Part I: Distress Prediction Models and Some Applications

1 Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models*

Edward I. Altman

Background

This paper discusses two of the venerable models for assessing the distress of industrial corporations. These are the so-called Z-Score model (Altman, 1968) and ZETA® (1977) credit-risk model. Both models are still being used by practitioners throughout the world. The latter is a proprietary model for subscribers to ZETA Services, Inc. (Hoboken, NJ).

The purpose of this summary is two-fold. First, those unique characteristics of business failures are examined in order to specify and quantify the variables which are effective indicators and predictors of corporate distress. By doing so, I hope to highlight the analytic as well as the practical value inherent in the use of financial ratios. Specifically, a set of financial and economic ratios will be analyzed in a corporate distress predication context using a multiple discriminant statistical methodology. Through this exercise, I will explore not only the quantifiable characteristics of potential bankrupts but also the utility of a much-maligned technique of financial analysis: ratio analysis. Although the models that we will discuss were developed in the late 1960s and mid-1970s, I will extend our tests and findings to include application to firms not traded publicly, to non-manufacturing entities, and also refer to a new bond-rating equivalent model for emerging markets corporate bonds. The latter utilizes a version of the Z-Score model called Z″. This paper also updates the predictive tests on defaults and bankruptcies through the year 1999.

As I first wrote in 1968, and it seems even truer in the late 1990s, academicians seem to be moving toward the elimination of ratio analysis as an analytical technique in

* This paper is adapted and updated from E. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, September 1968; and E. Altman, R. Haldeman and P. Narayanan, "Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations," *Journal of Banking & Finance*, 1, 1977. R. Haldeman is President of ZETA® Services Inc. and Paul Narayanan is a private financial consultant.

assessing the performance of the business enterprise. Theorists downgrade arbitrary rules of thumb (such as company ratio comparisons) widely used by practitioners. Since attacks on the relevance on ratio analysis emanate from many esteemed members of the scholarly world, does this mean that ratio analysis is limited to the world of "nuts and bolts?" Or, has the significance of such an approach been unattractively garbed and therefore unfairly handicapped? Can we bridge the gap, rather than sever the link, between traditional ratio analysis and the more rigorous statistical techniques which have become popular among academicians in recent years? Along with our primary interest, corporate bankruptcy, I am also concerned with an assessment of ratio analysis as an analytical technique.

It should be pointed out that the basic research for much of the material in this paper was performed in 1967 and that several subsequent studies have commented on the Z-Score model and its effectiveness, including an adaptation in 1995 for credit analysis of emerging market corporates. And this author co-developed a "second generation" model, ZETA® (ZETA®, 1977).

Traditional Ratio Analysis

The detection of company operating and financial difficulties is a subject which has been particularly amenable to analysis with financial ratios. Prior to the development of quantitative measures of company performance, agencies were established to supply a qualitative type of information assessing the creditworthiness of particular merchants. (For instance, the forerunner of the well-known Dun & Bradstreet, Inc. was organized in 1849 in Cincinnati, Ohio, to provide independent credit investigations). Formal aggregate studies concerned with portents of business failure were evident in the 1930s.

One of the classic works in the area of ratio analysis and bankruptcy classification was performed by Beaver (1967). In a real sense, his univariate analysis of a number of bankruptcy predictors set the stage for the multivariate attempts, by this author and others, which followed. Beaver found that a number of indicators could discriminate between matched samples of failed and nonfailed firms for as long as five years prior to failure. He questioned the use of multivariate analysis, although a discussant recommended attempting this procedure. The Z-Score model did just that. A subsequent study by Deakin (1972) utilized the same 14 variables that Beaver analyzed, but he applied them within a series of multivariate discriminant models.

The aforementioned studies imply a definite potential of ratios as predictors of bankruptcy. In general, ratios measuring profitability, liquidity, and solvency prevailed as the most significant indicators. The order of their importance is not clear since almost every study cited a different ratio as being the most effective indication of impending problems.

Although these works established certain important generalizations regarding the performance and trends of particular measurements, the adaptation of the results for assessing bankruptcy potential of firms, both theoretically and practically, is questionable. In almost every case, the methodology was essentially univariate in nature and emphasis was placed on individual signals of impending problems. Ratio analysis presented in this fashion is susceptible to faulty interpretation and is potentially confusing. For instance, a firm with a poor profitability and/or solvency record may be regarded as a potential bankrupt. However, because of its above average liquidity, the situation may not be considered serious. The potential ambiguity as to the relative performance of several firms is clearly evident. The crux of the shortcomings inherent in any univariate analysis lies therein. An appropriate extension of the previously cited studies, therefore, is to build on their findings and to combine several measures into a meaningful predictive model. In so doing, the highlights of ratio analysis as an analytical technique will be emphasized rather than downgraded. The questions are:

- Which ratios are most important in detecting bankruptcy potential?
- What weights should be attached to those selected ratios?
- How should the weights be objectively established?

Discriminant Analysis

After careful consideration of the nature of the problem and of the purpose of this analysis, I chose multiple discriminant analysis (MDA) as the appropriate statistical technique. Although not as popular as regression analysis, MDA has been utilized in a variety of disciplines since its first application in the 1930s. During those earlier years, MDA was used mainly in the biological and behavioral sciences. In recent years, this technique has become increasingly popular in the practical business world as well as in academia. Altman et al. (1981) discusses discriminant analysis in-depth and reviews several financial application areas.

