School of Business, Economics and Law
GÖTEBORG UNIVERSITY

Probabilistic Prediction of Bankruptcy with Financial Ratios
-An empirical study on Swedish market

Tugba Keskinkilic and Gunes Sari

Graduate Business School

Industrial and Financial Economics
Master Thesis No. 2006:9
Supervisor: Lennart Flood

Acknowledgements

Firstly, we would like to thank our supervisor, Lennart Flood, for his precious guidance, useful feedback and willingness to provide us with data which gave substance to our thesis.

Secondly, we are also grateful to Daniela Andrén for her contribution and valuable advice concerning a crucial part of our work.

Thirdly, we are so pleased that our patience and cooperation not only enabled us to do a great job but also gave us a great opportunity to have fun throughout the work.

Last but not least, special thanks go to our families for their emotional and financial supports and invaluable motivations.

ABSTRACT

Credit risk measurement has become more important during the last 20 years in response to a worldwide increase in the number of bankruptcies. This paper examines some of bankruptcy prediction models using financial accounting ratios. Logit and LPMs are employed in order to develop these prediction models. The purpose of this study is to assess the effects of the determined financial ratios and the selected industries on bankruptcy events that occurred between 2002 and 2006 in the Swedish market. These effects are calculated by measuring elasticity and marginal effect. In addition to prediction models calculating the effects of industries by means of dummy variables, industry normalized financial ratios are also used in order to control industry differences. The empirical results indicate that the company is more likely to go bankrupt if it is unprofitable, small, highly leveraged, and has liquidity problems and less financial flexibility to invest in itself. Furthermore, a company is more likely to enter bankruptcy if it operates in the wholesale and retail trade industry among the selected industries in our sample.

Key words: Credit Risk; Financial Accounting Ratios; Industry Relative Ratios; Linear Probability Model; Logit Model

TABLE OF CONTENTS

TABLES & FIGURES

1. INTRODUCTION

The question of what kind of factors can be helpful in order to understand the behaviour of bankruptcy has been addressed in a field of credit risk management and the academic world. According to literature generally accepted statistical models used for prediction are as follows:

(1) The linear probability model (LPM),
(2) The logit model,
(3) The probit model,
(4) The multiple discriminant analysis

Apart from all multivariate statistical models listed above, Beaver (1966) developed univariate analysis. This study is regarded as one of the classic studies in this field. Univariate analysis compares the key accounting ratios with industry or group norms at a point in time.

Altman (1968) improved on Beaver's univariate study by introducing the multivariate approach, which allows for the simultaneous consideration of several variables in the prediction of failure. Altman was the first to apply the multivariate technique known as linear discriminant analysis to develop a business failure prediction model for the United States manufacturing industry. This model, so called Z-score model, is built upon the values of both ratio-level and categorical univariate measures. These values are combined and weighted to obtain a measure which discriminates between failed and non-failed firms. According to Altman (1968), this model is applicable because firms that fail have ratios and financial trends that are discriminated easily from those firms that are financially sound.

Apart from Altman (1968), there have also been several studies using discriminant analysis applied to prediction of business failure. Some of them are as follows; Altman (1971) examining railroad bankruptcy propensity; Deakin (1972) replicating study of Beaver(1966), Edmister (1972) testing the usefulness of financial ratio in order to predict small business failure; Altman, Margaine, Schlosser and Vernimmen (1974) developing a model in order to determine the credit worthiness of commercial loan applicants in a cotton and wool textile sector in France, Blum (1974) examining general denominators underlying cash flow framework; Altman, Haldeman, and Narayan (1977) that is known Zeta Analysis which is the revision of the classical Z model, Norton and Smith (1979) who compared the prediction of bankruptcy using ratios computed from General Price Level (GPL) financial statements to the prediction of bankruptcy using ratios computed from traditional historical cost financial statements, Taffler (1982) who used linear discriminant analysis for the prediction of bankruptcies in UK with financial ratios; Altman and Eom (1995) attempting to construct and test a failure prediction model for Korean companies.

Although it is the mostly used technique in literature (Altman and Saunders, 1998), discriminant analysis contains some problems in terms of the assumptions it is based on. The first assumption is that financial ratios as independent variables are normally distributed and the second assumption is that the financial ratios of bankrupt and non-bankrupt firms have the same variance and covariance matrices. Even if Altman (1977) creates quadric discriminant analysis in order to relax the assumption of equal variance-covariance matrices, estimation process are very complicated (Eisenbeis, 1977). In fact, some studies comparing the logit model and discriminant analysis such as Martin (1977), Press and Wilson (1978) and Wiginton (1980) generally state that the logit model is preferable against discriminant analysis.

Since assumptions about normality and identical covariance matrices are not satisfied, Ohlson (1980) used the logit model to predict bankruptcy by using accounting ratios as independent variables since no assumptions should be made about the probabilities of bankruptcy and/or the distribution of independent variables. Martin (1977), West (1985), Platt and Platt (1991), Lawrence and Smith (1995) are other popular studies using the logit model in order to assess default probabilities. Nevertheless, Stone and Rasp (1991) and Maddala (1991) compare logit and OLS and have the same result that the logit model is preferred to OLS models for accounting studies, even in small samples.

Despite these results, Suzuki and Wright (1985) used multiple regression analysis to determine the business risk in Japanese companies, and the differences from U.S firms and Meyer and Pifer (1970) using LPM carried out the analysis of predicting bankruptcy of banks which happened between 1948 and 1965. There are also some studies using the probit model in order to assess default rate in literature such as Zmijewski (1984), Casey, McGee, and Stickney (1986), Noreen (1988).

However, as there is no widely accepted economic theory, every study has based their model specification on an empirical framework. This results in different accounting ratios used in different models. Generally speaking, these multivariate models are conducted on three types of data set. One of them is the match making procedure that is structured in such a way that an equal number of bankrupt and non-bankrupt firms are chosen randomly with respect to company size or industry. Others are large and small samples avoiding matching procedure.

This study utilizes linear probability and the logit model on the Swedish market. The authors try to keep the data set as large as possible and avoid match making procedure in order to examine the marginal effects of financial ratios together with size, and industry effects to probability of bankruptcy. Namely, small samples can cause over fitting problems and match making procedure can make it difficult to identify size and industry effect.

Other than examining industry effect on probability of bankruptcy, this paper also uses models consisting of industry normalized financial ratios in order to control industry differences and applies model specification test in order to compare this type of models with models including unadjusted ratios and dummy variables.

The structure of the rest of the paper is as follows. Section 2 outlines methodology and data used in the present study. Section 3 explains variable selection. Section 4 discusses the empirical findings. Section 5 gives evaluation of predictive accuracy of models and section 6 offers conclusions.

2. Methodology and Data

This study employs linear probability and the logit model in order to analyze bankruptcy. Firstly, the LPM can be written as:

$Y_j = \beta_0 + \beta_i X_{ij} + \varepsilon_j$ for $Y_j = 0$ and $Y_j = 1$; where β_i represents the coefficient of ith variable and ε_j represents stochastic terms of all observations denoted by j. Since LPMs are linear models estimated by Ordinary Least Square (OLS), they have the same assumptions as other linear models. Under the assumptions of the error terms having a mean of zero, being independent of one another, and of the independent variables, and having the same variance, OLS estimator is the best linear unbiased estimator (BLUE) for β_i.

Marginal effect of one variable is calculated by taking derivatives of a dependent variable with respect to an independent variable, which gives us a slope of regression line. Since this is a LPM estimated by OLS, the marginal effect of any variable to probability of bankruptcy is captured from directly coefficient. Another interpretation can be made by elasticity. Elasticity gives the percentage change in the probability of bankruptcy in response to a one percentage change in the independent variable.
On the other hand, the basic function of logistic analysis is,

$$P_i = E(Y_i = 1/X_i) = \frac{1}{1 + e^{-(\beta_0 + \beta_i X_i)}}$$

For the logit model, the estimated parameters do not have a direct interpretation in comparison to LPM. Measures which are familiar to economists are marginal effects and elasticities. In the logit model the probabilities are not linear in independent variables, leading us to the fact that there is no unique slope. Every point on this line gives us a different slope; i.e. marginal change on probability of bankruptcy. Hence, the marginal change is not constant. To compute marginal changes, the first partial derivative with respect to a corresponding independent variable should be taken. This leads us the following formula (Gujarati, 2003):

$$\frac{\partial P}{\partial \mathrm{X}_i} = \beta_i * \hat{P}_i (1 - \hat{P}_i)$$

Apparently, marginal change does not only depend on coefficient but also on the predicted probability of that variable. The definition of marginal change can be made as follows. For a unit change in X_k from the baseline, the probability of bankruptcy event is expected to change by the magnitude of the marginal change when all other variables are held constant. The mean value can be used as a baseline.

However marginal change in probability of bankruptcy is not tenable in order to interpret for dummy variables in this model. Instead, discrete change is the appropriate one, and in this case, this kind of change is defined as follows. By the change from X_k to $X_k+\delta$, the probability of bankruptcy changes by a magnitude of discrete change[1]; all other variables are kept at their given values. Continuous variables are kept at their mean, dummy variables are kept at their modal values. The formula for discrete change is as follows:

$$\frac{\Delta P}{\Delta X_k} = P(Y=1|X, X_k, X_k+\delta) - P(Y=1|X, X_k) \text{ Where } P = \frac{1}{1+e^{-(X\beta)}}$$

On the other hand, elasticity gives the percentage change in the probability of an event in response to a one percentage change in the independent variable. Since the elasticity is acquiring a different value on each point on the line of regression, it is plausible to calculate it at the point of the means, i.e. a representative point on the regression line. For the i^{th} independent variable elasticity is obtained using partial derivatives as:

$$\frac{\partial P(Y_t=1|X_j)}{\partial \mathrm{X}_{ij}} * \frac{X_{ij})}{P(Y_t=1|X_j)}$$

The data used in the present study was obtained from UC AB, named as “Upplysningscentralen”. UC is known as the largest and leading Swedish Business and Credit Information Agency. Through its large database, UC offers not only business reports but also credit monitoring and quantified financial analysis with its computer based systems. In other words, UC is accepted as one of the world’s widely respected and high quality information providers. The high quality of data strengthens the findings and the credibility of the models proposed in this paper.

[1] That is 1 in this case.

The rough data of analysis of the paper contains 262,769 Swedish firms with the number of 486,285 observations. Unfortunately, a complete panel data could not be obtained for each firm in the data set which is understandable for a large sample. In the analysis the companies are categorized according to their SNI Swedish Standard Industrial Classification 2002 codes (SNI-codes) involving 15 main industries. The data set covers 162 financial accounting ratios or items of companies. Additionally, the initial records of financial statements are within the time interval 2000-2003 and the closure records are within the time interval of 2002-2003.

