NORTH-HOLLAND

Identifying Failing Companies: A Re-evaluation of the Logit, Probit and DA Approaches

Clive Lennox

This paper examines the causes of bankruptcy for a sample of 949 UK listed companies between 1987–1994. The most important determinants of bankruptcy are profitability, leverage, cashflow, company size, industry sector and the economic cycle. Tests for heteroskedasticity revealed that cashflow and leverage have significant non-linear effects, and taking account of these non-linearities improves the model's explanatory power. In contrast to previous studies, the paper argues that well-specified logit and probit models can identify failing companies more accurately than discriminant analysis (DA).

Keywords: Bankruptcy; Discriminant analysis; Probit; Logit

JEL classification: C25, G33

I. Introduction

There have been numerous bankruptcy studies since the pioneering research of Beaver (1966) and Altman (1968). The contribution of this paper to the literature is two-fold. First, it evaluates the merits of using well-specified probit and logit models rather than discriminant analysis (DA). The earliest bankruptcy studies used DA to identify failing companies [Altman (1968); Altman et al. (1977); Beaver (1966); Blum (1974); Deakin (1972); Elam (1975); Norton and Smith (1979); Wilcox (1973)]. More recently, researchers have used probit and logit methods, which require less restrictive assumptions [Ohlson (1980); Zmijewski (1984); Koh (1991); Hopwood et al. (1994); Platt et al. (1994)]. Despite this, previous studies have argued that, in practice, the explanatory power of probit and logit models is similar to that of DA [Press and Wilson (1978); Lo (1986);

Economics Department, Bristol University, Bristol, England.

Address correspondence to: Dr. C. Lennox, University of Bristol, Department of Economics, 8 Woodland Road, Bristol BS8 1TN, England.

Journal of Economics and Business 1999; 51:347–364

0148-6195/99/$–see front matter
PII S0148-6195(99)00009-0

Collins and Green (1982)]. However, it remains unclear whether there are potential gains from using probit and logit rather than DA, as previous probit and logit studies have not reported tests for misspecification. This is rather surprising, because Lagrange Multiplier (LM) tests for omitted variable bias and heteroskedasticity are easily calculated [Davidson and MacKinnon (1984)]. Tests for heteroskedasticity are particularly important because heteroskedasticity causes bias in both the coefficient estimates and their standard errors in probit and logit models [Yatchew and Griliches (1985)].[1] This paper appears to be the first bankruptcy study to report tests for omitted variable bias and heteroskedasticity.

Secondly, this study analyzes the effects of industry sector, company size, and the economic cycle on the probability of bankruptcy. Previous bankruptcy studies have mostly adopted a matched pairs technique for drawing samples of failing and non-failing companies. A sample of non-bankrupt companies is usually drawn by matching against the characteristics of bankrupt companies. These characteristics are generally chosen to be company size, industry sector, and year of failure. The advantage of the matching procedure is that it helps to cut the cost of data collection, as the proportion of failing companies in the population is very small.[2] One disadvantage with the matching approach is that it is not possible to investigate the effects of industry sector, company size or year of failure on the probability of bankruptcy [Jones (1987)]. In addition, the use of relatively small samples could lead to over-fitting. This study avoided these problems by collecting data on a large number of companies over an eight-year period, and by evaluating the effects of company size, industry sector, and the economic cycle on the probability of bankruptcy.

The structure of the rest of this paper is as follows. Section II outlines an economic theory of bankruptcy, and describes the tests for misspecification in probit and logit models. Section III explains the sample-selection method and describes the data. Section IV discusses the estimation results, and Section V compares predictive accuracy of the DA, probit and logit models. Section VI offers conclusions.

II. Theoretical Developments and Statistical Methodology

Previous empirical research has found that a company is more likely to fail if it is unprofitable, highly leveraged, and suffers cashflow difficulties [Altman (1986)]. Myers (1977) has outlined a theoretical model which helps to explain these findings. The model predicts that investors will choose to liquidate if the company's liquidation value exceeds its going-concern value.[3] If profitable companies have higher going concern value, one should find that profitable companies are less likely to go bankrupt than unprofitable companies.

[1] In contrast, heteroskedasticity in regression models does not affect the consistency of coefficient estimates.

[2] In DA and logit estimation, a sample selection rule which results in the sample proportion being different from the population proportion of failing companies merely biases the constant term [Lachenbruch (1975); Anderson (1972)]. In contrast, probit estimation requires the likelihood function to be weighted so that the sample proportion of bankrupt companies is approximately equal to the population proportion—otherwise all coefficient estimates are biased.

[3] Variations on this theme have been proposed following this research. In Bulow and Shoven (1978), liquidation occurs if it leads to a positive gain for a coalition of the firm's claimants. Thus a company might be liquidated even when its going-concern value exceeds its liquidation value. In the investment irreversibility literature, there may be an option value to delaying bankruptcy if there are sunk costs and uncertainty. In this case, bankruptcy may not occur even if the expected returns from liquidating exceed the expected returns from not liquidating [Dixit and Pindyck (1994); Hubbard (1994)].

If contracting were complete, a company's owners would contract with the manager to liquidate when the company's liquidation value exceeds its going-concern value. When contracting is incomplete, financial structure can substitute for contracts [Aghion and Bolton (1992)]. In particular, managers may issue debt so that the company enters liquidation when it defaults on debt servicing. In this way, managers are able to signal that they are willing to act in the investors' interests [Hart (1995)]. This implies that a company is more likely to enter bankruptcy when leverage is high.

Bankruptcy is usually triggered by default on debt servicing, and this is less likely to occur if the company has access to internal or external finance. A company with healthy cashflow has relatively easy access to internal finance, and so it is less likely to go bankrupt than a company with cashflow problems. Large companies are less likely to encounter credit constraints in the market for external finance because of reputation effects. Therefore, company size may be an important determinant of bankruptcy. Finally, the economic cycle and industry sector may determine a company's access to finance.

Although early bankruptcy studies used DA to identify failing companies, its suitability rests on two assumptions. First, the explanatory variables are assumed to have a multivariate normal distribution. Nevertheless, it is well known that the variables typically used in bankruptcy studies are not normally distributed [Eisenbeis (1977); McLeay (1986)]. Secondly, the samples of failing and non-failing companies are assumed to be drawn at random from their respective populations. However, the matched pairs technique violates this assumption. For example, matching on the basis of company size leads to too many small companies in the non-bankrupt sample, because small companies are more likely to go bankrupt than large companies. Similarly, matching on the basis of industry will lead to too many companies from recession-hit industries in the non-bankrupt sample. In addition to these two assumptions, most DA studies have used a linear classification rule, which is only optimal if the restriction of equal group covariance matrices is satisfied. Yet, the evidence indicates that this restriction does not usually hold in bankruptcy studies [Hamer (1983)].

