempt was made to answer the question if the selected variables, determining elements of the structure of the market, have the essential influence on the decisions regarding the payment of the dividend by public limited companies listed on the Polish capital market.

The decisions which form the dividend policy belong along with investment decisions and decisions about the structure of the capital to the most important financial decisions made by companies. One of essential results of the payment of the dividend to stockholders is the growth of the share of the debt in liabilities which decreases the cost of capital and additionally, have an effect on the increase in the market value of equity. Both of these phenomena could be easily explained by analysis of the method of valuation of the cost of capital employing WACC formula[1] (Hozer and others, 1997, pp.90-91), (Machała, 2001, p.305).

The increase of dividend payout decreases the cost of capital, and simultaneously lowers the required rate of return, which leads eventually to an increase in the market value of companies through the influence on their stock price. This dependence is represented by the

---

[1] Both these positive phenomena are related to the tax-shield as a result of the interest on debt.

M. Gordon's Capital Assets Pricing Model[2]. There exist comparatively many theories on the method of estimation of the dividend. Most of them condition the level of dividend payout from the profits made by companies. An example of the model, where the level of dividend payout is to certain degree the weighted average of present and past profits, is the model introduced by J.Lintner (Brealey, Myers, 1991, p.375).

The decision whether to pay the dividend is the choice between the reinvestment of generated profit and payment of its part or all of it to stockholders. When the company does not use external financing, the value of the dividend is equal to the difference between generated and retained profit. It leads to the assumption that the present value of all expected dividends is equal to the present value of expected profits minus the present value of expected profits generated by reinvestments. In the case when the rate of return on equity attains the same value as the required rate of return the cost of capital is equal to the ratio of the dividend to the required rate of return.

The dividend payout plays also an essential part in the valuation of the option. When dividend is paid before option exercise date it is necessary to correct the classical formula of Black-Scholes (Bodie, Kane, Marcus, 1989, p.596).

Presently, the financial form of dividend payout dominates on the Polish capital market. The economic practice knows however many other manners of its payment. The choice of the form of the dividend is most often dependent on two factors: the fiscal policy of the state and expectations of stockholders. There also exist theories according to which the decision about the payment of dividend has the neutral character. The theory of Miller and Modigliani could be given as an example. According to this theory, decisions whether the dividend should be paid are not related to an existence of the perfect capital market and the lack of taxes and transaction costs.

So far the existing practice of public companies listed on the Polish capital market has shown that the number of companies that decide to award their stockholders in the form of dividend constantly grows. In the early years of the Warsaw Stock Exchange the payment of dividends could be observed only rarely. At present the fraction of companies which pay dividend increases with every year (in 2004 as much as 52 out of 125 companies which obtained the positive net financial result paid the dividend).

It seems that payment of the dividend by public companies becomes more and more important to investors. The high value of the dividend payout enlarges capital gains related to the rise in prices of these shares in the period of the prosperity on the market and diminishes the value of losses generated during the recession.

The methodology of the research is based on the application of selected measures of heterogeneity sets, correlations and distances used in classification problems. These measures are used in estimation homogeneity of dividing the set of examined objects into subsets and in selecting that variable from among all diagnostic variables which becomes exogenous in the following model (Gatnar, 2001, p.14):

$$y = f(x, \alpha) + \varepsilon, \tag{1}$$

where:
y – the endogenous variable,
x – exogenous variables (predicates of researched objects),
α – parameters of the model,
ε – random error.

The procedure of selecting the variable uses the function evaluating the quality of the

[2] In case of companies, in which dividend is on constant level, the mentioned formula reduces to: $P = D/r$. For comparison see the model of changing growth of dividend (Tarczyński, 1997, p.194).

received partition of examined objects. On the given stage of recursive partitioning, from all the variables that variable[3] is chosen, which assures the highest homogeneity of subsets received as result of the partitioning. E. Gatnar shows that a better solution is to minimize the function measuring heterogeneousness (formula (5)) than maximize the function estimating the homogeneity of subsets:

$$H(S_k) = \Phi[p(1 \mid k),...,p(J \mid k)], \tag{2}$$

where:

$p(j \mid k) = \frac{N_j(k)}{N(k)}$ - the probability that an object from subset k belongs to class j (j=1, 2, ..., J),

$N_j(k)$ –number of objects from class j belonging to subset k,

$N(k)$ – the number of objects in subset k.