Firstly, the time period between the bankruptcy event and the closure date of statements, and the time period between the bankruptcy event and the open date of statements of observations which entered bankruptcy are calculated. In the data set there are 27 observations having the closure date of financial statements which is later than the bankruptcy event, and 20 of those firms entered bankruptcy within the time period of the financial statements recording. Because the informative indicators have already been reflected with the financial accountants' reports of a companies financial statements, those firms may not be realistic representatives of bankrupted firms in the estimation of probability of failure. These 27 observations whose financial statements were audited after the bankruptcy event and 40,405 companies that do not have a SNI code label are dropped from the rough data set. Additionally, 36 observations are obtained with negative total asset values and 15 of them also do not have SNI codes. Since, some additional variables are generated by using total asset items such as the "SIZE" ratio which is defined in logarithmic form. This condition barely contradicts with the common sense of accounting; these observations are also omitted from the data set.

The models in this analysis only involve seven industries which are agriculture, hunting and forestry; fishing; mining and quarrying; manufacturing (involving the sub-classes of manufacturing[2]); electricity, gas and water supply; construction; and, wholesale and

[2] Manufacture of textiles and textile products (DB), manufacture of leather and leather products, manufacture of pulp (DC), paper and paper products; publishing and printing (DE), manufacturing of coke, refined petroleum products and nuclear fuel (DF), manufacture of chemicals, chemical products and man-made fibres (DG), manufacture of rubber and plastic products (DH), manufacture of other non-metallic mineral products (DI), manufacture of basic metals and fabricated metal products (DJ), manufacture of machinery and equipment n.e.c. (DK), manufacture of electrical and optical equipment (DL), manufacture of transport equipment (DM), and manufacturing n.e.c. (DN).

retail trade industries. The industries including utility companies, transportation, public companies, and financial intermediations, and financial services, i.e. banks, insurance companies, pension funds etc. are excluded from the initial data set. As Ohlson (1980) stated, it is acceptable not to include these industries because they differ with their financial structures and bankruptcy environment.

Some additional accounting ratios are obtained that are different from the ready ones given with the data set for the purpose of using throughout the analysis. These procedures are demonstrated in details with the variable selection part. Hence, after observation deletion procedure with respect to the selected industries, and considering the exclusion of firms having no industry label, the final data set including new variables has been acquired with 177,620 observations.

Table 1.The Number of Observations for the Initial and Analysis Data Sets According to Their Relative Industries

SSIC Code	Industry Sector	Data Labelled with Selected Industries	Data of Analysis
A	Agriculture, hunting and forestry	12,346	10,729
B	Fishing	383	329
C	Mining and quarrying	785	625
D	Manufacturing	50,946	43,148
E	Electricity, gas and water supply	972	789
F	Construction	42,968	36,804
G	Wholesale and retail trade	103,252	85,196
H	Hotels and restaurants	15,422	-
I	Transport, storage and communication	26,598	-
J	Financial Intermediation	10,939	-
K	Real estate, renting and business activities	149,624	-
L	Public administration and defence; compulsory soc. security	40	-
M	Education	4,923	-
N	Health and social work	12,124	-
O	Other community, social and personal service activities	14,631	-
		445,953	**177,620**

* The data analysis tables are obtained and reported by using SAS 9.1.

** The difference between 486,285- the initial number of observations- and 445,953 is equal to the number of observations having no industry label in the initial data set. These 40,332 observations have been omitted in the analysis.

The time interval of the bankruptcy event for the data set of analysis is 2002/06/20 – 2006/06/01 and the number of bankruptcies is 6,877 whereas the time interval of bankruptcy of rough data is 2002/05/10 – 2006/06/01 and there are 15,301 bankruptcies respectively. The data used in analyses involves observations having the mean of financial statement recording period equals to 1.0034 year, and this period lies within

January 2002 - December 2003. It shows that the fiscal year for the companies is found to be approximately 1 year[3].

Table 2.The Number of Bankruptcies over Years for All Industries of the Initial Data

The Number of Bankruptcies Over Years							
SSIC Index:	2002	2003	2004	2005	2006	Total	Percentage
A	0	31	73	64	19	187	1.32%
B	0	0	2	3	3	8	0.06%
C	0	0	8	6	3	17	0.12%
D	7	329	626	622	198	1,782	12.58%
E	0	0	9	6	0	15	0.11%
F	6	210	575	605	197	1,593	11.24%
G	20	528	1,484	1,583	638	4,253	30.02%
H	6	110	308	376	147	947	6.68%
I	3	122	347	404	139	1,015	7.16%
J	2	17	60	65	25	169	1.19%
K	23	476	1,190	1,366	486	3,541	24.99%
L	0	0	0	1	0	1	0.01%
M	2	17	47	53	17	136	0.96%
N	1	24	56	49	37	167	1.18%
O	1	41	109	135	52	338	2.39%
	71	1,905	4,894	5,338	1,961	14,169	

The tables show the number of bankruptcies with respect to the industry scale and relating fiscal years. Eventually the overall bankruptcy ratio of the population is found 3.4311% whereas it is 3.8717 % for the data set of analyses.

Table 3.The Number of Bankruptcies According to Years

	Observations in the Data of Analysis			All Observations in the Initial Data Set		
Bankruptcy Date	Frequency	Percentage	Cumulative Frequency	Frequency	Percentage	Cumulative Frequency
2002	32	0.47%	32	75	0.49%	75
2003	1,021	14.85%	1,053	2,017	13.18%	2,092
2004	2,483	36.11%	3,536	5,280	34.51%	7,372
2005	2,444	35.54%	5,980	5,806	37.95%	13,178
2006	897	13.04%	**6,877**	2,123	13.87%	**15,301**

Finally, for further investigation and information of the data structure for the scope of bankrupted firms, the frequency table of lead time between the date of last relevant financial statement report and the bankruptcy date is calculated and given with the table, e.g. there are 671 of 6877 bankrupted companies that have a lead time between 3 to 6

[3] The mean of financial statement closure and start time is 1.0034 year, the median of this period is 0.9973 year and minimum & maximum values are 0.0904 and 1.5151 respectively.

months. The mean of time lags of the data set of analysis is approximately 22.7 month; the minimum, maximum and median of lead times are 2.1 months, 49.7 months, and 21.9 months respectively. When these numbers are compared with the previous studies it is figured out that the lead times are satisfactory and long enough for reliability of the analyses. For instance, Ohlson (1980) obtained the same numbers of lead times as 13 months for the mean and 12.5 months for the median.

Table 4.The Time Lag between the Date of Bankruptcy and the Date of Last Relevant Reports in Monthly Basis

Lead Time	Frequency	Percent	Cumulative Frequency	Cumulative Percent
LT<3	6	0.09%	6	0.09%
3<LT<6	671	9.76%	677	9.84%
6<LT<9	697	10.14%	1,374	19.98%
9<LT<12	689	10.02%	2,063	30.00%
12<LT<15	645	9.38%	2,708	39.38%
15<LT<18	653	9.50%	3,361	48.87%
18<LT<21	691	10.05%	4,052	58.92%
21<LT<24	435	6.33%	4,487	65.25%
24<LT<27	369	5.37%	4,856	70.61%
27<LT<30	319	4.64%	5,175	75.25%
30<LT<33	271	3.94%	5,446	79.19%
33<LT<36	112	1.63%	5,558	80.82%
36<LT<39	136	1.98%	5,694	82.80%
39<LT<42	86	1.25%	5,780	84.05%
42<LT<45	27	0.39%	5,807	84.44%
45<LT<48	440	6.40%	6,247	90.84%
48<LT	630	9.16%	6,877	100.00%

The removal of outliers is a challenging issue. However, the nature of financial ratios is totally different because of their distribution. Namely, the distribution of financial ratios is not normal. Therefore, it is decided to remove only the extreme observations which are quite few in accordance with the number of observations. The univariate test of all determined variables[4] is employed, with respect to Box-Plots of the model variables extreme observations with their values has been deleted.

[4] The last set of variables used in the analyses is given in the following part: Variable Selection.

3. Variable Selection

Horrigon (1965) ascertained that one of the most fundamental aspects of the statistical nature of financial ratios is collinearity. Namely, some items in accounting statements tend to move in the same direction as other items, which mean that only a small number of financial ratios are needed to provide us with crucial information of corporate structure. Thus, this small number of ratios must be selected very carefully. A selection of collinear ratios which are related to a dependent variable in the same fashion would conceal and possibly worsen the results of the regression analyses.

According to the Michael A. Poole and Patrick N.O` Farrell (1971) if the absence of multi-collinearity which is one of the fundamental assumptions of the classical linear regression model is not satisfied and accordingly the independent variable is defined as multi-collinear, it results in the individual regression coefficients for each variable which are not identifiable. That means that the standard errors will be so high, and the t-tests are not reliable leading us to the fact that acceptance of null hypothesis is highly possible.

On the other hand, if the main purpose is only to predict the value of dependent variable, then multi-collinearity is not a serious problem. Even though such a problem exists, estimated parameters are still unbiased. Furthermore, if the objective of the analysis is not only prediction but also reliable estimation of the parameters, which complies with the purpose of this study, multi-collinearity will be a serious problem because of the large standard errors of the estimators revealed. Hence, it is obvious that large numbers of financial ratios cannot be used in an analysis. The collinearity of these ratios requires that a careful selection must be utilized.

Table 5. List of Financial Ratios Obtained

CASH FLOW RATIOS	**LIQUID ASSET RATIOS**
1) Cash Flow to Total Liabilities	14) Cash and Bank to Total Asset
2) Cash Flow to Financial Expenditures	15) Total Liquid Asset to Total Asset
PROFITABILITY RATIOS	16) Current Asset to Total Asset
3) Net profit to Net Sales	17) Working Capital to Total Asset
4) Operating Income to Net Sales	**SHORT TERM SOLVENCY RATIOS**
5) Net Income to Total Asset	18) Current Asset less Inventory to Current Liabilities

6) Net Income to Total Equity	19) Current Asset to Current Liabilities
7) Gross Income to Net Sales	20) Current Debt to Inventory
LEVERAGE RATIOS	**ACTIVITY RATIOS**
8) Total Current Liabilities to Total Asset	21) Cash to Sales
9) Total Debt to Total Asset	22) Accounts Receivables to Sales
10) Debt to Equity	23) Inventory to Sales
11) EBIT to Interest Expenditures	24) Liquid Asset to Sales
12) Equity to Asset	25) Current Asset to Sales
SIZE RATIOS	26) Working Capital to Sales
13) Total Asset	27) Total Asset to Sales
14) Number of Employees	28) Cost of Goods sold to Inventory

Some of the ratios can be defined as follows: Net profit to net sales is net margin, operating income to net sales is operating margin, net income to total asset is return on asset (ROA), net income to total Equity is return on equity (ROE), and gross income to net sales is gross margin. In addition to this, the components of some ratios are described in following manner: cash flow is defined as net income plus depreciation, depletion and amortization, working capital is defined as current asset minus current liabilities, liquid asset is defined as cash and bank plus accounts receivable.