Such problems with DA have led researchers to use probit and logit models. These can be written as follows:

$$Y_{it}^{*} = \beta_1' X_{it} + u_{it} \tag{1}$$

where

$$\begin{array}{lll} Y_{it} = 1 & \text{if} & Y_{it}^{*} \geq 0 \\ Y_{it} = 0 & & \text{otherwise.} \end{array}$$

The log-likelihood function is:

$$\ln(L) = \Sigma_i Y_{it} F(-\beta_1' X_{it}) + \Sigma_i (1 - Y_{it})(1 - F(-\beta_1' X_{it})), \tag{2}$$

where $F(\cdot)$ is the distribution function, the functional form of which depends on the assumptions made about u_{it}. In the logit model, the cumulative distribution of u_{it} is the logistic; in the homoskedastic probit model, the u_{it} are IN(0, σ^2). In the heteroskedastic probit model, the u_{it} are normally distributed with non-constant variance. Consider, for example, the case where the variance of u_{it} is a function of X_{it}:

$$Y_{it}^{*} = \beta_1' X_{it} + u_{it} \qquad u_{it} \sim \text{IN}(0, \exp(2\beta_2' X_{it})). \tag{3}$$

Clearly, the heteroskedastic probit model collapses to the homoskedastic probit model when $\beta_2 = 0$. The likelihood function for this heteroskedastic probit model is:

$$\ln(L) = \Sigma_i Y_{it}\Phi(\beta_1' X_{it} \exp(-\beta_2' X_{it})) + \Sigma_i(1 - Y_{it})(1 - \Phi(\beta_1' X_{it} \exp(-\beta_2' X_{it}))), \tag{4}$$

where $\Phi(\cdot)$ is the cumulative normal distribution.

Davidson and MacKinnon (1984) showed how to test for omitted variable bias and heteroskedasticity in logit and probit models. The procedure involves rescaling the residuals and explanatory variables from the maximum likelihood estimation. The scaled residuals ($R_{it}(\beta_1; Y_{it})$) and scaled explanatory variables ($X_{it}(\beta_1)$) are defined as follows:

$$R_{it}(\beta_1; Y_{it}) \equiv Y_{it}[F(-\beta_1' X_{it})/F(\beta_1' X_{it})]^{1/2} + (Y_{it} - 1)[F(-\beta_1' X_{it})/F(\beta_1' X_{it})]^{1/2}$$

$$X_{it}(\beta_1) \equiv [F(\beta_1' X_{it})\, F(-\beta_1' X_{it})]^{-1/2} f(\beta_1' X_{it})\, X_{it}, \tag{5}$$

where $f(\cdot)$ is the density function.

One, then, generates an artificial regression by regressing $R_{it}(\beta_1; Y_{it})$ on $X_{it}(\beta_1)$. The explained sum of squares from the artificial regression is an LM statistic which is used to test for omitted variable bias.

One can use a similar method to test for heteroskedasticity in the probit model by testing the restriction that $\beta_2 = 0$ in equations (3) and (4). A scaled polynomial variable can be defined as follows:

$$X_{it}(\beta_2) \equiv [F(\beta_1' X_{it})\, F(-\beta_1' X_{it})]^{-1/2} f(\beta_1' X_{it})(-X_{it}\beta_1 X_{it}). \tag{6}$$

The test for heteroskedasticity involves regressing $R_{it}(\beta_1; Y_{it})$ on $X_{it}(\beta_1)$ and $X_{it}(\beta_2)$. The explained sum of squares from the artificial regression is an LM statistic which is used to test for heteroskedasticity.

III. Data

The data for this study consist of 949 listed UK companies in the United Kingdom between 1987–1994. The sample was determined on the basis of availability of data from Datastream. Unfortunately, a complete panel of data was not available for each company in the sample; this means that the total number of companies observed in each year is lower than 949, as shown in the first row of Table 1.[4]

According to the Stock Exchange Yearbook, there were 160 listed companies that failed between 1987–1994. A company is deemed to have failed if it entered liquidation, receivership or administration as defined by the Yearbook. Financial information for 90 of these companies was found in Datastream. Thus, the frequency of corporate failure in the sample is 1.4% per annum.[5]

The dependent variable, $FAILS_{it}$, takes a value of 1 if two requirements are met: first, that the company entered bankruptcy and, secondly, that the company filed its final annual

[4] Two points should be made about the presence of missing observations in the data. First, the majority of companies (653) have a complete panel of data covering all eight-years. Secondly, when the sample proportion of companies differs from the population proportion of companies, the logit model has consistent coefficient estimates for all variables except the constant [Anderson (1972)]. The next section shows that the results from the probit and logit models are very similar. Thus, there is strong reason to believe that there are no sample selection problems.

[5] This is approximately equal to the population frequency of failure [Morris (1997)].

Table 1. Failing and Non-failing Companies Over Time

	1987	1988	1989	1990	1991	1992	1993	1994	Total
Total number of companies	706	773	813	836	843	832	823	790	6416
Number of failing companies[a]	0	7	24	26	23	6	3	1	90
F_t	4	5	6	36	46	41	15	7	160
CBI_t	29	18	−5	−23	−17	8	30	13	·

[a] A company is defined as failing if $FAILS_{it} = 1$.
$FAILS_{it} = 1$ if company i issued its last annual report in year t, prior to entering bankruptcy; = 0, otherwise.
F_t = total number of UK listed companies entering bankruptcy in year t.
CBI_t = proportion of respondents stating that business conditions had improved over the past four months minus the proportion of respondents stating that business conditions had deteriorated.

report prior to entering bankruptcy. For companies which did not go bankrupt, and for the earlier reports of failing companies, $FAILS_{it}$ takes a value of 0.[6] The average length of time between the final annual report of a failing company and its entry into bankruptcy was 14 months.[7]

The number of failures in the population of listed companies, F_t, is included in the bankruptcy model to capture cyclical effects. Rows 2 and 3 of Table 1 highlight the time lag which arose because companies did not immediately enter bankruptcy when they issued their final reports. The data also indicate that companies were more likely to fail in the recession of 1990–1992 than in the boom of the late 1980s and the economic recovery after 1992. To capture relative changes in business confidence, a variable (CBI_t) was constructed from the Confederation of British Industry (CBI) Quarterly Industrial Trends Surveys, in which the CBI published the results of a questionnaire asking the following question: "Are you more, or less, optimistic than you were four months ago about the general business situation in your industry?"[8] CBI_t is equal to the proportion of respondents answering "more" minus the proportion answering "less." Therefore, F_t captures the effect of *current* economic conditions while CBI_t is a leading indicator of future *changes* in economic conditions.

As the probability of bankruptcy is likely to vary across industry sectors, it is important to include industry dummies. The standard industrial classification (SIC) codes were obtained for each company's main activities from Extel. Data were collected at the one-digit level, with the exception of classification 8500, which refers to companies owning and dealing in real estate. This exception was made because of the volatile nature of the housing market over the period. Table 2 shows the number of companies operating in each industry sector.