The change of the level of the heterogeneousness obtained as a result of partitioning of set S into subsets S_k made by using variable x_m could be in certain cases measured by the following:

$$\Delta H(S, x_m) = H(S) - \sum_{k=1}^{K} H(S_k)p(k), \tag{3}$$

where:

p(k) - the probability that as a result of partitioning, one of the objects belongs to the subset S_k.

The maximum value of the change in formula (6) informs about the lowest degree of the heterogeneousness of subsets S_k. Variable x_m is then considered the best criterion of the partitioning on the given stage of discrimination. Simultaneously, it seems that the opinion that the above mentioned procedure of selecting variables could be the base for determination of the most important factors indicating whether the payment of the dividend is justified.

Table 1 presents selected measures used to estimate the heterogeneousness of subsets S_k and the measures of the correlation and distance used in the analysis.

Table 1. Selected measures of the quality of the partition

No.	Name of the measure	Method of the calculation[1] $\Phi[p(1 \mid k),...,p(J \mid k)]$
1	Error of the classification (M1)	$1 - \max_j \{p(j \mid k)\}$
2	Entropy function (M2)	$-\sum_{j=1}^{J} p(j \mid k) \log_2 p(j \mid k)$
3	Gini's coefficient of the variation (M3)	$1 - \sum_{j=1}^{J} p^2(j \mid k)$
4	Statistic χ^2 (M4)	$\chi^2 = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{[N_j(k) - L_j(k)]^2}{L_j(k)}$, where: $L_j(k) = \frac{N_j \cdot N(k)}{N}$

[3] Or some combination of variables.

5	Goldman-Kruskal's coefficient τ (M5)[2]	$\tau = \dfrac{\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{p_{jk}^2}{\pi(j)} + \sum_{k=1}^{K}\sum_{j=1}^{J}\frac{p_{jk}^2}{p(k)} - \sum_{j=1}^{J}\pi^2(j)}{2 - \sum_{j=1}^{J}\pi^2(j) - \sum_{k=1}^{K}p^2(k)}$, where: $p_{jk} = \frac{N_j(k)}{N}$, $\pi_j = \frac{N_j}{N}$
6	Minger's measure G (M6)	$G = 2N\log_e(2)\Delta E(S, x_m)$, where: $\Delta E(S, x_m)$ change of the function of the entropy
7	Kolmogorov-Smirnov's distance D (M7)[3]	$D = \max_m \left\| F_{y_1}(x_m) - F_{y_2}(x_m) \right\|$, where: $F_{y_1}(x_m)$ and $F_{y_2}(x_m)$ cumulative functions of distribution of variable x_m in class 1 and 2
8	Ripley's deviation M8)[4]	$O(k) = -2\sum_{j=1}^{J} N_j(k)\log_2 p(j \mid k)$
9	Measure of the number of objects I (M9)[5]	$SM = \sum_{k=1}^{K} M(k)$, where: $M(k) = \min\{N_1(k), N_2(k), ..., N_J(k)\}$
10	Measure of the number of objects II (M10)[5]	$MM = \max\{M(1), M(2), ..., M(K)\}$

[1] measure M(1), M(2), M(3) and M(8) were calculated as changes (6),

[2] probabilities a priori π(j) were calculated for set S,

[3] Utgoff, Berkman and Clouse made enlargements of this measure on case of J > 2 by grouping classes,

[4] it occurs relation O(k) = 2N(M2),

[5] in the case of measures M9 and M10, their lower value inform about higher homogeneity of subsets S_k.