MDA is a statistical technique used to classify an observation into one of several *a priori* groupings dependent on the observation's individual characteristics. It is used primarily to classify and/or make predictions in problems where the dependent variable appears in qualitative form, for example, male or female, bankrupt or non-bankrupt. Therefore, the first step is to establish explicit group classifications. The number of original groups can be two or more. Some analysts refer to discriminant analysis as "multiple" only when the number of groups exceeds two. We prefer that the multiple concepts refer to the multivariate nature of the analysis.

After the groups are established, data are collected for the objects in the groups; MDA in its most simple form attempts to derive a linear combination of these characteristics which "best" discriminates between the groups. If a particular object, for instance, a corporation, has characteristics (financial ratios) which can be quantified for all of the companies in the analysis, the MDA determines a set of discriminant coefficients. When these coefficients are applied to the actual ratios, a basis for classification into one of the mutually exclusive groupings exists. The MDA technique has the advantage of considering an entire profile of characteristics common to the relevant firms, as well as the interaction of these properties. A univariate study, on the other hand, can only consider the measurements used for group assignments one at a time.

Another advantage of MDA is the reduction of the analyst's space dimensionally, that is, from the number of different independent variables to G-1 dimension(s), where G equals the number of original *a priori* groups. This analysis is concerned with two groups, consisting of bankrupt and non-bankrupt firms. Therefore, the analysis is transformed into its simplest form: one dimension. The discriminant function, of the form

$$Z = V_1X_1 + V_2X_2 + \cdots + V_nX_n$$

transforms the individual variable values to a single discriminant score, or z value, which is then used to classify the object where

$V_1, V_2, \ldots, V_n$ = discriminant coefficients, and
$X_1, X_2, \ldots, X_n$ = independent variables

The MDA computes the discriminant coefficient; V_i while the independent variables X_i are the actual values.

When utilizing a comprehensive list of financial ratios in assessing a firm's bankruptcy potential, there is reason to believe that some of the measurements will have a high degree of correlation or collinearity with each other. While this aspect is not serious in discriminant analysis, it usually motivates careful selection of the predictive variables (ratios). It also has the advantage of potentially yielding a model with a relatively small number of selected measurements which convey a great deal of information. This information might very well indicate differences among groups, but whether or not these differences are significant and meaningful is a more important aspect of the analysis.

Perhaps the primary advantage of MDA in dealing with classification problems is the potential of analyzing the entire variable profile of the object simultaneously rather than sequentially examining its individual characteristics. Just as linear and integer programming have improved on traditional techniques in capital budgeting, the MDA approach to traditional ratio analysis has the potential to reformulate the problem correctly. Specifically, combinations of ratios can be analyzed together so as to remove possible ambiguities and misclassifications observed in earlier traditional ratio studies.

As we will see, the Z-Score model is a linear analysis in that five measures are objectively weighted and summed to arrive at an overall score that then becomes the basis for classification of firms into one of the *a priori* groupings (distressed and nondistressed).

Development of the Z-Score Model

The initial sample is composed of 66 corporations with 33 firms in each of the two groups. The bankrupt (distressed) group (Group 1) are manufacturers that filed a bankruptcy petition under Chapter X of the National Bankruptcy Act from 1946 through 1965. A 20-year period is not the best choice since average ratios do shift over time. Ideally, we would prefer to examine a list of ratios in time period t and to make predictions about other firms in the following period ($t + 1$). Unfortunately, it was not possible to do this because of data limitations. Recognizing that this group is not completely homogeneous (due to industry and size differences), I attempted to make a careful selection of non-bankrupt (nondistressed) firms. Group 2 consists of a paired sample of manufacturing firms chosen on a stratified random basis. The firms are stratified by industry and by size, with the asset size range restricted to between $1 million and $25 million. The mean asset size of the firms in Group 2 ($9.6 million) was slightly greater than that of Group 1, but matching exact asset size of the two groups seemed unnecessary. Firms in Group 2 were still in existence at the time of the analysis. Also, the data collected are from the same years as those compiled for the bankrupt firms. For the initial sample test, the data are derived from financial statements dated one annual reporting period prior to bankruptcy. The data were derived from *Moody's Industrial Manuals* and also from selected annual reports. The average lead-time of the financial statements was approximately seven and one-half months.

An important issue is to determine the asset-size group to be sampled. The decision to eliminate both the small firms (under $1 million in total assets) and the very large companies from the initial sample essentially is due to the asset range of the firms in Group 1. In addition, the incidence of bankruptcy in the large-asset-size firm was quite rare prior to 1966. This changed, starting in 1970, with the appearance of several very large bankruptcies, e.g., Penn-Central R.R. Large industrial bankruptcies also increased in appearance, since 1978. In all, there have been at least 100 Chapter 11 bankruptcies with over $1 billion in liabilities since 1978 (the year of the existing Bankruptcy Code's enactment).

A frequent argument is that financial ratios, by their very nature, have the effect of deflating statistics by size, and that therefore a good deal of the size effect is eliminated. The Z-Score model, discussed below, appears to be sufficiently robust to accommodate large firms. The ZETA model did include larger-sized distressed firms and is unquestionably relevant to both small and large firms.