Table 3 shows that there are a large number of companies with more than one main activity, reflecting the diversification of many listed companies.

Variables capturing company size, profitability, leverage, and cashflow were collected from Datastream and are defined in Table 4.[9] The number of employees (EMP_{it}) is used as a measure of company size, and the SIC data are used to create industry dummies (Dj). The debtor-turnover ($DBTN_{it}$), gross cashflow (GCF_{it}) and cash ratio ($CASHRAT_{it}$)

[6] None of the failing companies issued reports after filing for bankruptcy.
[7] This is consistent with the findings of Citron and Taffler (1992).
[8] For each year, the results from the April questionnaire were used.
[9] These financial ratios are defined by Datastream to aid analysts in evaluating companies' financials. They can be directly downloaded from Datastream without the need for any complex calculations, by using the codes given in Table 4.

Table 2. Main Activities

SIC Code	Industry Sector	Number of Companies
0	Agriculture	19
1	Energy and water	31
2	Extraction of minerals and ores	137
3	Metal goods	275
4	Other manufacturing	274
5	Construction	100
6	Distribution, hotels and catering	411
7	Transport and communication	47
8[a]	Banking, finance and insurance	255
8500	Owning and dealing in real estate	151
9	Other services	84
Total		1784[b]

[a] Classification 8 refers to all industries with a first digit which begins with an 8, but which does not belong to classification 8500.
[b] The total of 1784 is greater than the number of companies in the sample (949) because some companies operated in more than one sector.

variables are used as measures of cashflow. $DBTN_{it}$ captures the effect of debtor repayment on cashflow; if $DBTN_{it}$ is low, the company may have experienced problems in receiving payment for past sales. GCF_{it} is a measure of profit-generated cashflow. $CASHRAT_{it}$ captures the ability of the company to meet its short-term liabilities through cash reserves.[10]

Table 5 presents descriptive statistics for these variables. Compared to non-failing companies, failing companies are typically small, have poor cashflow and profitability, and are highly leveraged. All the variables failed normality tests for skewness and kurtosis [D'Agostino et al. (1990)], implying that DA is unlikely to provide satisfactory results.[11]

Table 3. Company Diversification

Number of Industry Sectors	Number of Companies
1	455
2	325
3	146
4	37
5	12
6	1
Total	949

[10] Previous bankruptcy studies have used very simple ratios; for example, the Zmijewski (1984) model comprised the return on assets (net income/total assets), the current ratio (current assets/current liabilities) and gearing (debt/total assets). These ratios are not readily available from Datastream and so are not included in the model. Other financial ratios provided by Datastream are income gearing (Datastream code, 732), creditor turnover (728), stock turnover (724), working capital ratio (741), and the quick ratio (742). These variables were found to have insignificant and/or non-constant effects on bankruptcy and were therefore omitted from the model.

[11] The number of employees (EMP_{it}), debtor-turnover ($DBTN_{it}$), and cash ratio ($CASHRAT_{it}$) variables are inevitably non-normal because they have a lower bound of zero. In addition, the use of industry dummies clearly violates the normality assumption. In this respect, it should be noted that DA may perform quite well when all

Table 4. Company Size, Industry Dummies, Cashflow, Leverage and Profitability Variables

Variable	Definition	Datastream Code	Interpretation
EMP_{it}	Total number of employees	219	Company size
Dj	$= 1$ if company i operates in industry j; $= 0$ otherwise	·	Industry effects
$DBTN_{it}$	$\frac{(\text{Total sales}) \times 100}{\text{Total debtors}}$	726	Debtor turnover ratio
$CASHRAT_{it}$	$\frac{(\text{Total cash}) \times 100}{\text{Current liabilities}}$	743	Cash ratio
GCF_{it}	$\frac{(\text{Profits earned for ordinary shareholders} + \text{depreciation} + \text{tax equalisation}) \times 100}{\text{Capital employed} + \text{current liabilities} - \text{intangibles}}$	735	Gross cashflow
$CAPG_{it}$	$\frac{(\text{Preference capital} + \text{subordinated debt} + \text{loan capital} + \text{short-term borrowings}) \times 100}{\text{Capital employed} + \text{short-term borrowing} - \text{intangibles}}$	731	Capital gearing
ROC_{it}	$\frac{(\text{Total interest charged} + \text{pre-tax profit}) \times 100}{\text{Capital employed} + \text{short-term borrowing} - \text{intangibles}}$	707	Return on capital

Table 5. Descriptive Statistics

	Non-Failing Companies (Number of Observations = 6326)		Failing Companies (Number of Observations = 90)	
Variables	Mean	Standard Deviation	Mean	Standard Deviation
EMP_{it}	6399.26	20160.45	1176.50	2244.24
$DBTN_{it}$	827.59	1646.87	580.52	532.46
$CASHRAT_{it}$	33.98	90.50	10.46	18.31
GCF_{it}	9.22	10.03	−0.87	14.78
$CAPG_{it}$	3062.15	8699.26	11008.68	28528.21
ROC_{it}	1836.18	6032.03	415.81	4506.76

These results were also confirmed using Shapiro and Wilk (1965) and Shapiro and Francia (1972) tests.

IV. Empirical Results

This section presents the results from estimating the following bankruptcy model:

$$FAILS^*_{it} = \beta_1 X_{it} + u_{it} \qquad u_{it} \sim IID(0, \exp(2\beta_2' X_{it})), \tag{7}$$

where

$$\begin{aligned} FAILS_{it} &= 1 \quad \text{if} \quad FAILS^*_{it} \geq 0 \\ FAILS_{it} &= 0 \quad \text{otherwise.} \end{aligned}$$

The X_{it} variables used to predict bankruptcy are those shown in Tables 1, 2 and 4. Table 6 summarizes the key characteristics and assumptions of the seven probit, logit and DA models that were estimated.

Model 1 is a linear probit model which imposes the restriction of homoskedasticity ($\beta_2 = 0$). The assumption of a linear functional form and homoskedasticity is the same as that imposed by previous probit and logit studies. However, LM tests revealed that the homoskedasticity restriction can be rejected in Model 1. Models 2–4 were estimated to show that the heteroskedasticity of Model 1 is caused by incorrectly assuming a linear

Table 6. Summary of Models 1–7

Model	Assumes Normality and Equal Group Covariance Matrices?	Functional Form	Assumes Homoskedasticity?	Distribution of u_{it}
1. Linear homoskedastic probit	No	Linear	Yes	Normal
2. Linear heteroskedastic probit	No	Linear	No	Normal
3. Non-linear heteroskedastic probit	No	Non-linear	No	Normal
4. Non-linear homoskedastic probit	No	Non-linear	Yes	Normal
5. Non-linear homoskedastic logit	No	Non-linear	Yes	Logistic
6. Discriminant analysis	Yes	Non-linear	No	Not applicable

explanatory variables are dichotomous; however, DA's performance is less likely to be satisfactory when variables have a lower bound [Amemiya and Powell (1983)].

functional form. In particular, the gross cashflow (GCF_{it}) and leverage ($CAPG_{it}$) variables have non-linear effects. Model 3 is the most general model, as it allows for non-linearities and heteroskedasticity. Once the non-linear effects of GCF_{it} and $CAPG_{it}$ are taken into account, the null hypothesis of homoskedasticity can no longer be rejected. Therefore, the homoskedasticity restriction can be imposed, as in Models 4 and 5. Models 4 and 5 enable a comparison of the logit and probit results to investigate whether the choice between the logistic and normal distributions is important.