The decision as to which variables should be used in a model ought to be based first on theoretical considerations. However, in the case of bankruptcy prediction models, there is no widely accepted theory. Therefore, the choice becomes an empirical issue.

In this study, twenty-eight potentially helpful explanatory variables are compiled due to the fact that these variables are to be found as a significant in past studies dealing with bankruptcy or business failure. While the multi-collinearity problem exists within financial ratios, and a small number of ratios provide us with crucial information, the variables are classified into seven common ratio categories which are consistent with Beaver's (1966) study. These include cash flow, profitability, leverage, size, liquid asset, short-term solvency and activity ratios[5]. Some ratios are excluded because they are simply the transformation of other ratios and at least two variables are selected from each category according to their popularity and performance in an attempt to explain the bankruptcy in previous studies[6]. In addition to these ratios we use a dummy variable called NW as an independent variable which is defined in such a way that equals 1 if total liabilities exceeds total asset, otherwise 0. Since, in our study, bankruptcy as a

[5] Variables are listed in Table 5.
[6] Two variables are selected from liquid asset, activity, leverage and short term solvency ratios. One variable is selected from profitability, size and cash flow ratios.

dependent variable is regarded as liquidation bankruptcy we used the condition of negative net worth as an independent variable[7].

First, the stepwise analysis was adopted. Here, the change in R-square as well as F statistics and significance values are accepted as the criteria of stepwise analysis. The F value for each variable shows whether or not this variable has a statistically significant effect on a model, i.e. if any contribution in the coefficient of determination, (R^2), is statistically significant then the conclusion is that the added variable is necessary to explain the variation in dependent variable. The decision as to parameter is statistically significant or not depends on the probability value of F[8]. According to the F statistics of the general linear model restricted and unrestricted models are evaluated step-by-step for each additional relevant financial ratio[9]. Later, LPM and the logit model are used in order to check the signs and significance of the parameters of these variables in the model as to whether or not financial ratios are the most important predictors in explaining bankruptcy. Hence, a set of eight variables are chosen in conformance with the following considerations: (i) the degree of collinearity of variables between each other, (ii) the significant change in the coefficient of determination (R^2) emanating from the addition of variable to the logistic regression and LPM, (iii) the relative importance of each variable as indicated by the standardized regression coefficients (betas), and (iv) the magnitude of multivariate F ratio conducted on regression coefficient.

Additionally, some other combinations of financial ratios are also checked; stepwise analysis is applied to the best 23 of them. After checking the signs and significance of the coefficients of these variables by running LPM and logistic regressions, 10 of these variables are selected for the repetition of the procedure. It is also known that this procedure gives the best results with at most 10 variables. Without any interference, ROE is found not statistically significant in any of possible combinations. As is seen in the correlation matrix, TDTA and WCTA are highly correlated in the opposite direction with each other. When we include one or both of them into the model then neither the variable CASHCL nor the included variable(s) become(s) significant. So WCTA and CASHCL are deleted.

[7] Ohlson (1980) also used OENEG instead of this variable with the same definition.
[8] R^2 and F ratios for each variable are shown in table 2
[9] The macro codes are given in the Appendix part.

Table 6.The Correlation Matrix of Variables

Var.	NW	TDTA	SIZE	TDTE	WCTA	STA	CFEX	ROE	RMG	LATA	CASHCL
NW	1.00	0.10	-0.23	-0.11	-0.08	0.08	-0.02	0.04	0.00	-0.08	-0.01
TDTA		1.00	-0.08	0.00	**-0.76**	0.12	0.00	0.00	0.00	0.00	0.00
SIZE			1.00	0.06	0.06	-0.07	0.04	0.00	0.00	-0.20	0.00
TDTE				1.00	0.00	0.00	0.00	-0.40	0.00	-0.03	0.00
WCTA					1.00	-0.07	0.00	0.00	0.00	0.01	0.00
STA						1.00	0.00	0.01	0.00	0.00	-0.01
CFEX							1.00	0.02	0.01	0.03	0.00
ROE								1.00	-0.17	0.01	0.00
RMG									1.00	0.00	0.00
LATA										1.00	0.06
CASHCL											1.00

The nine variables including dependent variable and dummies for industries used in models are as follows:

1) **SIZE =** log (total asset/100). A logarithmic transformation was applied to help normalize the distribution of the variable because of the outlier it exhibits.
2) **LATA =** Liquid assets divided by total assets. It is a measure of company's short term solvency.
3) **RMG =** Gross profit minus cost of sales divided by sales turnover. It is a profit margin (operating margin) which measures the size of profit in relation to sales turnover.
4) **CFEX =** Cash flow divided by financial expenditures. This ratio is also divided by 100 to make it consistent with other ratios. It is a measure of company's financial flexibility to invest in itself.
5) **STA**= Sales divided by total assets is a measure of firm's ability to generate sales from its total assets.
6) **TDTE** = Total debt divided by total assets, which is a measure of company's leverage.
7) **TDTA =** Total debt divided by total assets, which is another leverage ratio which measures the percentage of the company's total assets which are financed with total debt.
8) **NW** is a dummy variable which is defined in such a way that one if a company has negative equity, zero otherwise.
9) **BR** is a dummy variable used as a dependent variable and it is defined in a way that one if firm went bankrupt, zero otherwise.

10) INDUSTRY DUMMIES d1-d6 There are six dummies used to represent seven industries and named as d1, d2, d3, d4, d5, and d6. The dummy variables are defined with dk which equals to 1 if the observation is in the industry set K consisting the industries coded with A, B, C, D, E, and F; otherwise it equals to 0. If the observation is not in the industry set of K then it belongs to the industry G which is wholesale and retail trade.

The mean and standard deviation of ratios were computed for bankrupt and non-bankrupt firms. The comparison of mean values for both groups is called profile analysis which should not be regarded as a predictive test. According to Altman (1968) and Beaver (1966), it can be a convenient way of capturing an opinion about the general relationships and differences between the bankrupt and non-bankrupt firms.

The table of profile analysis shows the means of the seven variables for bankrupt and non-bankrupt firms with t statistics. In order to test the differences of means within two groups, an independent group t-test is employed under the assumption that variances for both groups are not the same[10]. It is clear that ratios deteriorate as one moves from non-bankrupt firms to bankrupt firms. Compared to non-bankrupt firms, bankrupt firms are typically small, highly leveraged, having poor financial flexibility and liquidity. However, STA appears strange since it is believed that the more companies have the ability to generate sales from their assets, the less likely will bankruptcy occur.

The t statistics for all variables except for RMG are statistically significant at 5 % significance level, meaning that the differences in mean values of these variables between two groups are statistically significant. Put differently, the greater t-values, the better the variables in terms of univariate predictive ability. Some ratios such as LATA and SIZE have higher univariate discriminatory power than others, indicating that their contribution to the estimated probability of bankruptcy is assumed to be more than others in multivariate analysis.

[10] This is also tested by Folded F test by SAS and the hypothesis that variance are equal for both groups is rejected.

Table 7.Profile Analysis

	Bankrupt Firms			Non-Bankrupt Firms				
Variable	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**t Value**	**Pr > |t|**
CFEX	6877	-0.05212	1.23788	170743	0.29336	1.91080	22.11	<.0001
STA	6877	3.03969	3.07900	170743	2.40104	2.30776	-17.01	<.0001
RMG	6877	-1.16295	63.22504	170743	-0.27358	14.51579	1.17	0.2439
SIZE	6877	0.14059	0.01387	170743	0.14665	0.01581	35.36	<.0001
TDTE	6877	5.37612	35.84874	170743	3.82260	20.84104	-3.57	0.0004
LATA	6877	0.08757	0.15098	170743	0.18222	0.22040	49.89	<.0001
TDTA	6877	1.10377	1.69992	170743	0.68306	0.97970	-20.39	<.0001

Previous studies would suggest that the sign of the coefficients of the different financial ratios used multivariate analysis should be as in the table given below.

Table 8.Expected sign of variables

Positive	Negative	Indeterminate
TDTA	CFEX	NW
TDTE	STA	
	RMG	
	SIZE	
	LATA	

NW used here is a discontinuity correction for leverage ratios. Ohlson (1980) cited that a firm that has a negative net worth is a special case. The condition of non-bankruptcy would tend to depend on many sophisticated factors, and the effect of extreme leverage condition needs to be corrected. A positive sign accounts for almost certain bankruptcy, while a negative sign accounts for the situation which is very bad due to TDTA and TDTE, but not that bad.

4. Empirical Results

This section presents the results from 4 different models and the Davidson and MacKinnon J test, one of model specification tests for non-nested models. Model 1 and model 2 are linear probability and logit models respectively. These models use selected financial ratios and dummy variables representing industry groups. Model 3 and model 4 are linear probability and logit models using industry relative financial ratios. However, the interpretation of coefficients is not made directly in logit model compared to LPM. A parameter estimates just give us the expected change in logit, not in probability of bankruptcy. The purpose of this study is not to examine the effect of variables to logit, instead, the purpose is to examine the direct effect of variables to bankruptcy probability, leading us to the marginal effects and elasticities.

The following industry relative ratios are used in model 3 and 4:

$ratX_i = X_i / X_{id}$ where

X_i = ratio i,

d = industry d,

X_{id} = industry d's median for ratio i.[11]

The main reason to use industry relative ratios in models is to control industry differences. Horrigon (1965) contends that one of the common characteristics regarding the statistical nature of financial ratios is the extent of the dispersion in ratio distribution within industries. Wide dispersion in financial ratio distributions may make discrimination between firms based on the financial ratios difficult. One remedy to solve that problem, according to Horrigon(1965), is industry stratification. Since this paper is regarding bankruptcy prediction models using accounting ratios, this subject should be regarded as an important factor affecting the performance of models regarding bankruptcy. Altman and Izan (1984) used industry relative ratios in discriminant analysis for approximately 100 Australian firms and captured robust results.

It should be remembered that in bankruptcy prediction models, since there is not a widely accepted theory as to whether which variables should be used, then model specification ought to be an empirical issue. As for models using industry relative ratios and models using unadjusted ratios, one can test these models so as to which one should be used by means of Davidson and MacKinnon J Test which a is model specification test for non-nested models.

There are two sets of independent variables for each LPM and logit model. X_1 (model 1 and 2- unadjusted financial ratios and dummy variables for selected industries) and X_2 (model 3 and 4- industry adjusted financial ratios). Models 1-3 and models 2-4 are being compared separately by J test for predicting bankruptcy probabilities. The null

[11] Industry median ratios are calculated from our row data and the main reason to use median values instead of mean is that the the distribution of financial variables are highly skewed.

hypothesis examines prediction of probability of bankruptcy based on one model. The alternative hypothesis combines the two models. Hence, there are two null hypotheses, $H1_0$ and $H2_0$. Since two null hypotheses based on two models are tested independently by z test, one can follow the possible outcomes; accepting unadjusted model, accepting adjusted model, accepting or rejecting both models.

$$H1_0 : Y = X_1\beta_1$$
$$H1_a : Y = X_1\beta_1 + X_2\beta_2$$
(Model 2 does not add incrementally)

$$H2_0 : Y = X_2\beta_2$$
$$H2_a : Y = X_2\beta_2 + X_1\beta_1$$
(Model 1 does not add incrementally)

Platt and Platt (1991) carried out the logit model to compare the predictive accuracy of models with relative industry ratios and unadjusted ratios by means of Davidson and MacKinnon J Test, which resulted in a better performance of model with industry relative ratios over unadjusted model.