Variable Selection

After the initial groups are defined and firms selected, balance sheet and income statement data are collected. Because of the large number of variables found to be significant indicators of corporate problems in past studies, a list of 22 potentially helpful variables (ratios) was complied for evaluation. The variables are classified into five standard ratio categories: liquidity, profitability, leverage, solvency, and activity. The ratios are chosen on the basis of their popularity in the literature and their potential relevancy to the study, and there are a few "new" ratios in this analysis. The Beaver study (1967) concluded that the cash flow to debt ratio was the best single ratio predictor. This ratio was not considered in my 1968 study because of the lack of consistent and precise depreciation and cash flow data. The results obtained, however, were still superior to the results Beaver attained with his single best ratio. Cash flow measures were included in the ZETA® model tests (see later discussion).

From the original list of 22 variables, five are selected as doing the best overall job together in the prediction of corporate bankruptcy. This profile did not contain all of the most significant variable measured independently. This would not necessarily improve on the univariate, traditional analysis described earlier. The contribution of the entire profile is evaluated and, since this process is essentially iterative, there is no claim regarding the optimality of the resulting discriminant function. The function, however, does the best job among the alternatives which include numerous computer runs analyzing different ratio profiles.

To arrive at a final profile of variables, the following procedures are utilized:

1. Observation of the statistical significance of various alternative functions, including determination of the relative contributions of each independent variable
2. Evaluation of intercorrelations among the relevant variables
3. Observation of the predictive accuracy of the various profiles
4. Judgment of the analyst

The final discriminant function is as follows:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

where

X_1 = working capital/total assets
X_2 = retained earnings/total assets
X_3 = earnings before interest and taxes/total assets
X_4 = market value of equity/book value of total liabilities
X_5 = sales/total assets
Z = overall index

Note that the model does not contain a constant (y-intercept) term. This is due to the particular software utilized and, as a result, the relevant cut-off score between the two groups is not zero. Other software programs, like SAS and SPSS, have a constant term, which standardizes the cut-off score at zero if the sample sizes of the two groups are equal.

X_1, WORKING CAPITAL/TOTAL ASSETS (WC/TA)

The working capital/total assets ratio, frequently found in studies of corporate problems, is a measure of the net liquid assets of the firm relative to the total capitalization. Working capital is defined as the difference between current assets and current liabilities. Liquidity and size characteristics are explicitly considered. Ordinarily, a firm experiencing consistent operating losses will have shrinking current assets in relation to total assets. Of the three liquidity ratios evaluated, this one proved to be the most valuable. Two other liquidity ratios tested were the current ratio and the quick ratio. There were found to be less helpful and subject to perverse trends for some failing firms.

X_2, RETAINED EARNINGS/TOTAL ASSETS (RE/TA)

Retained earnings is the account which reports the total amount of reinvested earnings and/or losses of a firm over its entire life. The account is also referred to as earned surplus. It should be noted that the retained earnings account is subject to "manipulation" via corporate quasi-reorganizations and stock dividend declarations. While these occurrences are not evident in this study, it is conceivable that a bias would be created by a substantial reorganization or stock dividend, and appropriate readjustments should be made to the accounts.

This measure of cumulative profitability over time is what I referred to earlier as a "new" ratio. The age of a firm is implicitly considered in this ratio. For example, a relatively young firm will probably show a low RE/TA ratio because it has not had time to build up its cumulative profits. Therefore, it may be argued that the young firm is somewhat discriminated against in this analysis, and its chance of being classified as bankrupt is relatively higher than that of another older firm, *ceteris paribus*. But, this is precisely the situation in the real world. The incidence of failure is much higher in a firm's earlier years. In 1993, approximately 50 percent of all firms that failed did so in the first five years of their existence (Dun & Bradstreet, 1994).

In addition, the RE/TA ratio measures the leverage of a firm. Those firms with high RE, relative to TA, have financed their assets through retention of profits and have not utilized as much debt.

X_3, EARNINGS BEFORE INTEREST AND TAXES/TOTAL ASSETS (EBIT/TA)

This ratio is a measure of the true productivity of the firm's assets, independent of any tax or leverage factors. Since a firm's ultimate existence is based on the earning power of its assets, this ratio appears to be particularly appropriate for studies dealing with corporate failure. Furthermore, insolvency in a bankrupt sense occurs when the total liabilities exceed a fair valuation of the firm's assets with value determined by the earning power of the assets. As we will show, this ratio continually outperforms other profitability measures, including cash flow.

X_4, MARKET VALUE OF EQUITY/BOOK VALUE OF TOTAL LIABILITIES (MVE/TL)

Equity is measured by the combined market value of all shares of stock, preferred and common, while liabilities include both current and long term. The measure shows how much the firm's assets can decline in value (measured by market value of equity plus debt) before the liabilities exceed the assets and the firm becomes insolvent. For example, a company with a market value of its equity of $1,000 and debt of $500 could experience a two-thirds drop in asset value before insolvency. However, the same firm with $250 equity will be insolvent if assets drop only one-third in value. This ratio adds a market value dimension which most other failure studies did not consider. The reciprocal of X_4 is a slightly modified version of one of the variables used effectively by Fisher (1959) in a study of corporate bond yield–spread differentials. It also appears to be a more effective predictor of bankruptcy than a similar, more commonly used ratio: net worth/total debt (book values). At a later point, we will substitute the book value of net worth for the market value to derive a discriminant function for privately held firms (Z') and for non-manufacturers (Z'').