Table 7 presents the results for the probit and logit models. The LM1 test statistics indicate no problems of omitted variable bias; however, Model 1's LM2 statistic shows that the null hypothesis of homoskedasticity is rejected in the linear model. The significance of the gross cashflow (GCF_{it}) and leverage ($CAPG_{it}$) variables in the heteroskedastic part of Model 2 confirms that these variables were the cause of the heteroskedasticity in Model 1.

The LM test for heteroskedasticity indicates that a non-constant variance for the residuals can be caused by imposing an incorrect functional form.[12] To investigate whether leverage and cashflow have non-linear effects, polynomial terms in $CAPG_{it}$ and GCF_{it} were added in Model 3. Model 3 is the least restrictive, because it allows for both polynomial terms in $CAPG_{it}$ and GCF_{it} and for heteroskedasticity. These polynomial terms have highly significant effects, and including them means that one can no longer reject the null hypothesis of homoskedasticity ($\beta_2 = 0$). The LM2 test statistics for Models 4 and 5 confirm that once the non-linear effects of cashflow and leverage are taken into account, the null hypothesis of homoskedasticity can no longer be rejected. The results for Models 4 and 5 are very similar, indicating that there is little to choose between the probit and logit approaches.

The results show that bankruptcy is more likely when the economy moves from boom to recession. The negative coefficient on the number of failing companies in the population (F_t) implies that a company is less likely to go bankrupt in the future if the economy is currently in a recession. The negative coefficient on the CBI indicator (CBI_t) implies that an improvement in business confidence is correlated with a fall in the probability of bankruptcy (an increase in CBI_t implies that business confidence is improving). The signs on F_t and CBI_t show that a company is less (more) likely to fail over the next 12–18 months if the economy is currently in a recession (boom) and business conditions are expected to improve (worsen).

Another important determinant of bankruptcy is company size. The coefficient on the number of employees (EMP_{it}) is negative, showing that corporate failure is more likely if a company is small.[13] The industry dummies are also important; a company was more likely to enter bankruptcy if it operated in the construction ($D5_i$) or financial services ($D8_i$) sectors. This reflects the fact that high interest rates badly affected the building industry and, in contrast to the 1979–1981 recession, the 1990–1992 recession badly hit the financial services sector.

[12] Equation (6) shows that the heteroskedasticity may be captured by including polynomial terms rather than by allowing the error term to have a non-constant variance.

[13] It might be argued that company size can affect the probability of bankruptcy because large companies tend to be more highly diversified and are less vulnerable to sector-specific shocks. To investigate this possibility, a variable capturing the number of industry sectors in which companies had main activities was included in the model. No evidence was found to suggest that the degree of diversification helped improve predictive accuracy.

Table 7. Probit and Logit Models of Bankruptcy 1987–1994 (z statistics in parentheses)

$$FAILS^*_{it} = \beta_1' X_{it} + u_{it} \qquad u_{it} \sim \mathrm{IID}(0,\ \exp(2\beta_2' X_{it}))$$

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5
F_t	−0.011	−0.019	−0.016	−0.015	−0.034
	(−2.880)	(−3.585)	(−3.397)	(−3.682)	(−3.506)
CBI_t	−0.026	−0.036	−0.030	−0.030	−0.065
	(−6.697)	(−5.912)	(−5.287)	(−6.775)	(−6.494)
EMP_{it}	−0.379e-04	−0.582e-04	−0.457e-04	−0.455e-04	−1.013e-04
	(−2.193)	(−2.204)	(−2.201)	(−2.342)	(−2.175)
$D1_i$	0.451	0.617	0.545	0.540	1.199
	(1.744)	(1.974)	(1.990)	(2.127)	(2.111)
$D2_i$	−0.239	−0.183	−0.144	−0.145	−0.376
	(−1.420)	(−0.860)	(−0.807)	(−0.812)	(−0.886)
$D4_i$	0.084	0.138	0.108	0.108	0.299
	(0.746)	(0.922)	(0.856)	(0.865)	(1.803)
$D5_i$	0.416	0.530	0.457	0.455	0.959
	(3.116)	(2.930)	(2.888)	(3.103)	(3.077)
$D8_i$	0.399	0.462	0.394	0.392	0.840
	(3.800)	(3.183)	(3.064)	(3.339)	(3.282)
$D8500_i$	−0.054	−0.119	−0.186	−0.185	−0.362
	(−0.408)	(−0.660)	(−1.239)	(−1.260)	(−1.104)
$DBTN_{it}$	−0.205e-03	−0.255e-03	−0.227e-03	−0.226e-03	−0.499e-03
	(−1.876)	(−1.606)	(−1.789)	(−1.804)	(−1.718)
$CASHRAT_{it}$	−0.777e-02	−0.4592	−0.355e-02	−0.353e-02	−0.955e-02
	(−3.296)	(−1.808)	(−1.602)	(−1.627)	(−1.769)
GCF_{it}	−0.177e-01	0.761e-02	−0.122e-01	−0.117e-01	−0.281e-01
	(−5.272)	(0.643)	(−0.896)	(−1.250)	(−1.413)
GCF^2_{it}	·	·	−0.283e-03	−0.279e-03	−0.657e-03
	·	·	(−1.838)	(−2.020)	(−2.089)
$CAPG_{it}$	0.116e-04	0.513e-04	0.254e-03	0.254e-03	0.539e-03
	(3.842)	(2.054)	(6.789)	(6.986)	(6.807)
$CAPG^2_{it}$	·	·	−0.848e-08	−0.849e-08	−1.800e-08
	·	·	(−4.622)	(−4.662)	(−4.650)
$CAPG^3_{it}$	·	·	0.594e-13	0.597e-13	1.270e-13
	·	·	(3.432)	(3.884)	(3.912)
$CAPG^4_{it}$	·	·	−0.115e-18	−0.116e-18	−0.245e-18
	·	·	(−2.365)	(−3.171)	(−3.218)
ROC_{it}	0.960e-06	−0.632e-04	−0.639e-04	−0.638e-04	−1.218e-04
	(0.154)	(−2.307)	(−1.776)	(−1.821)	(−1.767)
Constant	−1.759	−2.267	−2.533	−2.529	−4.952
	(−12.591)	(−9.247)	(−10.584)	(−11.324)	(−9.789)
Heteroskedasticity					
GCF_{it}	·	−0.992e-02	0.300e-03	·	·
	·	(−2.793)	(0.050)	·	·
$CAPG_{it}$	·	0.501e-04	0.571e-06	·	·
	·	(3.823)	(0.032)	·	·
LM1[a]	1.050	·	·	0.025	1.145
95% critical value	6.571	·	·	9.390	9.390
LM2[b]	49.979	·	·	9.663	14.168
95% critical value	16.928	·	·	23.300	23.300
Pseudo R^2	0.2169	·	·	0.3305	0.3249

[a] LM1: LM test for omitted variable bias.
[b] LM2: LM test for heteroskedasticity and incorrect functional form.
6416 observations.