Table 9 summarize the empirical findings of four models [12]. The results indicate that all parameters in model 2 are statistically significant at 5% significance level, which contend that all selected variables in model 2 have additional information in order to explain bankruptcy behaviour. Moreover, parameters of d4 and d5 in model 1, and parameters of industry adjusted RMG ratios in model 3 and 4 are not statistically significant at 5% significance level.

One can notice that TDTA in all models and STA in model 1 and 2 are not as expected in accordance with the previous studies regarding prediction of bankruptcy. Since STA, so called capital turnover ratio, is illustrating the company’s ability to generate sales from its asset, the more sales generated from assets the less likely company goes bankruptcy. This is also the case for TDTA, which means that the more debt the company has the more likely it goes bankruptcy.

[12] This table summarizes the results, all tables regarding four models are presented Appendix A

Table 9. Results of LPMs and Logit Models

	Industry Unadjusted				Industry Adjusted		
Exp. Var.		**Model 1**	**Mode 2**	**Exp. Var.**		**Model 3**	**Model 4**
NW*	dy/dx	0,13033	0,07847	**NW***	dy/dx	0,14088	0,13558
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
	ey/ex	0,20479	0,08882		ey/ex	0,22137	0,10948
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
CFEX	dy/dx	-0,00119	-0,00246	**ratCFEX**	dy/dx	-0,00009	-0,00012
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
	ey/ex	-0,00861	-0,02505		ey/ex	-0,00914	-0,01524
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
STA	dy/dx	0,00169	0,00084	**ratSTA**	dy/dx	-0,00419	-0,00256
	P>|z|	<.0001	<.0001		P>|z|	0,001	0,012
	ey/ex	0,10604	0,07384		ey/ex	-0,06554	-0,04641
	P>|z|	<.0001	<.0001		P>|z|	0,001	0,012
RMG	dy/dx	-0,00008	-0,00002	**ratRMG**	dy/dx	0,00000	0,00000
	P>|z|	0,0005	0,045		P>|z|	**0,847**	**0,594**
	ey/ex	0,00066	0,00020		ey/ex	-0,00003	-0,00018
	P>|z|	0,001	0,045		P>|z|	0,847	0,594
SIZE	dy/dx	-0,65089	-0,53372	**ratSIZE**	dy/dx	-0,28940	-0,26816
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
	ey/ex	-2,46147	-2,84784		ey/ex	-0,00467	-0,00503
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
TDTE	dy/dx	0,00044	0,00015	**ratTDTE**	dy/dx	0,00091	0,00034
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
	ey/ex	0,04453	0,02083		ey/ex	0,02648	0,01152
	P>|z|	<.0001	<.0001		P>|z|	<.0001	<.0001
LATA	dy/dx	-0,06526	-0,07836	**ratLATA**	dy/dx	0,00020	0,00012
	P>|z|	<.0001	<.0001		P>|z|	<.0001	0,007
	ey/ex	-0,30098	-0,50990		ey/ex	0,06610	0,04418
	P>|z|	<.0001	<.0001		P>|z|	<.0001	0,007
TDTA	dy/dx	-0,00089	-0,00078	**ratTDTA**	dy/dx	-0,00104	-0,00051
	P>|z|	0,067	<.0001		P>|z|	0,028	0,053
	ey/ex	-0,01611	-0,01996		ey/ex	-0,01468	-0,00831
	P>|z|	0,067	<.0001		P>|z|	0,028	0,053
d1*	dy/dx	-0,02460	-0,01870	•	•	•	•
	P>|z|	<.0001	<.0001				
	ey/ex	-0,03838	-0,06130				
	P>|z|	<.0001	<.0001				
d2*	dy/dx	-0,02789	-0,01649	•	•	•	•
	P>|z|	0,008	<.0001				
	ey/ex	-0,00133	-0,00168				
	P>|z|	0,008	0,016				
d3*	dy/dx	-0,01712	-0,01421	•	•	•	•
	P>|z|	0,024	<.0001				
	ey/ex	-0,00156	-0,00254				
	P>|z|	0,024	0,016				
d4*	dy/dx	-0,00154	-0,00204	•	•	•	•
	P>|z|	**0,176**	0,012				
	ey/ex	-0,00966	-0,01838				
	P>|z|	0,176	0,014				
d5*	dy/dx	-0,00807	-0,01172	•	•	•	•
	P>|z|	**0,237**	0,011				
	ey/ex	-0,00093	-0,00245				
	P>|z|	0,237	0,054				
d6*	dy/dx	-0,00299	-0,00216	•	•	•	•
	P>|z|	0,012	0,011				
	ey/ex	-0,01601	-0,01667				
	P>|z|	0,012	0,013				

(*) dy/dx is for discrete change of dummy variable from 0 to 1

The negative coefficient of cash flow to financial expenditure (CFEX) ratio indicates that the marginal effect of this variable to probability of bankruptcy event is negative. In other words, bankruptcy is more likely when the company has less financial flexibility to invest in itself. Profitability of a company also has a negative effect to probability of bankruptcy since the coefficient of the operating margin (RMG) is also negative.

Positive coefficient of total debt to total equity (TDTE) implies that a company is more likely to go bankruptcy if it is highly leveraged. The negative coefficient of liquid asset to total asset (LATA) ratio shows negative correlation between liquidity of a company and probability of bankruptcy. Another important factor is size in terms of assets which have a negative coefficient saying that size has a negative marginal affect to probability of bankruptcy. In other words, the company is more likely to go bankruptcy if it is relatively small.

It is obvious that the dummy variable (NW) has a positive effect to bankruptcy. As a result of the values of parameter estimates in all models, it can be said that this variable has the most powerful effect of explaining bankruptcy behaviour in all models.

The industry dummies also important factors explaining the bankruptcy event. It seems in model 2 that a company is more likely to enter bankruptcy if it operates in the wholesale and retail trade industry. In model 1, a company is less likely to enter bankruptcy if it operates in thefollowing industries rather than in wholesale and retail trade industry since they have negative coefficients[13]:

- **i)** Agriculture, hunting and forestry (d1)
- **ii)** Fishing (d2)
- **iii)** Mining and quarrying (d3)
- **iv)** Construction (d6)

The coefficients of manufacturing industry (d4) and electricity, gas and water supply industry (d5) are not statistically significant at 5% level of significance, indicating that the mean probability of bankruptcy in these two industries, and the wholesale and retail trade industry are about the same, i.e. for a company being in one of these three industries does not affect the bankruptcy probability.

[13] The wholesale and retail trade industry is chosen as a benchmark category as a result of high bankruptcy frequency compared to others.

Table 10 shows results of J test. Since the both of the null hypotheses that adjusted and unadjusted models do not add incrementally are rejected for LPM and the logit model. We can therefore conclude that both industry adjusted and unadjusted models help us in explaining the behaviour of bankruptcy event. According to Gujarati (2003) the data may not be rich enough to discriminate between two models if both models are accepted according to J test.[14]

Table 10.J test Results by Model Specification for LPM and Logit Model

J test Results by Model Specification for LPM (Model 1-3)		
Estimate Parameter	**z-ratio**	**p-value**
$H1_0$ Industry-relative ratios do not add incrementally		
-0.274840	**-6.06**	**<0.001**
$H2_0$ Unadjusted ratios do not add incrementally		
0.605285	**33.55**	**<0.001**
J test Results by Model Specification for Logit Model (Model 2-4)		
Estimate Parameter	**z-ratio**	**p-value**
$H1_0$ Industry-relative ratios do not add incrementally		
-0.274840	**-6.06**	**<0.001**
$H2_0$ Unadjusted ratios do not add incrementally		
0.605285	**33.55**	**<0.001**

Since model specification tests did not give exact results as to which model should be used in order to explain the behaviour of the bankruptcy event, the purpose of study should be the main determining factor. Thus, if the purpose of the study is to capture industry effects on bankruptcy one can use the model including unadjusted ratios and dummy variables for industries. But if the purpose is to calculate only the marginal effects of financial ratios to probability of bankruptcy without looking at industry effect, then one can use the model including adjusted financial ratios since this kind of model controls the industry differences.

To summarize, the company is more likely to go bankrupt if it is unprofitable, small, highly leveraged, and has liquidity problems, negative equity situation and less financial flexibility to invest in itself. Additionally, a company is more likely to enter bankruptcy if it operates in the wholesale and retail trade industry among the selected industries in our sample.

[14] Gujarati also refer this statement to Kmenta, op. cit., p.597.

5. Evaluation of Predictive Accuracy

One can evaluate the predictive accuracy by looking at the percent correctly predicted statistic which is shown in Table 11. Suppose, for example, that the cut off value is 0.4., the company is predicted as a bankrupt if its probability of bankruptcy is higher than this cut off point, if not it is assumed to be nonbankrupt. At this point, the percentage of correctly predicted statistics is 96.1 percent for all models. To rely on this number is misleading since if we classified all firms as nonbankrupt, then 96.13 percent (170743/ (170743+6877)) would be correctly classified due to the extremely high number of nonbankrupt firms compared to the small number of nonbankrupt contained in data sample.

In order to get a clearer picture of the prediction accuracy of the models, it is helpful to define type 1 and type 2 errors. Type 1 error takes place when a company goes bankrupt but is predicted to be non-bankrupt and type 2 errors takes place when a company is non-bankrupt but is predicted to be bankrupt. It is obvious that type 1 and type 2 error rates depend on the number of firms that are predicted to go bankruptcy. As can be seen from the classification tables, a type 1 error rate is relatively low and a type 2 error rate is relatively high for a large number of firms that are predicted to go bankruptcy. Apparently, the number of firms predicted to go bankrupt depends on the cut off value chosen. Thus, it seems tricky, that is, one can increase the number of firms as a bankrupt by decreasing the cut off value since the consequence of having a type 1 error seems more serious than having a type 2 error.

Lennox (1999) stated that type 1 and type 2 error rates depend on the sample selection criterion, i.e. studies in which samples that have an equal number of failing and non-failing companies have much smaller error rates. Since the sample which this study uses does not have a proportional rate of bankruptcy event, relatively large error rates are captured.