More recent models, such as the KMV[1] approach, are essentially based on the market value of equity and its volatility. The equity market value serves as a proxy for the firm's asset values.

X_5, SALES/TOTAL ASSETS (S/TA)

The capital–turnover ratio is a standard financial ratio illustrating the sales generating ability of the firm's assets. It is one measure of management's capacity in dealing with competitive conditions. This final ratio is quite important because it is the least significant ratio on an individual basis. In fact, based on the univariate statistical significance test, it would not have appeared at all. However, because of its unique relationship to other variables in the model, the sales/total assets ratio ranks second in its contribution to the overall discriminating ability of the model. Still, there is a wide variation among industries in asset turnover, and we will specify an alternative model (Z''), without X_5 at a later point.

[1] KMV is a San Francisco based credit risk modeling company.

A Clarification

The reader is cautioned to utilize the model in the appropriate manner. Due to the original computer format arrangement, variables X_1 through X_4 must be calculated as absolute percentage values. For instance, the firm whose net working capital to total assets (X_1) is 10 percent should be included as 10.0 percent and not 0.10. Only variable X_5 (sales to total assets) should be expressed in a different manner: that is, a S/TA ratio of 200 percent should be included as 2.0. The practical analyst may have been concerned by the extremely high relative discriminant coefficient of X_5. This seeming irregularity is due to the format of the different variables. Table 1.1 illustrates the proper specification and form for each of the five independent variables.

Over the years, many individuals have found that a more convenient specification of the model is of the form:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5$$

Using this formula, one inserts the more commonly written percentage, for example, 0.10 for 10 percent, for the first four variables (X_1–X_4) and rounds the last coefficient off to equal 1.0 (from 0.99). The last variable continues to be written in terms of number of times. The scores for individual firms and related group classification and cutoff scores remain identical. We merely point this out and note that we have utilized this format in some practical application, for example, Altman and LaFleur (1981).

Table 1.1 Variable means and test significance

Variable	*Bankrupt group mean*[a]	*Non-bankrupt group mean*[a]	*F ratio*[a]
X_1	−6.1%	41.4%	32.50*
X_2	−62.6%	35.5%	58.86*
X_3	−31.8%	15.4%	26.56*
X_4	40.1%	247.7%	33.26*
X_5	1.5X	1.9X	2.84

[a] N = *33.*
$F_{1.60}(0.001) = 12.00$; $F_{1.60}(0.01) = 7.00$; $F_{1.60}(0.05) = 4.00$.
* *Significant at the 0.001 level.*

Variable Tests

A test to determine the overall discriminating power of the model is the F-value which is the ratio of the sums-of-squares between-groups to the within-groups sums-of-squares. When this ratio is maximized, it has the effect of spreading the means (centroids) of the groups apart and, simultaneously, reducing dispersion of the individual points (firm Z-values) about their respective group means. Logically, this test (commonly called the F-test) is appropriate because the objective of the MDA is to identify and utilize those variables which best discriminate between groups and which are most similar within groups.

The group means of the original two-group sample are:

- For Group 1: –0.29, F = 20.7
- For Group 2: +5.02, $F_{4n}(0.01) = 3.84$

The significance test therefore rejects the null hypothesis that the observations come from the same population.

Variable means measured at one financial statement prior to bankruptcy and the resulting F-statistics a shown in table 1.1. Variables X_1 through X_4 are all significant at the 0.001 level, indicating extremely significant differences in these variables among groups. Variable X_5 does not show a significant difference among groups and the reason for its inclusion in the variable profile is not apparent as yet. On a strictly univariate level, all of the ratios indicate higher values for the non-bankrupt firms. Also, all of the discriminant coefficients display positive signs, which is what one would expect. Therefore, the greater a firm's distress potential, the lower its discriminant score. It is clear that four of the five variables display significant differences between groups, but the importance of MDA is its ability to separate groups using multivariate measures.

Once the values of the discriminant coefficients are estimated, it is possible to calculate discriminant scores for each observation in the samples, or any firm, and to assign the observations to one of the groups based on this score. The essence of the procedure is to compare the profile of an individual firm with that of the alternative groupings. The comparisons are measured by a χ^2 value and assignments are made based on the relative proximity of the firms' score to the various group centroids.

Initial Sample (Group 1)

The initial sample of 33 firms in each of the two groups is examined using data compiled one financial statement prior to distress. Since the discriminant coefficients and the group distributions are derived from this sample, a high degree of successful classification is expected. This should occur because the firms are classified using a discriminant function which, in fact, is based upon the individual measurements of these same firms. The classification matrix for the original sample is shown in table 1.2.

The model is extremely accurate in classifying 95 percent of the total sample correctly. The Type I error proved to be only 6 percent while the Type II error was even

Table 1.2 Classification results, original sample

						Predicted	
	Number correct	*% correct*	*% error*	*n*	*Actual*	*Group 1*	*Group 2*
					Group 1	31	2
					Group 2	1	32
Type I	31	94	6	33			
Type II	32	97	3	33			
Total	63	95	5	66			

lower at 3 percent. The results, therefore, are encouraging, but the obvious upward bias should be kept in mind, and further validation techniques are appropriate.