A company is also more likely to go bankrupt when it is suffering cashflow difficulties. The negative coefficient on the debtor-turnover ratio ($DBTN_{it}$) implies that a company is more likely to fail if it is having problems receiving payment from debtors. Similarly, the negative coefficient on the cash ratio ($CASHRAT_{it}$) indicates that companies are more likely to fail if cash reserves are low. The negative coefficient on gross cashflow (GCF_{it}) shows that a company is more likely to go bankrupt when profit-generated cashflow is low.

Models 3–5 show that GCF_{it} has non-linear effects on $FAILS^*_{it}$. Differentiating $FAILS^*_{it}$ with respect to GCF_{it} shows how a change in gross cashflow affects the probability of bankruptcy. For values of GCF_{it} that lie within two standard deviations of the mean of GCF_{it}, an increase in cashflow always reduces the probability of bankruptcy. However, the size of this effect depends on whether the company is suffering from cashflow problems. For low values of GCF_{it}, an increase in cashflow has a small negative effect on the probability of bankruptcy; for high values of GCF_{it}, an increase in cashflow has a large negative effect. In other words, companies are less likely to go bankrupt as cashflow improves, and this effect is larger for companies with relatively healthy cashflow.

Company indebtedness is also an important determinant of bankruptcy. The positive coefficient on capital gearing ($CAPG_{it}$) in Model 1 shows that a company is more likely to fail when leverage is high. Models 3–6 show that there is a non-linear relationship between leverage and bankruptcy. Differentiating $FAILS^*_{it}$ with respect to $CAPG_{it}$ shows how a change in leverage affects the probability of bankruptcy. For values of $CAPG_{it}$ that lie within two standard deviations of the mean of $CAPG_{it}$, an increase in leverage always increases the probability of bankruptcy. For low values of $CAPG_{it}$, an increase in leverage has a large positive effect on the probability of bankruptcy. For high values of $CAPG_{it}$, an increase in leverage has a small positive effect on the probability of bankruptcy. Thus, an increase in leverage raises the probability of bankruptcy, but this effect diminishes as leverage increases.[14]

The negative coefficient on the return on capital (ROC_{it}) means that a company is more likely to fail when profitability is low. It is noteworthy that the profitability effect is absent in Model 1, which incorrectly assumes a linear functional form, and is only detected when the non-linear effects of leverage and cashflow are taken into account. This emphasizes the importance of testing for heteroskedasticity.

To summarize, a company is most likely to go bankrupt when it is unprofitable, highly leveraged, and has cashflow problems.[15] Although these results are similar to those of previous studies, the finding of non-linear effects for cashflow and leverage is new. These non-linear effects were only discovered by testing for heteroskedasticity. This reinforces the importance of testing for misspecification in probit and logit models. In contrast to studies using the matching approach, company size, industry sector and the economic cycle have also been shown to have important effects on the probability of bankruptcy.

---

[14] A relatively small proportion of the observations for GCF_{it} (6.9%) and $CAPG_{it}$ (0.3%) took negative values. For these observations, the interpretation of the coefficients on the second and fourth powers is less clear than for positive observations. Various checks were carried out to ensure that this was not a major problem. Omitting the small number of negative $CAPG_{it}$ observations had no effect on the sign or significance of the coefficients on $CAPG^2_{it}$ or $CAPG^4_{it}$. Similarly, omitting the negative GCF_{it} observations had no effect on the sign or significance of GCF^2_{it} (however, omitting negative GCF_{it} observations involved dropping 34 failing observations and resulted in a less robust model). Retaining negative observations but omitting the $CAPG^2_{it}$, $CAPG^4_{it}$ and GCF^2_{it} variables was not found to solve the heteroskedasticity problem, and so it seemed sensible to retain these polynomial terms.

[15] It should be noted that lagged variables were not found to be important predictors of bankruptcy.

Table 8. Results for DA (Model 6)[a]

Panel A: Canonical Discriminant Functions

Eigenvalue[b]	Canonical Correlation[c]	Wilks' Lambda[d]	χ^2 (18)	Prob $> \chi^2$
0.048	0.214	0.954	300.555	0.000

Panel B: Pooled Within-Groups Correlations Between Variables and Discriminant Function (variables ordered by size of correlation within function)

Variable	Correlation	Variable	Correlation	Variable	Correlation
GCF_{it}	0.535	$CAPG_{it}^2$	−0.242	$D2_i$	0.117
$CAPG_{it}$	−0.460	$D5_i$	−0.228	$D8500_i$	−0.089
CBI_t	0.445	F_t	−0.221	$DBTN_{it}$	0.081
$CAPG_{it}^3$	−0.272	$CASHRAT_{it}$	0.141	$D1_i$	−0.049
$CAPG_{it}^4$	−0.270	EMP_{it}	0.140	$D4_i$	0.035
$D8_i$	−0.265	ROC_{it}	0.127	GCF_{it}^2	−0.022

Panel C: Test of Equality of Group Covariance Matrices Using Box's M (the ranks and natural logarithms of determinants are those of the group covariance matrices)

Group	Rank	Log Determinant		
$FAILS_{it} = 0$	18	291.02		
$FAILS_{it} = 1$	18	260.04		
Pooled within-groups covariance matrix	18	293.06		
Box's M	Approximate F(171)	Significance	Prob $> F$	Number of Observations
15480	86.013	72733.4	0.000	6416

[a] DA calculates β_i ($i = 1, \ldots, n$), giving a Z score which can be used to classify observations into one of the two samples:

$$Z = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n.$$

These weights were estimated so that they resulted in the best separation between the samples, given prior probabilities and costs of misclassification. In other words, the weights were chosen so as to maximize the (between-groups sum of squares/within-groups sum of squares) ratio.

[b] Eigenvalue ≡ (between group sum of squares/within group sum of squares).

[c] Canonical correlation ≡ (between group sum of squares/total sum of squares)$^{1/2}$.

In the two group case, the canonical correlation is the correlation coefficient between the discriminant score and the group variable.

[d] Wilks' lambda ≡ (within group sum of squares/total sum of squares).