Table 11. Classification Table for 4 Models

LPM(Model 1)				
Prob Level	**0.1**	**0.2**	**0.3**	**0.4**
Correct Event	1,824	114	7	1
Correct Non-event	161,67	170,115	170,719	170,742
Incorrect Event	9,073	628	24	1
Incorrect Non-event	5,053	6,763	6,87	6,876
Correct	92.0%	95.8%	96.1%	96.1%
Sensitivity	26.5%	1.7%	0.1%	0.0%
Specifity	94.7%	99.6%	100.0%	100.0%
TYPE 1	3.0%	3.8%	3.9%	3.9%
TYPE 2	83.3%	84.6%	77.4%	50.0%

LOGIT(Model 2)				
Prob Level	**0.1**	**0.2**	**0.3**	**0.4**
Correct Event	1,783	264	33	13
Correct Non-event	162	169	171	171
Incorrect Event	8,863	1,65	162	73
Incorrect Non-event	5,094	6,613	6,844	6,864
Correct	92.1%	95.3%	96.1%	96.1%
Sensitivity	25.9%	3.8%	0.5%	0.2%
Specifity	94.8%	99.0%	99.9%	100.0%
TYPE 1	3.1%	3.8%	3.9%	3.9%
TYPE 2	83.3%	86.2%	83.1%	84.9%

INDUSTRY ADJUSTED LPM (Model 3)				
Prob Level	**0.1**	**0.2**	**0.3**	**0.4**
Correct Event	1818	9	1	0
Correct Non-event	161699	170657	170729	170739
Incorrect Event	9047	86	14	4
Incorrect Non-event	5059	6868	6876	6877
Correct	92.1%	96.1%	96.1%	96.1%
Sensitivity	26.4%	0.1%	0.0%	0.0%
Specifity	94.7%	99.9%	100.0%	100.0%
TYPE 1	3.0%	3.9%	3.9%	3.9%
TYPE 2	83.3%	90.5%	93.3%	100%

INDUSTRY ADJUSTED LOGIT MODEL(Model 4)				
Prob Level	**0.1**	**0.2**	**0.3**	**0.4**
Correct Event	1,67	568	60	24
Correct Non-event	163	168	170	171
Incorrect Event	7,683	2,899	310	99
Incorrect Non-event	5,207	6,309	6,817	6,853
Correct	92.7%	94.8%	96.0%	96.1%
Sensitivity	24.3%	8.3%	0.9%	0.3%
Specifity	95.5%	98.3%	99.8%	99.9%
TYPE 1	3.1%	3.6%	3.8%	3.9%
TYPE 2	82.1%	83.6%	83.8%	80.5%

Correct: the percentage of correct classification
Sensitivity: the proportion of correctly classified events divided by the total number of events
Specificity: the number of correctly classified non-events divided by the total number of non events

Here, the purpose is to have a minimum sum of error rates. It is clear that the 0.4 cut off value leads to a minimum sum of errors for both unadjusted models, model 1 and 2. Additionally, the unadjusted LPM is preferable to other models in terms of type 1 and type 2 error rates since it has minimum errors for a cut off value of 0.4. But it should be remembered that the aim of this study is not the comparison of models with respect to their accuracy.

6. SUMMARY & CONCLUSION

The main purpose of this study was to examine the effects of financial ratios and industries to bankruptcy events that occurred between 2002 and 2006 in Swedish market. This is also a kind of analysis which investigates the general characteristics of a company that is likely to go bankrupt. Even if the comparison of this study with previous ones conducted on different markets and in different years is not appropriate but consistent with previous studies, this study shows that size, financial flexibility, profitability, liquidity and leverage ratios statistically significantly affect the probability of bankruptcy. Put differently, the company is more likely to go bankrupt if it is unprofitable, small, highly leveraged, has liquidity problems and suffers financial flexibility to invest in itself. Negative equity situation is also an important factor, namely, bankruptcy is more likely if a company has negative equity.

In addition to this, this study investigated the industry effects of the probability of bankruptcy. The wholesale and retail trade market was chosen as a benchmark industry since this sector contains a higher bankruptcy frequency compared to others. It was encountered in LPM that there is no difference for a company being in the electricity, gas and water supply industry or the wholesale-retail trade industry to affect probability of bankruptcy. However, all other selected industries are statistically significantly different from wholesale-retail trade industry in affecting the probability of event. It can be stated for two models, LPM and logit model, that bankruptcy is more likely if a company operates in the latter.

A model specification test was also employed to see whether or not models using industry normalized ratios have better performance compared to others. In order to generate industry adjusted ratios, financial ratios were divided by industry medians. The main reason behind using the median was that all financial ratios are highly skewed. Davidson and McKinnon J test is used for comparison but unfortunately the test did not give proper information as to which model should be used.

The overall significance of models are confirmed by F statistics and likelihood ratios, which means that models are successful in explaining the variation in probability of bankruptcy. In addition to this, parameters of all financial ratios in four models are statistically significant with the exception of RMG ratios which are in industry adjusted models. However, the magnitudes of marginal effects of financial ratios to the probability of bankruptcy are small, leading us to further suggestions such as using variables bearing information based on equity prices, economic conditions or business cycles, and non quantitative variables including managerial elements in addition to financial ratios.

More robust results can be obtained by carrying out analysis on sub samples without ruining the randomness of the data. This study avoided the matching approach. Thus, by doing so, it was able to calculate the marginal effects of company size and industries on the probability on bankruptcy, since Lennox (1999) states that in small samples over fitting problem can arise.

The models in this study have relatively large type 2 errors explaining predictive accuracy in part. One of the possible reason why this is the case here is that the frequency of bankruptcy events is almost the same with frequency of population. The predictive quality of the models may be improved in cases of robust estimation, match-making procedure or analysis based on sub samples.

REFERENCES

Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, Vol. 23, No.4.(Sep.,1968), pp.598-609.

Altman, E. (1971). Railroad Bankruptcy Propensity. The Journal of Finance, Vol.26, No.2 (May, 1971), pp. 333-345.

Altman, E., Eom, Y., Kim, W.(1995), Failure Prediction: Evidince from Korea. Journal of International Financial Management and Accounting 6:3(1995).

Altman, E., Haldeman, R., Narayan, P.(1977), A new model to identify bankruptcy risk of corporations. Journal of Banking Finance 1(1977)29-54. North-Holland Publishing Company.

Altman, E., Margaine, M., Schlosser, M., Vernimmen, P.(1974), Financial and Statistical Analysis for Commercial Loan Evaluation: A French Experience. The Journal of Financial and Quantitative Analysis, Vol.9, No.2 (Mar., 1974), pp. 195-211.

Altman, E., Saunders, A.(1998). Credit risk measurement: Developments over the last 20 years. Journal of Banking and Finance 21(1998) 1721-1742.

Beaver,W.(1966). Financial Ratios As Predictors of Failure. Journal of Accounting Research, Vol. 4, Empirical Research in Accounting: Selected Studies (1966), pp. 71-111.

Blum, M.(1974). Failing Company Discriminant Analysis. Journal of Accounting Research, Vol. 12, No.1.(Spring, 1974), pp. 1-25. The Institute of Professional Accounting, Graduate School of Business, University of Chicago.

Casey, C., J., McGee, E., V., Stickney, P., C.(1986). Discriminant between Reorganized and Liquidated Firms in Bankruptcy. The Accounting Review, Vol. 61, No. 2 (1986), pp. 249-262. American Accounting Association.

Deakin, E.(1972). A Discriminant Analysis of Predictors of Business Failure. Journal of Accounting Research, Vol.10, No.1, (Spring, 1972), pp. 167-179. The Institute of Professional Accounting, Graduate School of Business, University of Chicago.

DeMaris, A., Teachman, J., Morgan, P.S.(1990), Interpreting Logistic Regression Results: A Critical Commentary. Journal of Marriage and the Family, Vol. 52, No. 1.(Feb., 1990), pp. 271-277. National Council on Family Relations.

Edmister, R.(1972), An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. The Journal of Financial and Quantitative Analysis, Vol.7, No. 2, Supplement: Outlook for the Securities Industry. (Mar., 1972), pp.1477-1493.

Eisenbeis, R, A.(1977), Pitfalls in the Application of Discriminant Analysisin Business, Finance, and Economics. The journal of Finance, Vol. 32, pp. 875-900.

Gujarati, D.(2003). Basic Econometrics. Fourth Edition,The Mc Graw-Hill Company.

Horrigon, O.J.(1965). Some Emprical Bases of Financial Ratio Analysis. The Accounting Review, Vol. 40, No. 3. (Jul., 1965), pp. 558-568. American Accounting Association.

Lennox, C. (1999). Identifying Failing Companies: A Reevaluation of the Logit, Probit and DA Approaches. North Holland. Elsevier Science Inc.

Maddala, G.,S.(1991). A Perspective on the Use of Limited-Dependent and Qualitative Variables Models in Accounting Research. The Accounting Review, Vol. 66, No. 4. (Oct., 1991), pp. 788-807. American Accounting Association.

Martin, D.(1977). Early Warning of Bank Failure: A logit regression approach. Journal of Banking and Finance 1(1977) 249-276. North-Holland Publishing Company.

Meyer, P., A, Pifer, H., W.(1970). Prediction of Bank Failures. The journal of Finance, Vol. 25, No.4, (Sep., 1970), pp. 853-868.

Noreen, E. (1988). An Empirical Comparison of Probit and OLS Regression Hypothesis Tests. Journal of Accounting Research, Vol. 26, No. 1. (Spring, 1988), pp. 119-133. The Institute of Professional Accounting, Graduate School of Business, University of Chigago.

Northon, L., Smith, E.(1979). A Comprasion of General Price Level and Historical Cost Financial Astatment in the Prediction of Bankruptcy. The Accounting Review, Vol. 54, No. 1 (Jan., 1979), pp. 72-87. American Accounting Association.

Ohlson, J.(1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, Vol. 18, No. 1. (Spring, 1980), pp. 109-131. The Institute of Professional Accounting, Graduate School of Business, University of Chicago.

Platt D., H., Platt, B., M.(1991). A note on the use of industry-relative ratios in bankruptcy prediction. Journal of Banking and Finance 15 (1991) 1183-1194. North Holland.

Poole, A. M., O`Farell,N.P.(1971). The Assumptions of the Linear Regression Model. Transactions of the Institute of British Geographers, No. 52. (Mar.,1971), pp. 145-158. The Royal Geographical Society.

Press, J.,S., Wilson, S.(1978). Choosing Between Logistic Regression and Discriminant Analysis. Journal of the American Statistical Association, Vol. 73, No. 364. (Dec., 1978), pp. 699-705. American Statistical Association.

Smith, D.,L., Lawrence, C., E.(1995).Forecasting losses on a liquidating long-term loan portfolio. Journal of Banking and Finance 19(1995) 959-985.

Stone, M., Rasp J.(1991). Tradeoff in Choice between Logit and OLS for Accounting Choice Studies. The Accounting Review, Vol.66, No. 1. (Jan., 1991), pp.170-187. American Acounting Association.