Results Two Statements Prior to Bankruptcy

The second test observes the discriminating ability of the model for firms using data compiled two statements prior to distress. The two-year period is an exaggeration since the average lead time for the correctly classified firms is approximately 20 months, with two firms having a 13-month lead. The results are shown in table 1.3. The reduction in accuracy is understandable because impending bankruptcy is more remote and the indications are less clear. Nevertheless, 72 percent correct assignment is evidence that bankruptcy can be predicted two years prior to the event. The Type II error is slightly larger (6 percent vs. 3 percent) in this test, but still it is extremely accurate. Further tests are applied below to determine the accuracy of predicting bankruptcy as much as five years prior to the actual event.

Table 1.3 Classification results, two statements prior to bankruptcy

						Predicted	
	Number correct	*% correct*	*% error*	n	*Actual*	*Group 1 (Bankrupt)*	*Group 2 (Non-bankrupt)*
					Group 1	23	9
					Group 2	2	31
Type I	23	72	28	32			
Type II	31	94	6	33			
Total	54	83	17	65			

Potential Bias and Validation Techniques

When the firms used to determine the discriminant coefficients are reclassified, the resulting accuracy is biased upward by

1 sampling errors in the original sample; and
2 search bias.

The latter bias is inherent in the process of reducing the original set of variables (22) to the best variable profile (5). The possibility of bias due to intensive searching is inherent in any empirical study. While a subset of variables is effective in the initial sample, there is no guarantee that it will be effective for the population in general.

The importance of secondary sample testing cannot be overemphasized. One type of secondary sample testing is to estimate parameters for the model using only a subset of the original sample, and then to classify the remainder of the sample based on the parameters established. A simple t-test is then applied to test the significance of the results. Five different replications of the suggested method of choosing subsets (16 firms) of the original sample are tested.

The test results reject the hypothesis that there is no difference between the groups and substantiate that the model does, in fact, possess discriminating power on observations other than those used to establish the parameters of the model. Therefore, any search bias does not appear significant.

Secondary Sample of Bankrupt Firms

To test the model rigorously for both bankrupt and non-bankrupt firms, two new samples are introduced. The first contains a new sample of 25 bankrupt firms whose asset size range is similar to that of the initial bankrupt group. On the basis of the parameters established in the discriminant model to classify firms in this secondary sample, the predictive accuracy for this sample as of one statement prior to bankruptcy is described in table 1.4.

The results here are surprising in that one would not usually expect a secondary sample's results to be superior to the initial discriminant sample (96 percent vs. 94 percent). Two possible reasons are that the upward bias normally present in the initial sample tests is not manifested in this investigation and/or that the model, as stated before, is not optimal.

Table 1.4 Classification results, secondary sample of bankrupt firms

	Bankrupt	*Group*	*(Actual)*	*Predicted*	
	Number correct	*% correct*	*% error*	*Bankrupt*	*Non-bankrupt*
				24	1
Type I (Total)	24	96	4	$N = 25$	

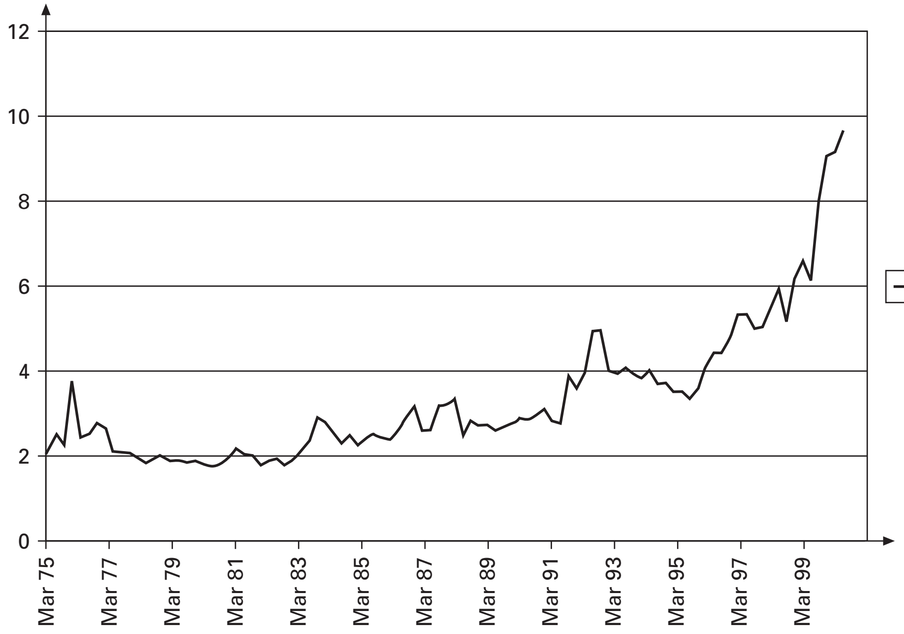
Testing the Model on Subsequent Distressed Firm's Samples

In three subsequent tests, I examined 86 distressed companies from 1969–75, 110 bankrupts from 1976–95 and 120 from 1997–99. I found that the Z-Score model, using a cutoff score of 2.675, was between 82 percent and 94 percent accurate. For an in-depth discussion of these studies, see below. In repeated tests up to the present (1999), the accuracy of the Z-Score model on samples of distressed firms has been in the vicinity of 80–90 percent, based on data from one financial reporting period prior to bankruptcy.