Wilks' lambda captures the proportion of the total variance in the discriminant scores not explained by the differences between the groups.

Having developed a model of bankruptcy which appears to be well-specified, the analysis turns to consider the results from DA, so as to evaluate whether there are gains in predictive accuracy from using probit and logit rather than DA. Table 8 reports the results using DA (Model 6), where cashflow and leverage are allowed to have non-linear effects as in Models 4 and 5.[16] Panel A shows that the eigenvalue and canonical correlation statistics are rather low, suggesting that the DA model is not exceptionally good at discriminating between the samples of failing and non-failing companies. The significance of the Wilks' lambda statistic shows that the null hypothesis that the mean of the discriminant scores is the same in the groups of failing and non-failing companies can

[16] The estimation results for the DA model are shown separately from those for the probit and logit models, as the reported test statistics are somewhat different.

be rejected. Panel B assesses the contribution of each variable to the discriminant function. Selecting variables on the basis of each variable's contribution to the discriminant function resulted in models with low explanatory power. Moreover, the size of the correlation between each variable and the discriminant function did not reflect the levels of significance in the probit model and logit models. This suggests that the level of correlation with the discriminant function is a poor criterion for choosing which variables should be included in the model.

Box's M tests the null hypothesis of equal covariance matrices assuming multivariate normality. Panel C shows that the null hypothesis of equality in the group covariance matrices can be strongly rejected. This may be due either to a failure of multivariate normality or because the group covariance matrices are not equal. In either case, the DA approach, together with a linear classification rule, is not an appropriate way in which to estimate the bankruptcy model.

V. Explanatory Power

In evaluating the explanatory power of the bankruptcy models, it is helpful to define two types of prediction error. A type I error occurs when a company fails but is predicted to survive; a type II error occurs when a company survives but is predicted to fail. Clearly, the type I and type II error rates depend on the number of companies predicted to fail. The higher (lower) the number of companies predicted to go bankrupt, the smaller (larger) is the type I error rate and the larger (smaller) is the type II error rate. The number of predicted bankruptcies depends on the cut-off probabilities chosen for the models. For example, if the cut-off is equal to 0.1, a company for which the expected probability of bankruptcy exceeds 10% is predicted to go bankrupt, whereas a company for which the expected probability of bankruptcy is less than 10% is predicted to survive. One can therefore increase the number of companies predicted to fail by reducing the cut-off probability. Table 9 presents the type I and type II error rates for the probit, logit and DA models for different cut-offs; the lowest type I and type II error rates are highlighted in bold.

In comparing these results to previous studies, it is important to note that reported type I and type II error rates depend critically on the sample selection criterion. Studies which sample approximately equal numbers of failing and non-failing companies typically have much smaller error rates compared to those where the sample frequency of failure is close to the population frequency, as in this study [see Zmijewski (1984, Table 1)]. The type I and II error rates reported in this paper are comparable to those of previous studies with sample frequency failure rates of less than 3% [White and Turnbull (1975); Zmijewski (1983)].

The results indicate that the non-linear models (3, 4 and 5) predicted better than the linear models (1 and 2). This shows that a failure to take account of the non-linear effects of cashflow and leverage results in models with worse explanatory power. Models 3–5 also had lower type I and type II error rates than Model 6, suggesting that well-specified probit and logit models are superior to the DA model. For all cut-off probabilities, the type I error rates of Models 3–5 were significantly less than those of Models 1 and 6 (at confidence levels of at least 80%); the differences between the type II error rates were not statistically significant. As type I errors tend to be much more costly than type II errors,

Table 9. Type I and Type II Error Rates for Models Estimated Between 1987–1994

No. of Predicted Failures	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)
80	78.89	0.96	80.00	0.98	**73.33**[c,f]	**0.89**	**73.33**[c,f]	**0.89**	**73.33**[c,f]	**0.89**	78.89	0.96
100	76.67	1.25	76.67	1.25	**68.89**[c,f]	**1.14**	**68.89**[c,f]	**1.14**	70.00[c,f]	1.15	76.67	1.25
120	73.33	1.52	72.22	1.50	**64.44**[c,f]	**1.39**	**64.44**[c,f]	**1.39**	**64.44**[c,f]	**1.39**	72.22	1.50
140	71.11	1.80	68.89	1.77	**61.11**[c,f]	**1.66**	**61.11**[c,f]	**1.66**	**61.11**[c,f]	**1.66**	70.00	1.79
180	66.67	2.37	63.33	2.32	**55.56**[b,e]	**2.21**	**55.56**[b,e]	**2.21**	**55.56**[b,e]	**2.21**	65.56	2.36
220	63.33	2.96	56.67	2.86	52.22[b,e]	2.80	52.22[b,e]	2.80	**51.11**[a,d]	**2.78**	63.33	2.96
260	58.89	3.53	52.22	3.43	48.89[c,f]	3.38	48.89[c,f]	3.38	**47.78**[c,f]	**3.37**	56.67	3.49
300	54.44	4.09	51.11	4.05	42.22[b,e]	3.92	42.22[b,e]	3.92	**41.11**[b,e]	**3.90**	52.22	4.06
500	43.33	7.10	36.67	7.00	32.22[b,e]	6.94	**31.11**[b,e]	**6.92**	32.22[b,e]	6.94	42.22	7.08
1000	30.00	14.81	14.44	14.59	**13.33**[a,d]	**14.57**	**13.33**[a,d]	**14.57**	**13.33**[a,d]	**14.57**	25.56	14.75
2000	13.33	30.38	7.78	30.30	**6.67**[b,e]	**30.29**	**6.67**[b,e]	**30.29**	**6.67**[b,e]	**30.29**	13.33	30.38
4000	1.11	61.82	**0**[c,f]	**61.81**	**0**[c,f]	**61.81**	**0**[c,f]	**61.81**	**0**[c,f]	**61.81**	1.11	61.82

Total number of observations = 6416. Number of failures = 90.
[a] Significantly lower than corresponding error rate for Model 6 at 95% confidence level (one-tailed test).
[b] Significantly lower than corresponding error rate for Model 6 at 90% confidence level (one-tailed test).
[c] Significantly lower than corresponding error rate for Model 6 at 80% confidence level (one-tailed test).
[d] Significantly lower than corresponding error rate for Model 1 at 95% confidence level (one-tailed test).
[e] Significantly lower than corresponding error rate for Model 1 at 90% confidence level (one-tailed test).
[f] Significantly lower than corresponding error rate for Model 1 at 80% confidence level (one-tailed test).

this suggests that there may be real gains to users when models are tested for misspecification problems.[17]

It is worth noting that the type I and type II error rates of Models 1 and 2 are very similar to those of Model 6. In other words, there is virtually no difference between the explanatory power of misspecified probit models and the DA model. This reflects the findings of previous studies which argued that there is little gain from using probit or logit approaches rather than DA [Press and Wilson (1978); Lo (1986); Collins and Green (1982)]; however, these studies did not test the probit and logit models for omitted variable bias or heteroskedasticity and did not allow for non-linear effects.