Source: (Gatnar, 2001, p.33 and further).

The rule of partitioning used by Breiman in the CART method and the Fayyad and Irani measure leaning on orthogonality of vectors of probabilities were not presented in Table 1. These measures are only applied when the partition is a binary character. Other kind of the measure that could be used in the procedure of partitioning is a measure based on discriminating coordinates. This measure is used in case when we seek predictors which have different distributions in each class instead of examining distributions of the dependent variable in subsets S_k (Gatnar, 2001, p.46).

One of the key problems connected with the quality of received partition is the number of variables characterizing classified objects. The number of variables used in the classification cannot be too large, and one of the criteria of their selection should be the ability to discriminate the set of examined objects (Walesiak, 2004, p.321).

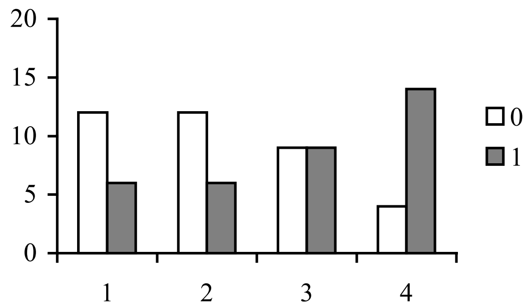
Empirical results

The original set of factors which can determine dividend decision, sorted according to theoretical assumptions contained the following variables characterizing the examined companies: the trading value, the type of industry, the coefficient the price/book value ratio, the net profitability of sales, the value of cash flows, the value of investment outlays and the level of debt. The statistical analysis of the strength of the relationship between the above variables and the dependent variable (y = 1 if the dividend for year 2004 is paid, y = 0 if no dividend for 2004 year is paid) led to the choice of two from the above set of potential variables:

- X_1 - market value of the company,
- X_2 - the type of industry.

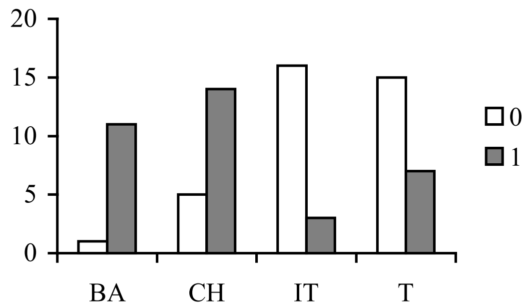
The selection of the variable decides about the number of received subsets S_k. This problem is close to the fixing of the optimum number of classes. This topic was explored by some authors (Milligan, Cooper, 1985), (Siedlecka, 1999), (Sokołowski, 1992), (Walesiak, 1988). The number of subsets for variable x_1 was settled using the discretization procedure of its value on the ground of quartile partition[4]. In the case of variable x_2 the number of subsets was the same[5]. Thereby we reduce the incomparability of used measures and we avoid the necessity of their normalization. In the analysis, 72 public companies listed on the Warsaw Stock Exchange 17.02.2006 were taken into consideration. Figures 1 and 2 present a distribution of dividend decision according to the size of company and kind of business. We can observe that positive decision about dividend payment are made mainly by big companies and the firms operating in the banking sector and chemical industry.

Fig. 1. Distribution of dividend decisions for variable X_1



Source: author's own calculation.

Fig. 2. Distribution of dividend decisions for variable X_2



Source: author's own calculation.

Basing on the measures presented in the Table 1, the variable will be selected that should serve as the criterion of the best partitioning of companies according to the value of dependent variable. Table 2 presents the numbers of each subset and the probabilities which were used for calculation of M1-M10 measures.

[4] The values of median, first and third quartile were the base of partition of set S on subsets S_k. Companies from each subset received values 1, 2, 3 and 4 accordingly to their market value (1 for the lowest market value).
[5] Chosen sectors: banking sector (BA - 12 companies), chemical industry (CH – 19 companies), IT sector (IT - 19 companies) and trade (T – 22 companies).