The Type II error (classifying the firm as distressed when it does not go bankrupt), however, has increased substantially with as much as 15–20 percent of all firms and 10 percent of the largest firms having Z-Scores below 1.81. Recent tests, however, show the average Z-Score increasing significantly with the average rising from the 4–5 level in 1970–95 period to almost 10 (ten) in 1999; see Osler and Hong (2000) for these results, shown also in figure 1.1. But, the median level has not increased much. The majority of increase in average Z-Scores was due to the dramatic climb in stock prices and its impact on X_4.

I advocate using the lower bond of the zone-of-ignorance (1.81) as a more realistic cutoff Z-Score than the score 2.675. The latter resulted in the lowest overall error in the

Figure 1.1 Average Z-Scores: US industrial firms 1975–99



Source: Osler and Hong (2000).

original tests. In 1999, the proportion of US industrial firms, comprised in the *Compustat* data tapes, that had Z-Scores below 1.81 was over 20 percent.

Secondary Sample of Non-Bankrupt Firms

Up to this point, the sample companies were chosen either by their bankruptcy status (Group I) or by their similarity to Group I in all aspects except their economic well-being. But what of the many firms which suffer temporary profitability difficulties, but actually do not become bankrupt? A bankruptcy classification of a firm from this group is an example of a Type II error. An exceptionally rigorous test of the discriminant model's effectiveness would be to search out a large sample of firms that have encountered earning problems and then to observe the Z-Score's classification results.

To perform the above test, a sample of 66 firms is selected on the basis of net income (deficit) reports in the years 1958 and 1961, with 33 from each year. Over 65 percent of these firms had suffered two or three years of negative profits in the previous three years. The firms are selected regardless of their asset size, with the only two criteria being that they were manufacturing firms which suffered losses in the year 1958 or 1961. The companies are then evaluated by the discriminant model to determine their bankruptcy potential.

The results show that 14 of the 66 firms are classified as bankrupt, with the remaining 52 correctly classified. Therefore, the discriminant model correctly classified 79 percent of the sample firms. This percentage is all the more impressive when one considers that these firms constitute a secondary sample of admittedly below-average performance. The t-test for the significance of the result is $t = 4.8$; significant at the 0.001 level. Another interesting facet of this test is the relationship of these "temporarily" sick firms' Z-Scores and the "zone of ignorance." The zone of ignorance is that range of Z-Scores where misclassification can be observed.

Of the 14 misclassified firms in this secondary sample, 10 have Z-Scores between 1.81 and 2.67, which indicates that although they are classified as bankrupt, the prediction of their bankruptcy is not as definite as it is for the vast majority in the initial sample of bankrupt firms. In fact, just under one-third of the 66 firms in this last sample have Z-Scores within the entire overlap area, which emphasizes that the selection process is successful in choosing firms which showed signs (profitability) of deterioration. Although these tests are based on data from over 40 years ago, they do indicate the robustness of the model which is still in use in the year 2000.

Long-Range Accuracy

The previous results give important evidence of the reliability of the conclusions derived from the initial and holdout samples of firms. An appropriate extension would be to examine the overall effectiveness of the discriminant model for a longer period of time prior to bankruptcy.

To answer this question, data are gathered for the 33 original firms from the third, fourth, and fifth years prior to bankruptcy. One would expect on an *a priori* basis that, as the lead time increases, the relative predictive ability of any model would decrease. This was true in the univariate studies cited earlier, and it is also quite true for the multiple discriminant model. We will shortly see, however, that the more recent model (e.g., ZETA®) has demonstrated higher accuracy over a longer period of time.

Based on the above results, it is suggested that the Z-Score model is an accurate forecaster of failure up to two years prior to distress and that accuracy diminishes substantially as the lead time increases. We also performed a trend analysis on the individual ratios in the model. The two most important conclusions of this trend analysis are that

1. all of the observed ratios show a deteriorating trend as bankruptcy approaches; and
2. the most serious change in the majority of these ratios occurred between the third and the second years prior to bankruptcy.

The degree of seriousness is measured by the yearly change in the ratio values. The latter observation is extremely significant as it provides evidence consistent with conclusions derived from the discriminant model. Therefore, the important information inherent in the individual ratio measurement trends takes on deserved significance only when integrated with the more analytical discriminant analysis findings.

Average Z-Scores Over Time

As table 1.5 shows, we have tested the Z-Score model for various sample periods over the last 30 years. In each test, the Type I accuracy using a cutoff score of 2.67 had a

Table 1.5 Classification and prediction accuracy Z-Score (1968) failure model*

Year prior to failure	*Original sample (33)*	*Holdout sample (25)*	*1969–75 Predictive sample (86)*	*1976–95 Predictive sample (110)*	*1997–99 Predictive sample (120)*
1	94% (88%)	96% (92%)	82% (75%)	85% (78%)	94% (84%)
2	72%	80%	68%	75%	74%
3	48%	–	–	–	–
4	29%	–	–	–	–
5	36%	–	–	–	–

** Using 2.67 as cutoff score (1.81 cutoff accuracy in parenthesis).*

range of 82–94 percent, based on data from one financial statement prior to bankruptcy or default on outstanding bonds. Indeed, in the most recent test, based on 120 firms which defaulted on their publicly held debt during 1997–99, the default prediction accuracy rate was 94 percent (113 out of 120). Using the more conservative 1.81 cutoff, the accuracy rate was still an impressive 84 percent. The 94 percent, 2.67 cutoff accuracy is comparable to the original sample's accuracy which was based on data used to construct the model itself.