Suzuki, S., Wright, W., R.(1985). Financial Structure and Bankruptcy Risk in Japanise Companies. Journal of International Business Studies, Vol. 16, No. 1. (Spring, 1985), pp. 97-110. Palgrave Macmillan Journals.

Taffler, R.,J.(1982). Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. Journal of the Royal Statistical Society. Series A(General), Vol. 145, No. 3, (1982), pp. 342-358. Royal Statistical Society.

West, C., R.(1984). A Factor-Analytic Approach to Bank Condition. Journal of Banking and Finance 9 (1985) 253-266. North Holland.

Wiginton, J.(1980). A note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior. The Journal of Financial and Quantitative Analysis, Vol. 15, No. 3 (Sep., 1980), pp. 757-770. University of Washington School of Business Administration.

Zmijewski, E., M.(1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research, Vol. 22, Studies of Current Econometric Issues in Accounting Research. (1984), pp. 59-82. The Institute of Professional Accounting, Graduate School of Business, University of Chigago.

APPENDIX A: Empirical Results

Table 12. Results of Linear Probability Model (Model 1)

BR	Coef.	Std. Err.	z	P>|z|	ey/ex	Std. Err.	z	P>|z|
NW	0,13033	0,00214	60,97	<.0001	0,20479	0,00411	49,81	0
CFEX	-0,00119	0,00024	-4,94	<.0001	-0,00861	0,00175	-4,93	0
STA	0,00169	0,00020	8,45	<.0001	0,10604	0,01260	8,41	0
RMG	-0,00008	0,00002	-3,47	0,0005	0,00066	-0,00019	-3,47	0,001
SIZE	-0,65089	0,03090	-21,06	<.0001	-2,46147	0,12028	-20,46	0
TDTE	0,00044	0,00002	20,65	<.0001	0,04453	0,00222	20,09	0
LATA	-0,06526	0,00214	-30,56	<.0001	-0,30098	0,01045	-28,81	0
TDTA	-0,00089	0,00049	-1,83	0,067	-0,01611	0,00878	-1,83	0,067
d1	-0,02460	0,00197	-12,51	<.0001	-0,03838	0,00310	-12,38	0
d2	-0,02789	0,01045	-2,67	0,008	-0,00133	0,00050	-2,67	0,008
d3	-0,01712	0,00760	-2,25	0,024	-0,00156	0,00069	-2,25	0,024
d4	-0,00154	0,00114	-1,35	**0,176**	-0,00966	0,00714	-1,35	0,176
d5	-0,00807	0,00682	-1,18	**0,237**	-0,00093	0,00078	-1,18	0,237
d6	-0,00299	0,00119	-2,51	0,012	-0,01601	0,00637	-2,51	0,012
Intercept	0,13547	0,00480	28,21	<.0001				

Coef. =dy/dx Ey/ex* = elasticity

Number of obs = 177620	F(14,177605) = 536.77 (Pr<.0001)
Prob > F = 0.0000	R-squared = 0.0406
Adj R-squared = 0.0405	Root MSE = .18897

Table 13. Results of Logit Model (Model 2)

BR	Coef.	z	P>|z|	dy/dx	z	P>|z|	ey/ex	z	P>|z|
NW	1,50108	43,2	<.0001	0,07847	25,2	0	0,08882	43,06	0
CFEX	-0,09200	-10,47	<.0001	-0,00246	-10,54	0	-0,02505	-10,46	0
STA	0,03130	7,76	<.0001	0,00084	7,75	0	0,07384	7,76	0
RMG	-0,00066	-2	0,0452	-0,00002	-2	0,045	0,00020	-2	0,045
SIZE	-19,99872	-21,59	<.0001	-0,53372	-21,98	0	-2,84784	-21,55	0
TDTE	0,00552	15,43	<.0001	0,00015	15,23	0	0,02083	15,42	0
LATA	-2,93631	-31,38	<.0001	-0,07836	-35,82	0	-0,50990	-31,17	0
TDTA	-0,02934	-3,83	0,0001	-0,00078	-3,83	0	-0,01996	-3,83	0
d1	-1,04349	-12,59	<.0001	-0,01870	-19,87	0	-0,06130	-12,58	0
d2	-0,93301	-2,4	0,0163	-0,01649	-3,89	0	-0,00168	-2,4	0,016
d3	-0,74127	-2,41	0,0161	-0,01421	-3,5	0	-0,00254	-2,41	0,016
d4	-0,07779	-2,47	0,0136	-0,00204	-2,52	0,012	-0,01838	-2,47	0,014
d5	-0,56692	-1,92	0,0543	-0,01172	-2,55	0,011	-0,00245	-1,92	0,054
d6	-0,08271	-2,49	0,0129	-0,00216	-2,54	0,011	-0,01667	-2,49	0,013
Intercept	-0,15210	-1,09	0,2744						

dy/dx = marginal effect and Ey/ex = elasticity

Logistic regression	Number of obs = 177620
Log likelihood = -26393.544	LR chi2(14) = 5417.78
Prob > chi2 = 0.0000	Pseudo R2 = 0.0931

Table 14. Results of Industry adjusted LPM (Model 3)

BR	Coef.	Std. Err.	z	P>|z|	ey/ex	Std. Err.	z	P>|z|
NW	0,14088	0,00198	71,19	0	0,22137	0,00404	54,82	0
ratCFEX	-0,00009	0,00001	-7,36	0	-0,00914	0,00125	-7,33	0
ratSTA	-0,00419	0,00129	-3,25	0,001	-0,06554	0,02019	-3,25	0,001
ratRMG	0,00000	0,00000	0,19	0,847	-0,00003	-0,00013	0,19	0,847
ratSIZE	-0,28940	0,04080	-7,09	0	-0,00467	0,00066	-7,07	0
ratTDTE	0,00091	0,00006	16,03	0	0,02648	0,00168	15,76	0
ratLATA	0,00020	0,00006	3,52	0	0,06610	0,01877	3,52	0
ratTDTA	-0,00104	0,00047	-2,19	0,028	-0,01468	0,00669	-2,19	0,028
Intercept	0,03020	0,00056	53,81	0				

Ey/ex* = elasticity F(8,177610) = 711.66 Prob > F = 0.0000

R-squared = 0.0311 Adj R-squared = 0.0310 Root MSE = .18991

Table 15. Results of Industry Adjusted Logit Model (Model 4)

BR	Coef.	Wald Chi-Square	Pr > ChiSq	dy/dx	z	P>|z|	ey/ex	z	P>|z|
NW	1,86170	3576,7782	<.0001	0,13558	35,32	0	0,10948	59,23	0
ratCFEX	-0,00387	84,5833	<.0001	-0,00012	-9,23	0	-0,01524	-9,19	0
ratSTA	-0,07930	6,3711	0,0116	-0,00256	-2,52	0,012	-0,04641	-2,52	0,012
ratRMG	0,00004	0,2779	0,5981	0,00000	0,53	0,594	-0,00018	0,53	0,594
ratSIZE	-8,31530	49,0445	<.0001	-0,26816	-7,03	0	-0,00503	-7	0
ratTDTE	0,01050	134,5481	<.0001	0,00034	11,57	0	0,01152	11,6	0
ratLATA	0,00363	7,3695	0,0066	0,00012	2,71	0,007	0,04418	2,71	0,007
ratTDTA	-0,01570	3,7349	0,0533	-0,00051	-1,93	0,053	-0,00831	-1,93	0,053
Intercept	-3,45960	46823,695	<.0001						

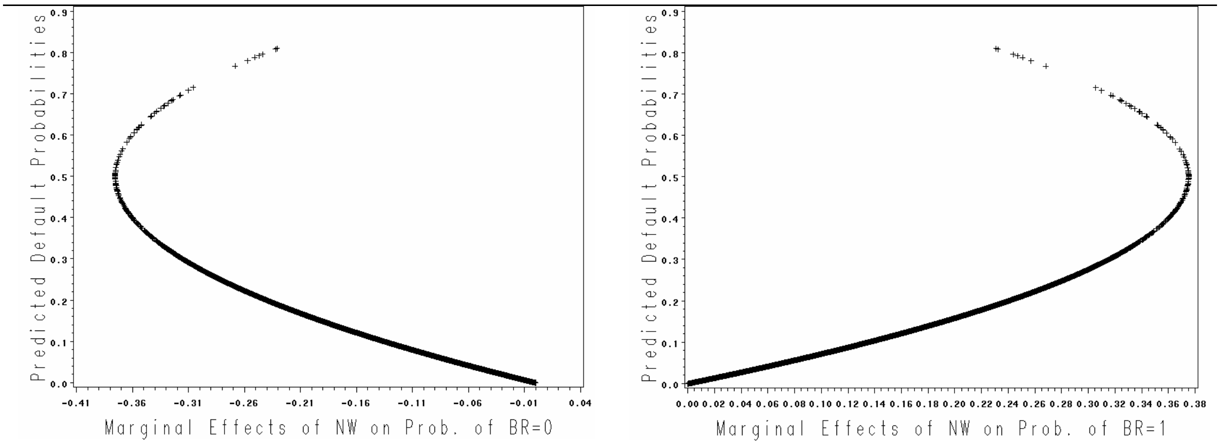
dy/dx = marginal effect and ey/ex = elasticity

Log likelihood = -27435.661 LR chi2(8) = 3333.55

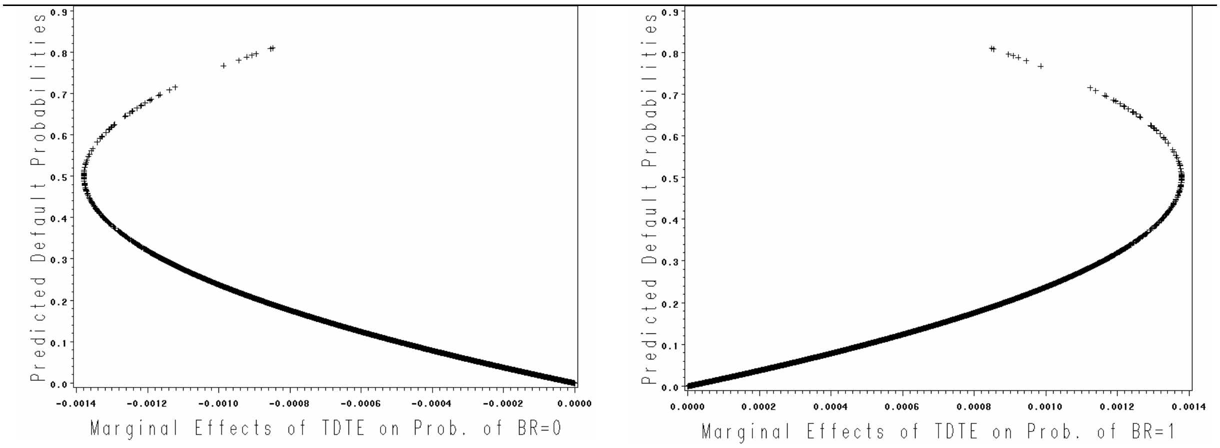
Prob > chi2 = 0.0000 Pseudo R2 = 0.0573

Table 16. The graphs of marginal effects of unadjusted logit model[15]