Table 10 shows how Model 5's accuracy was affected by alternative bankruptcy horizons. The benchmark case is where the model predicts one period ahead (the error rates for column 1 correspond to those reported in Table 9). The second (third and fourth) column(s) show Model 5's type I and type II error rates for failures predicted to occur within the next two (three and four) reporting periods.

Consistent with previous research, the bankruptcy model showed a decline in accuracy for more distant bankruptcy horizons. However, the deterioration in accuracy was not very great; for example, a comparison of Tables 9 and 10 shows that Model 5 had lower type I and type II error rates, when the bankruptcy horizon was *two* periods, than DA where the horizon was just *one* period.

[17] Altman (1977) estimated the relative costs of type I and type II errors for commercial bank loans as being 7:1.

Table 10. Type I and Type II Error Rates of Model 5 for Alternative Bankruptcy Horizons

No. of Predicted Failures	Number of Reporting Periods Prior to Failure							
	One Period[a]		Two Periods[b]		Three Periods[c]		Four Periods[d]	
	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)
80	73.33	0.89	79.21	0.69	83.53	0.62	86.23	0.62
100	70.00	1.15	78.84	0.91	81.18	0.84	84.26	0.81
120	64.44	1.39	73.03	1.15	78.82	1.07	82.30	1.08
140	61.11	1.66	70.79	1.41	76.86	1.31	80.66	1.33
180	55.56	2.21	65.73	1.91	71.76	1.75	76.07	1.75
220	51.11	2.78	62.36	2.45	69.02	2.29	73.11	2.26
260	47.78	3.37	58.43	2.98	66.27	2.82	70.49	2.78
300	41.11	3.90	52.81	3.46	61.96	3.29	66.56	3.24
500	32.22	6.94	41.57	6.35	51.37	6.10	57.05	6.04
1000	13.33	14.57	25.28	13.90	36.08	13.59	42.30	13.48
2000	6.67	30.29	16.29	29.67	24.31	29.33	30.49	29.26
4000	0	61.81	4.49	61.40	7.45	61.09	9.84	60.96

[a] There were 90 observations where companies issued their final reports prior to entering bankruptcy (i.e., where $FAILS_{it} = 1$).
[b] There were 178 observations where companies issued their final or penultimate reports prior to entering bankruptcy (i.e., where $FAILS_{it-1} = FAILS_{it} = 1$).
[c] There were 255 observations where companies issued their last three reports prior to entering bankruptcy (i.e., where $FAILS_{it-2} = FAILS_{it-1} = FAILS_{it} = 1$).
[d] There were 305 observations where companies issued their last four reports prior to entering bankruptcy (i.e., where $FAILS_{it-3} = FAILS_{it-2} = FAILS_{it-1} = FAILS_{it} = 1$).
For each cut-off, the type I error rate tended to rise and the type II error rate fell as the bankruptcy horizon increased. This means that a statistical test for each cut-off, similar to those reported in Tables 9 and 11, would not be meaningful. This is because, as the bankruptcy horizon increased, the number of failing observations rose (from 90 to 305), whereas the number of predicted failures was constant for each cut-off.

To evaluate out-of-sample accuracy, Models 1–6 were re-estimated using data for 1987–1990, and were used to identify failing companies between 1991–1994.[18] Table 11 shows the type I and type II error rates for these re-estimated models during the hold-out period (1991–1994). A comparison of Tables 9 and 11 show that there is little difference between the accuracy of models in the estimation and hold-out samples. This suggests that the models do not suffer from over-fitting.

As in Table 9, the non-linear models (3, 4 and 5) generally performed better than the DA model (Model 6). The difference in type I error rates was statistically significant (at the 95% confidence level) for the 20–40 observations with the highest predicted bankruptcy probabilities. This suggests that the difference in accuracy is greatest when the number of predicted failures is close to the number of actual failures (33 in the hold-out sample). The difference in accuracy between the linear and non-linear probit models was not statistically significant in the hold-out sample. The linear models (1 and 2) generally performed better than DA, but their superior performance was less significant statistically. This is consistent with previous studies which imposed linear specifications and did not find significant differences between the probit, logit and DA approaches.

[18] The Lachenbruch approach does not accurately measure explanatory power when coefficients are non-constant over the sample period. Therefore, a hold-out sample was used in preference to the Lachenbruch method.

Table 11. Type I and Type II Error Rates in Hold-out Sample (1991–1994) for Models Estimated Between 1987–1990

No. of Predicted Failures	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)
20	87.88[a]	0.49	96.97	0.58	90.91[b]	0.52	**75.76**[a]	**0.39**[c]	90.91[b]	0.52	100.00	0.61
30	78.79[b]	0.71	78.79[b]	0.71	75.76[a]	0.68	**69.70**[a]	**0.61**	72.73[a]	0.65	93.94	0.86
40	72.73[b]	0.95	72.73[b]	0.95	**69.70**[a]	**0.92**	**69.70**[a]	**0.92**	**69.70**[a]	**0.92**	90.91	1.14
60	**63.64**	**1.47**	**63.64**	**1.47**	69.70	1.54	**63.64**	**1.47**	66.67	1.51	72.73	1.57
80	**57.58**	**2.03**	60.61	2.06	63.64	2.09	63.64	2.09	63.64	2.09	66.67	2.12
100	54.54	2.61	**51.52**	**2.58**	57.58	2.64	57.58	2.64	60.61	2.67	60.61	2.67
120	54.54	3.23	**48.48**	**3.16**	51.52	3.20	51.52	3.20	51.52	3.20	57.58	3.26
140	45.45	3.75	48.48	3.78	45.45	3.75	**42.42**	**3.72**	45.45	3.75	57.58	3.87
160	42.42	4.33	48.48	4.39	39.39	4.30	**36.36**	**4.27**	39.39	4.30	51.52	4.42
180	42.42	4.95	48.48	5.01	36.36	4.88	36.36	4.88	**33.33**	**4.85**	42.42	4.95
200	42.42	5.56	48.48	5.62	36.36	5.50	36.36	5.50	**33.33**	**5.47**	36.36	5.50
240	36.36	6.73	39.39	6.76	33.33	6.70	33.33	6.70	**30.30**	**6.67**	33.33	6.70
300	33.33	8.54	36.36	8.57	**27.27**	**8.48**	**27.27**	**8.48**	30.30	8.51	30.30	8.51
500	24.24	14.59	27.27	14.62	24.24	14.59	24.24	14.59	24.24	14.59	**21.21**	**14.56**
1000	18.18	29.89	15.15	29.86	**12.12**	**29.83**	**12.12**	**29.83**	**12.12**	**29.83**	**12.12**	**29.83**

Number of observations in hold-out sample = 3288. Number of failures in hold-out sample = 33.
[a] Significantly lower than corresponding error rate for Model 6 at 95% confidence level (one-tailed test).
[b] Significantly lower than corresponding error rate for Model 6 at 90% confidence level (one-tailed test).
[c] Significantly lower than corresponding error rate for Model 6 at 80% confidence level (one-tailed test).