We can, therefore, conclude that the Z-Score model has retained its reported high accuracy and is still robust despite its development over 30 years ago. In the last decade, however, the Type II accuracy has increased to about 15–20 percent of those manufacturing firms listed on *Compustat*.

Adaptation for Private Firms' Application

Perhaps the most frequent inquiry that I have received from those interested in using the Z-Score model is, "What should we do to apply the model to firms in the private sector?" Credit analysts, private placement dealers, accounting auditors, and firms themselves are concerned that the original model is only applicable to publicly traded entities (since X_1 requires stock price data). And, to be perfectly correct, the Z-Score model is a publicly traded firm model and *ad hoc* adjustments are not scientifically valid. For example, the most obvious modification is to substitute the book value of equity for the market value and then recalculate V_4X_4. Prior to this writing, analysts had little choice but to do this procedure since valid alternatives were not available.

A Revised Z-Score Model

Rather than simply insert a proxy variable into an existing model to calculate z-scores, I advocate a complete reestimation of the model, substituting the book values of equity for the market value in X_4. One expects that all of the coefficients will change (not only the new variable's parameter) and that the classification criterion and related cutoff scores would also change. That is exactly what happens.

The results of our revised Z-Score model with a new X_4 variable is:

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.998X_5$$

The equation now looks different than the earlier model (page 14); note, for instance, the coefficient for X_1 went from 1.2 to 0.7. But, the model looks quite similar to the one using market values. The actual variable that was modified, X_4, showed a coefficient change to 0.42 from 0.60; that is, it now has less of an impact on the Z-Score. X_3 and X_5 are virtually unchanged. The univariate F-test for the book value of X_4 (25.8) is lower than the 33.3 level for the market value but the scaled vector results show that the revised book value measure is still the third most important contributor.

Table 1.6 lists the classification accuracy, group means, and revised cutoff scores for the Z′-Score model. The Type I accuracy is only slightly less impressive than the model utilizing market value of equity (91 percent vs. 94 percent) but the Type II accuracy is identical (97 percent). The non-bankrupt group's mean Z′-Score is lower than that of the original model (4.14 vs. 5.02). Therefore, the distribution of scores is now tighter with larger group overlap. The gray area (or ignorance zone) is wider, however, since the lower boundary is now 1.23 as opposed to 1.81 for the original Z-Score model. All of this indicates that the revised model is probably somewhat less reliable than the original, but only slightly less. Due to lack of a private firm data base, we have not tested this model extensively on secondary sample distressed and nondistressed entities. A recent model from Moody's (2000) utilizing data on middle market firms and over 1600 defaults, concentrates on private firms.

Table 1.6 Revised Z′-Score model: Classification results, group means, and cutoff boundaries

	Classified		
Actual	*Bankrupt*	*Non-bankrupt*	*Total*
Bankrupt	30 (90.9%)	3 (9.1%)	33
Non-bankrupt	1 (3.0%)	32 (97.0%)	33

Note: Bankrupt group mean = 0.15; non-bankrupt group mean = 4.14.
$Z' < 1.23$ = Zone I (no errors in bankruptcy classification):
$Z' > 2.90$ = Zone II (no errors in non-bankruptcy classification):
gray area = 1.23 to 2.90.

A Further Revision – Adapting the Model for Non-Manufacturers

The next modification of the Z-Score model analyzed the characteristics and accuracy of a model without X_1 – sales/total assets. We do this to minimize the potential industry effect which is more likely to take place when such an industry-sensitive variable as asset turnover is included. In addition, I have used this model to assess the financial health of non-US corporates. In particular, Altman et al. (1995; ch. 5, this volume) have applied this enhanced Z″-Score model to emerging markets corporates, specifically Mexican firms that had issued Eurobonds denominated in US dollars. The book value of equity was used for X_4 in this case.

The classification results are identical to the revised five-variable model (Z′-Score). The new Z″-Score model is:

$$Z'' = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$$

All of the coefficients for variables X_1–X_4 are changed as are the group means and cutoff scores. This particular model is also useful within an industry where the type of financing of assets differs greatly among firms and important adjustments, like lease capitalization, are not made. In the emerging market model, we added a constant term of +3.25 so as to standardize the scores with a score of Zero (0) equated to a D (default) rated bond.

Emerging Market Scoring (EMS) Model and Process

Emerging markets credits may initially be analyzed in a manner similar to that used for traditional analysis of US corporates. Once a quantitative risk assessment has emerged, an analyst can then use a qualitative assessment to modify it for such factors as currency and industry risk, industry characteristics, and the firm's competitive position in that industry. It is not often possible to build a model specific to an emerging market country based on a sample from that country because of the lack of credit experience there. To deal with this problem, Altman et al. (1995; ch. 5, this volume) have modified the original Altman Z-Score model to create the EMS model.