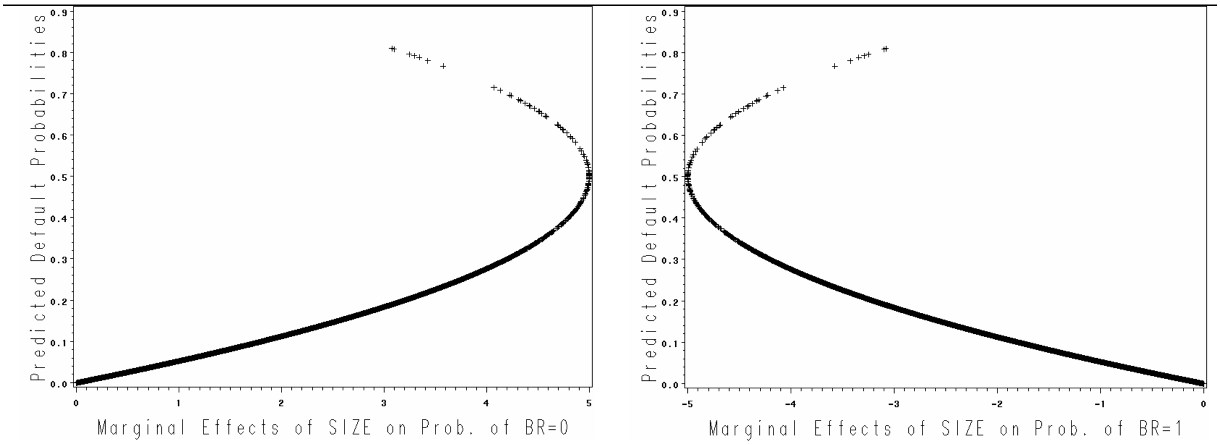
Predicted Probabilities vs Marginal Effects of NW on Prob. of BR=0 & BR=1



Predicted Probabilities vs Marginal Effects of TDTE on Prob. of BR=0 & BR=1



Predicted Probabilities vs Marginal Effects of SIZE on Prob. of BR=0 & BR=1



[15] The most effective variables of industry unadjusted logit model were consider for the graphical illustration.

Predicted Probabilities vs Marginal Effects of LATA on Prob. of BR=0 & BR=1

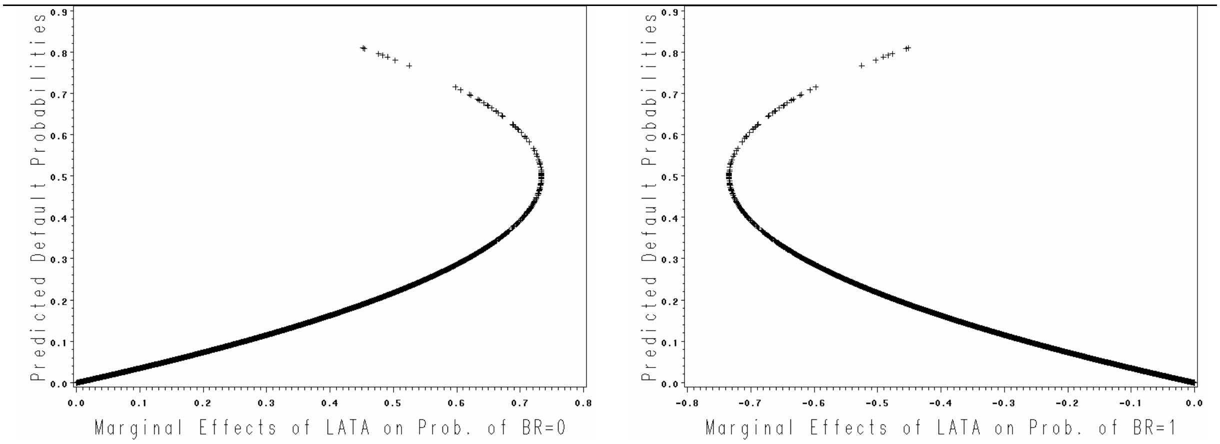
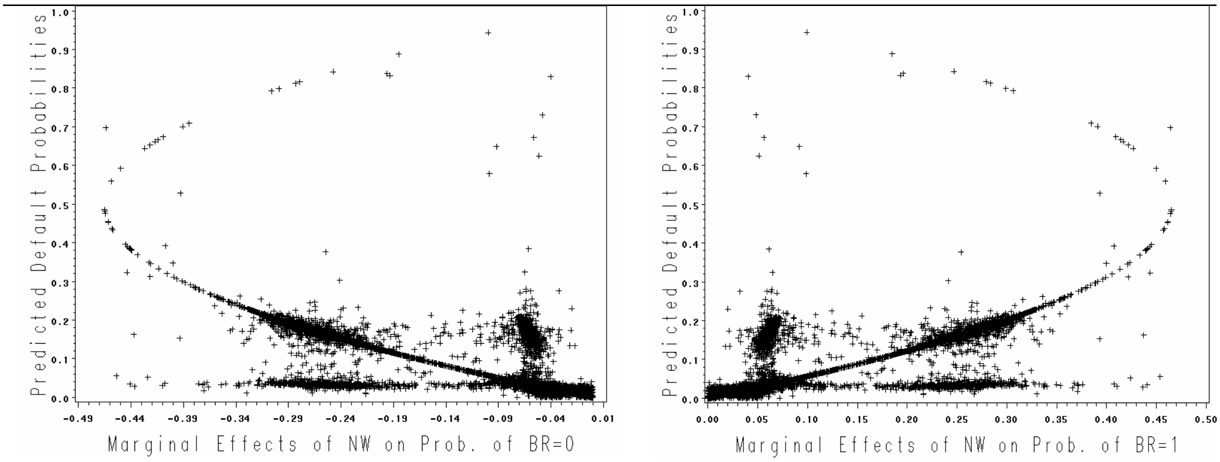


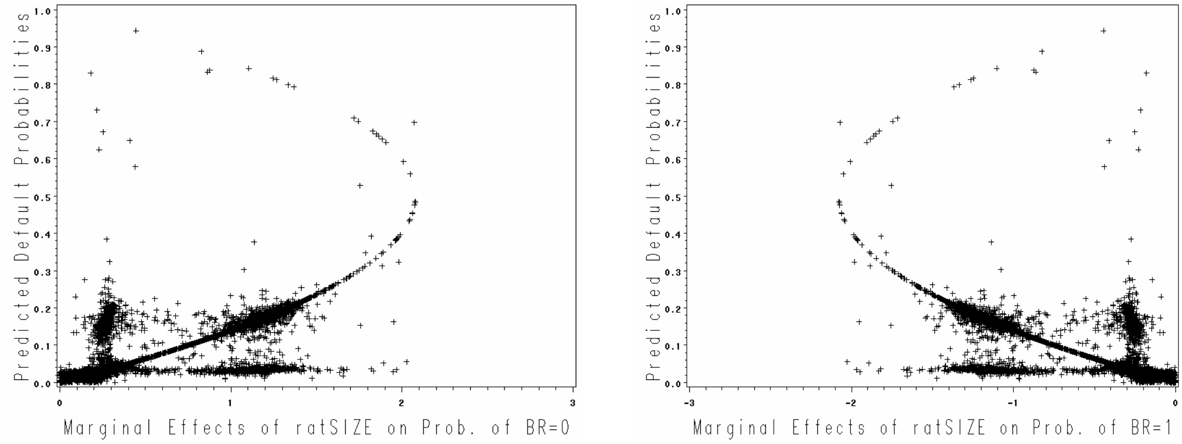
Table 17. The graphs of marginal effects of adjusted logit model[16]

Predicted Probabilities vs Marginal Effects of NW on Prob. of BR=0 & BR=1

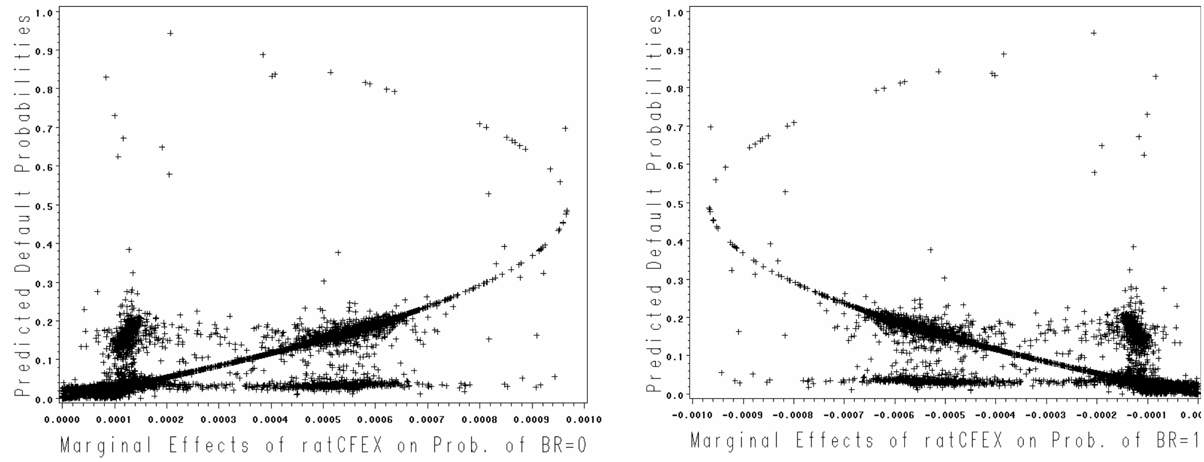


[16] The most effective variables of industry adjusted logit model were consider for the graphical illustration.