Table 2. Table of contingency for variables X_1 and X_2

y	X_1				
	1	2	3	4	N_j
0	12	12	9	4	37
1	6	6	9	14	35
N(k)	18	18	18	18	72
p(0 | k)	0.67	0.67	0.50	0.22	0.51
p(1 | k)	0.33	0.33	0.50	0.78	0.49
p(k)	0.25	0.25	0.25	0.25	
y	X_2				
	BA	CH	IT	T	N_j
0	1	5	16	15	37
1	11	14	3	7	35
N(k)	12	19	19	22	72
p(0 | k)	0.08	0.26	0.84	0.68	0.51
p(1 | k)	0.92	0.74	0.16	0.32	0.49
p(k)	0.17	0.26	0.26	0.31	

Source: author's own calculation.

Values of calculated measures are presented in the table 3.

Table 3. Values of measures M1-M10

Variable	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
X_1	0.153	0.099	0.066	9.507	0.079	9.906	0.306	14.291	25	9
X_2	0.328	0.269	0.169	24.364	0.200	26.877	0.552	38.775	16	7

Source: author's own calculation.

On the ground of the above results, one can ascertain that in the case of all measures, their values determine the choice of variable x_2 as variable which should be introduced to the model (4) in the first place. So it seems that the assumption about the greater influence of the type of industry of public company than its trading value on the decision concerning the payment of the dividend is justified.

Conclusions

It seems that the information on the existence of the significant relationship between the kind of business of the company and its propensity to the payment of the dividend can be the most important factor in making the decision whether to buy and sell the shares by participants of the capital market. The statistical analysis also points out that market value of the company has a significant influence on the dividend policy of public companies.

The payment of the dividend by public companies constitutes a more and more relevant factor to the formulation of investment strategies on the capital market. The payment of the dividend, except apparent capital gains, can be also examined in the aspect of the formation of the benchmark for the company, the emission of signals for potential investors about the profitable economic-financial situation of the company. It can also be perceived as the method of the reduction of the level of tax rates (Lumby, 1991, p.480).

REFERENCES:

1. Bodie Z., Kane A., Marcus A.J., (1989), Investments, Irwin, Homewood, Boston.
2. Brealey R.A., Myers S.C., (1991), Principles of Corporate Finance, McGraw-Hill, International Editions.
3. Gatnar E., (2001), Nieparametryczna metoda dyskryminacji i regresji, Wydawnictwo Naukowe PWN, Warszawa.
4. Hozer J., Tarczyński W., Gazińska M., Wawrzyniak K., Batóg J., (1997), Metody ilościowe w analizie finansowej przedsiębiorstwa, GUS, Warszawa.
5. Lumby S., (1991), Investment Appraisal and Financing Decisions, Chapman & Hall, London, New York, Tokyo, Melbourne, Madras.
6. Machała R., (2001), Praktyczne zarządzanie finansami firmy, Wydawnictwo Naukowe PWN, Warszawa.
7. Milligan G.W., Cooper M.C., (1985), An examination of procedures for determining the number of clusters analysis in a data set, Psychometrika, no. 2.
8. Sokołowski A., (1992), Empiryczne testy istotności w taksonomii, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, Monografie no. 108, Kraków.
9. Siedlecka U., (1999), in: Statystyczne metody analizy danych, ed. W.Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.
10. Tarczyński W., (1997), Rynki kapitałowe. Metody ilościowe, vol.I, Agencja Wydawnicza Placet, Warszawa.
11. Walesiak M., (1988), Sposoby wyznaczania optymalnej liczby klas w zagadnieniu klasyfikacji hierarchicznej, Prace Naukowe Akademii Ekonomicznej we Wrocławiu no. 449, Wrocław.
12. Walesiak M., (2004), in: The method of statistical multidimensional analysis in marketing investigations, ed. E. Gatnar, M. Walesiak, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.