VI. Conclusion

Consistent with previous research, this study has shown that profitability, leverage and cashflow have important effects on the probability of bankruptcy. Whereas previous studies did not report tests for misspecification, heteroskedasticity tests used in this paper revealed that cashflow and leverage have non-linear effects on the probability of bankruptcy; incorporating these effects into the model improved its predictive accuracy. In contrast to previous studies, this paper avoided the matching approach. Therefore, it was able to evaluate the effects of company size, industry sector, and the economic cycle on the probability of bankruptcy, and avoided the over-fitting problems which can arise in small samples.

In contrast to previous studies, probit and logit models were found to perform better than DA. In hold-out samples, this difference in predictive accuracy was found to be greatest for realistic cut-off probabilities (where the numbers of predicted and actual failures were similar), and for type I errors (which, in practice, are much more costly than type II errors). Moreover, this superiority over DA was greatest for well-specified non-linear probit and logit models. This may explain why previous studies which did not report tests for misspecification found that probit, logit and DA models were very similar in terms of predictive accuracy.

References

Aghion, P., and Bolton, P. July 1992. An "incomplete contracts" approach to financial contracting. *Review of Economic Studies* 59(3):473–494.

Altman, E. I. Sept. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23: 589–609.

Altman, E. I., Haldeman, R. G., and Narayanan, P. June 1977. Zeta analysis. *Journal of Banking and Finance*: 29–54.

Altman, E. October 1977. Some estimates of the cost of lending errors for commercial banks. *Journal of Commercial Bank Lending*: 51–58.

Altman, E. 1986. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. In *Issues and Readings in Managerial Finance* (R. E. Johnson, ed.). New York: Dryden Press, pp. 53–75

Amemiya, T., and Powell, J. 1983. A comparison of the logit model and normal discriminant analysis when the independent variables are binary. In *Studies in Econometrics, Time Series, and Multivariate Statistics* (S. Karlin, T. Amemiya and L. A. Goodman, eds.). New York: Academic Press, pp. 1–24

Anderson, J. A. 1972. Separate sample logistic discrimination. *Biometrika* 59(1):19–35.

Beaver, W. H. 1966. Financial ratios as predictors of bankruptcy. *Journal of Accounting Research* (Supplement):71–102.

Blum, M. Spring 1974. Failing company discriminant analysis. *Journal of Accounting Research* 12(1):1–25.

Bulow, J., and Shoven, J. Autumn 1978. The bankruptcy decision. *The Bell Journal of Economics* 9(2):437–456.

Citron, D., and Taffler, R. J. Autumn 1992. The audit report under going concern uncertainties: An empirical analysis. *Accounting and Business Research* 22:337–345.

Collins, R., and Green, R. 1982. Statistical methods for bankruptcy prediction. *Journal of Economics and Business* 34(4):349–354.

D'Agostino, R. B., Balanger, A., and D'Agostino, Jr., R. B. Nov. 1990. A suggestion for using powerful and informative tests of normality. *The American Statistician* 44:316–321.

Davidson, R., and MacKinnon, J. G. July 1984. Convenient specification tests for logit and probit models. *Journal of Econometrics* 25(3):241–262.

Deakin, E. B. Spring 1972. A discriminant analysis of predictors of business failure. *Journal of Accounting Research* 10:167–179.

Dixit, A., and Pindyck, R. 1994. *Investment Under Uncertainty*. Princeton: Princeton University Press.

Eisenbeis, R. A. June 1977. Pitfalls in the application of discriminant analysis in business, finance and economics. *Journal of Finance* 22(3):875–890.

Elam, R. Jan. 1975. The effect of lease data on the predictive ability of financial ratios. *The Accounting Review* 5(1): 25–43.

Hamer, M. 1983. Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets. *Journal of Accounting and Public Policy* 2:289–307.

Hart, O. 1995. *Firms, Contracts and Financial Structure*. Oxford: Clarendon Lectures in Economics, OUP.

Hopwood, W. S., McKeown, J. C., and Mutchler, J. F. 1994. A Re-examination of auditor versus model accuracy within the context of the going-concern opinion decision. *Contemporary Accounting Research* 10(2):409–431.

Hubbard, G. December 1994. Investment under uncertainty: Keeping one's options open. *Journal of Economic Literature* 32(4):1816–1831.

Jones, F. L. 1987. Current techniques in bankruptcy prediction. *Journal of Accounting Literature* 6:131–164.

Koh, H. C. Autumn 1991. Model predictions and auditor assessments of going concern status. *Accounting and Business Research* 21(84):331–338.

Lachenbruch, P. A. 1975. *Discriminant Analysis*. Hafner Press. New York

Lo, A. W. March 1986. Logit versus discriminant analysis: A specification test and application to corporate bankruptcies. *Journal of Econometrics* 31(2):151–178.

McLeay, S. Summer 1986. Student's *t* and the distribution of financial ratios. *Journal of Business Finance and Accounting* 13(2): 209–222.

Morris, R. 1997. *Early Warning Indicators of Corporate Failure: A Critical Review of Previous Research and Further Empirical Evidence*. Ashgate Publishing in association with the Institute of Chartered Accountants in England and Wales. Aldershot, England.

Myers, S. November 1977. Determinants of corporate borrowing. *Journal of Financial Economics* 5(2):147–175.

Norton, C. L., and Smith, R. E. Jan. 1979. A comparison of general price level and historical cost financial statements in the prediction of bankruptcy. *The Accounting Review* 54:72–87.

Ohlson, J. S. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18(1):109–131.

Platt, H., Platt, M., and Pedersen, J. June 1994. Bankruptcy discrimination with real variables. *Journal of Business, Finance and Accounting* 21(4):491–510.

Press, J., and Wilson, S. Dec. 1978. Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association* 73(364): 699–705.

Shapiro, S. S., and Francia, R. S. 1972. An approximate analysis of variance test for normality. *Journal of the American Statistical Association* 67(337):215–216.

Shapiro, S. S., and Wilk, M. B. 1965. An analysis of variance test for normality. *Biometrika* 52(3):591–611.

White, R. W., and Turnbull, H. 1975. The probability of bankruptcy for American Industrial Firms. Working paper. London Graduate School of Business.

Wilcox, J. W. 1973. A prediction of business failure using accounting data. *Journal of Accounting Research* (Supplement) 11:163–179.

Yatchew, A., and Griliches, Z. February 1985. Specification error in probit models. *The Review of Economics and Statistics* 66(1):134–139.

Zmijewski, M. E. 1983. Essays on corporate bankruptcy. Ph.D. dissertation. State University of New York at Buffalo.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* (Supplement) 22:59–86.