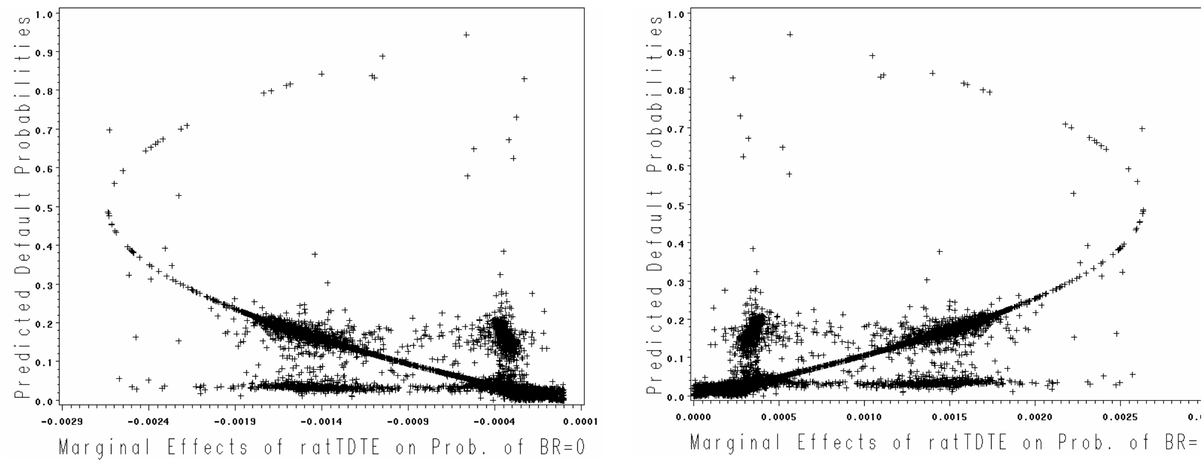
Predicted Probabilities vs Marginal Effects of ratSIZE on Prob. of BR=0 & BR=1



Predicted Probabilities vs Marginal Effects of ratCFEX on Prob. of BR=0 & BR=1



Predicted Probabilities vs Marginal Effects of ratTDTE on Prob. of BR=0 & BR=1



APPENDIX B: Davidson and MacKinnon J Test

Table 18. Predicted Y value obtained from adjusted LPM is included as an additional regressor to unadjusted LPM

BR	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
NW	0.16897	0.00673	25.12	<.0001	0.15578	0.18215
CFEX	-0.00153	0.00025	-6.19	<.0001	-0.00202	-0.00105
STA	0.00170	0.00020	8.47	<.0001	0.00130	0.00209
RMG	-0.00008	0.00002	-3.4	0.001	-0.00013	-0.00003
SIZE	-0.69019	0.03157	-21.86	<.0001	-0.75207	-0.62832
TDTE	0.00051	0.00002	21.27	<.0001	0.00046	0.00055
LATA	-0.06488	0.00214	-30.37	<.0001	-0.06907	-0.06069
TDTA	-0.00109	0.00049	-2.23	0.026	-0.00204	-0.00013
d1	-0.02501	0.00197	-12.71	<.0001	-0.02887	-0.02116
d2	-0.02851	0.01044	-2.73	0.006	-0.04898	-0.00804
d3	-0.01679	0.00760	-2.21	0.027	-0.03168	-0.00190
d4	-0.00111	0.00114	-0.98	0.33	-0.00335	0.00112
d5	-0.00568	0.00683	-0.83	0.406	-0.01907	0.00771
d6	-0.00346	0.00119	-2.9	0.004	-0.00580	-0.00112
Est. Prob(Adj)(**)	-0.27484	0.04536	-6.06	<.0001	-0.36375	-0.18593
Intercept	0.14945	0.00533	28.06	<.0001	0.13901	0.15989

(**) Estimated Probabilities of Industry Relative Model

regress BR NW CFEX STA RMG SIZE TDTE LATA TDTA d1 d2 d3 d4 d5 d6 pat

Number of obs = 177620	F(15,177604)= 503.53
Prob > F = 0.0000	R-squared = 0.0408
Adj R-squared = 0.0407	Root MSE = .18895

Table 19. Predicted Y value obtained from unadjusted LPM is included as an additional regressor to industry adjusted LPM

BR	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
NW	0.05592	0.00321	17.44	<.0001	0.04963	0.06220
ratCFEX	-0.00004	0.00001	-3.64	<.0001	-0.00007	-0.00002
ratSTA	-0.00014	0.00129	-0.11	0.913	-0.00267	0.00239
ratRMG	0.00000	0.00000	0.6	0.551	0.00000	0.00000
ratSIZE	-0.02519	0.04142	-0.61	0.543	-0.10638	0.05599
ratTDTE	0.00040	0.00006	6.81	<.0001	0.00028	0.00051
ratLATA	0.00002	0.00006	0.42	0.675	-0.00009	0.00014
ratTDTA	-0.00009	0.00047	-0.2	0.842	-0.00102	0.00083
Est. Prob(Unadj)(**)	0.60529	0.01804	33.55	<.0001	0.56992	0.64065
Intercept	0.01146	0.00079	14.49	<.0001	0.00991	0.01301

(**) Estimated Probabilities of Unadjusted Logit Model with Dummies

regress BR NW ratCFEX ratSTA ratRMG ratSIZE ratTDTE ratLATA ratTDTA phatie

Number of obs = 177620	F(9,177610) = 761.11
Prob > F = 0.0000	R-squared = 0.0371
Adj R-squared = 0.0371	Root MSE = .18931

Table 20. Predicted Y value obtained from adjusted logit model is included as an additional regressor to unadjusted logit model

BR	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
NW	2.31075	0.12833	18.01	<.0001	2.05924	2.56226
CFEX	-0.11862	0.01038	-11.42	<.0001	-0.13897	-0.09827
STA	0.03121	0.00404	7.73	<.0001	0.02330	0.03912
RMG	-0.00057	0.00034	-1.7	0.09	-0.00123	0.00009
SIZE	-21.15940	0.94763	-22.33	<.0001	-23.01671	-19.30208
TDTE	0.00718	0.00044	16.34	<.0001	0.00632	0.00804
LATA	-2.91534	0.09358	-31.15	<.0001	-3.09875	-2.73194
TDTA	-0.03356	0.00769	-4.36	<.0001	-0.04863	-0.01848
d1	-1.04497	0.08288	-12.61	<.0001	-1.20741	-0.88252
d2	-0.93519	0.38827	-2.41	0.016	-1.69618	-0.17420
d3	-0.71059	0.30787	-2.31	0.021	-1.31401	-0.10718
d4	-0.03953	0.03189	-1.24	0.215	-0.10204	0.02298
d5	-0.47886	0.29472	-1.62	0.104	-1.05649	0.09878
d6	-0.08287	0.03329	-2.49	0.013	-0.14812	-0.01763
Est. Prob(Adj)(**)	-5.73179	0.88576	-6.47	<.0001	-7.46784	-3.99573
Intercept	0.17548	0.14910	1.18	0.239	-0.11674	0.46771

(**) Estimated Probabilities of Industry Relative Model

Logistic regression	Number of obs = 177620
Log likelihood = -26364.34	LR chi2(15) = 5476.19
Prob > chi2 = 0.0000	Pseudo R2 = 0.0941

Table 21. Marginal effects of unadjusted logit model including fitted values

variable	dy/dx	Std. Err.	z	P>|z|	[95% Conf. Interval]		X
NW*	0.173461	0.0192	9.03	<.0001	0.135831	0.211091	0.06084
CFEX	-0.003148	0.00027	-11.54	<.0001	-0.003683	-0.002614	0.27998
STA	0.000828	0.00011	7.73	<.0001	0.000618	0.001038	2.42576
RMG	-1.52E-05	0.00001	-1.69	0.09	-3.30E-05	2.40E-06	-0.30802
SIZE	-0.561601	0.02469	-22.74	<.0001	-0.610001	-0.5132	0.14642
TDTE	0.000191	0.00001	16.15	<.0001	0.000167	0.000214	3.88275
LATA	-0.077377	0.00218	-35.48	<.0001	-0.081652	-0.073103	0.17855
TDTA	-8.91E-04	0.0002	-4.37	<.0001	-0.00129	-4.91E-04	0.69935
d1*	-0.018613	0.00094	-19.9	<.0001	-0.020446	-0.01678	0.0604
d2*	-0.016423	0.0042	-3.91	<.0001	-0.024659	-0.008187	0.00185
d3*	-0.013723	0.00415	-3.31	0.001	-0.02186	-0.005586	0.00352
d4*	-0.001039	0.00083	-1.25	0.211	-0.002667	0.000588	0.24292
d5*	-0.010226	0.00497	-2.06	0.040	-0.019964	-0.000488	0.00444
d6*	-0.00215	0.00084	-2.55	0.011	-0.003805	-0.000495	0.20721
Est. Prob(Adj)(**)	-0.15213	0.02352	-6.47	<.0001	-0.198225	-0.106035	0.03872

(*) dy/dx is for discrete change of dummy variable from 0 to 1
(**) Estimated Probabilities of Industry Relative Model

Marginal effects after logit
y = Pr(BR) (predict)
= .02728595

Table 22. Predicted Y value obtained from unadjusted logit model is included as an additional regressor to industry adjusted logit model

BR	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
NW	0.98967	0.05672	17.45	<.0001	0.87850	1.10084
ratCFEX	-0.00276	0.00045	-6.17	<.0001	-0.00363	-0.00188
ratSTA	-0.05030	0.03157	-1.59	0.111	-0.11218	0.01157
ratRMG	0.00004	0.00008	0.49	0.624	-0.00011	0.00019
ratSIZE	-4.67720	1.21575	-3.85	<.0001	-7.06002	-2.29438
ratTDTE	0.00454	0.00108	4.19	<.0001	0.00242	0.00666
ratLATA	0.00203	0.00140	1.45	0.146	-0.00071	0.00477
ratTDTA	-0.00524	0.00813	-0.65	0.519	-0.02117	0.01068
Est. Prob(Unadj)(**)	5.89479	0.30607	19.26	<.0001	5.29491	6.49468
Intercept	-3.64422	0.01881	-193.78	<.0001	-3.68108	-3.60736

(**) Estimated Probabilities of Unadjusted Logit Model with Dummies

logit BR NW ratCFEX ratSTA ratRMG ratSIZE ratTDTE ratLATA ratTDTA phatie

Logistic regression		Number of obs	= 177620
LR chi2(9)	= 3700.89	Prob > chi2	= 0.0000
Log likelihood	= -27251.99	Pseudo R2	= 0.0636

Table 23. Marginal effects of adjusted logit model including fitted values

variable	dy/dx	Std. Err.	z	P>|z|	[95% Conf. Interval]		X
NW*	0.04865	0.00398	12.21	<.0001	0.04084	0.05646	0.06084
ratCFEX	-0.00009	0.00001	-6.19	<.0001	-0.00012	-0.00006	4.07417
ratSTA	-0.00161	0.00101	-1.59	0.111	-0.00360	0.00037	0.60575
ratRMG	0.00000	0.00000	0.49	0.624	0.00000	0.00001	-4.33383
ratSIZE	-0.14995	0.03890	-3.85	<.0001	-0.22620	-0.07370	0.00063
ratTDTE	0.00015	0.00003	4.19	<.0001	0.00008	0.00021	1.12977
ratLATA	0.00007	0.00004	1.45	0.146	-0.00002	0.00015	12.60390
ratTDTA	-0.00017	0.00026	-0.65	0.519	-0.00068	0.00034	0.54719
Est. Prob(Unadj)(**)	0.18898	0.00996	18.97	<.0001	0.16946	0.20851	0.03872

(*) dy/dx is for discrete change of dummy variable from 0 to 1

(**) Estimated Probabilities of Unadjusted Logit Model with Dummies

Marginal effects after logit
y = Pr(BR) (predict)
= .03315902