The process of deriving the rating for a Mexican corporate credit is as follows:

1. The EMS score is calculated, and equivalent rating is obtained based on the calibration of the EMS scores with US bond-rating equivalents (table 1.7).
2. The company's bond is then analyzed for the issuing firm's vulnerability concerning the servicing of its foreign currency-denominated debt. This vulnerability is based on the relationship between the nonlocal currency revenues minus costs, compared with nonlocal currency expense. Then the level of nonlocal currency cash flow is compared with the debt coming due in the next year. The analyst adjusts the rating downward depending on the degree of vulnerability seen.
3. The rating is further adjusted downward (or upward) if the company is in an industry considered to be relatively riskier (or less risky) than the bond-rating equivalent from the first EMS result.
4. The rating is further adjusted up or down depending on the dominance of the firm's position in its industry.
5. If the debt has special features, such as collateral or a bona fide guarantor, the rating is adjusted accordingly.
6. Finally, the market value of equity is substituted for the book value in variable X_4, and the resulting bond-rating equivalents are compared. If there are significant differences in the bond-rating equivalents, the final rating is modified, up or down.

For relative value analysis, the corresponding US corporates' credit spread is added to the sovereign bond's option-adjusted spread. Only a handful of the Mexican companies were rated by the rating agencies. Thus, risk assessments such as those provided by EMS are often the only reliable indicators of credit risk to overseas investors in Mexico. Altman et al. (1995) report that the modified ratings have proven accurate in

Table 1.7 US bond rating equivalent based on EM score

US equivalent rating	*Average EM score*
AAA	8.15
AA+	7.60
AA	7.30
AA–	7.00
A+	6.85
A	6.65
A–	6.40
BBB+	6.25
BBB	5.85
BBB–	5.65
BB+	5.25
BB	4.95
BB–	4.75
B+	4.50
B	4.15
B–	3.75
CCC+	3.20
CCC	2.50
CCC–	1.75
D	0

Source: In-depth Data Corp. Average based on over 750 US corporates with rated debt outstanding: 1994 data.

anticipating both downgrades and defaults – Group Synkro (10/95), Situr (3/96), GMD (8/97), Tribasa (3/99), etc. – and upgrades (Aeromexico in July 1995).

THE ZETA® CREDIT-RISK MODEL

In 1977, Altman et al. (1977) constructed a second generation model with several enhancements to the original Z-Score approach. The purpose of this study was to construct, analyze and test a new bankruptcy classification model which considers explicitly recent developments with respect to business failures. The new study also incorporated refinements in the utilization of discriminant statistical techniques. Several reasons for building a new model, despite the availability of several fairly impressive "old" models, are presented below and the empirical results seem to substantiate the effort. The new model, which we call ZETA®, was effective in classifying bankrupt companies up to five years prior to failure on a sample of corporations consisting of manufacturers and retailers. Since the ZETA® model is a proprietary effort, I cannot fully disclose the parameters of the market.

Reasons for Attempting to Construct a New Model

There are at least five valid reasons why a revised Z-Score bankruptcy classification model can improve on and extend those statistical models which had been published in the literature in the prior decade:

1 One such reason is the change in the size, and perhaps the financial profile, of business failures. The average size of bankrupt firms had increased dramatically with the consequent greater visibility and concern from financial institutions, regulatory agencies and the public at large. Most of the past studies used relatively small firms in their samples with the exception of Altman's (1973) railroad study and the commercial bank studies. Any new model should be as relevant as possible to the population to which it will eventually be applied. This present study utilizes a bankrupt firm sample where the average asset size two annual reporting periods prior to failure was approximately \$100 million. No firm had less than \$20 million in assets.
2 Following (1) above, a new model should be as current as possible with respect to the temporal nature of the data.
3 Past failure models concentrated either on the broad classification of manufacturers or on specific industries. I feel that with the appropriate analytical adjustments, retailing companies, a particularly vulnerable group, could be analyzed on an equal basis with manufacturers.
4 An important feature of this study is that the data and footnotes to financial statements have been scrupulously analyzed to include the most recent changes in financial reporting standards and accepted accounting practices. Indeed, in at least one instance, a change which was scheduled to be implemented in a very short time was applied. The purpose of these modifications was to make the model not only relevant to past failures, but to the data that will appear in the future. The predictive as well as the classification accuracy of the ZETA® model is implicit in our efforts.
5 It is also important to test and assess several of the then recent advances and still controversial aspects of discriminant analysis.

Principal Findings

We concluded that the new ZETA® model for bankruptcy classification appeared to be accurate for up to five years prior to failure with successful classification of well over 90 percent of our sample one year prior and 70 percent accuracy up to five years. We also observed that the inclusion of retailing firms in the same model as manufacturers does not seem to affect our results negatively. This is probably true due to the adjustments to our data based on recent and anticipated financial reporting changes – primarily the capitalization of leases.

We also find that the ZETA® model outperformed alternative bankruptcy classification strategies in terms of expected cost criteria utilizing prior probabilities and explicit cost of error estimates. In our investigation, we were surprised to observe that, despite the statistical properties of the data which indicate that a quadratic structure is appropriate, the linear structure of the same model outperformed the quadratic in tests of model validity. This was especially evident regarding the long-term accuracy of the model and in holdout sample testing.

Sample and Data Characteristics, and Statistical Methodology

Our two samples of firms consist of 53 bankrupt firms and a matched sample of 58 non-bankrupt entities. The latter are matched to the failed group by industry and year of data. Our sample is almost equally divided into manufacturers and retailer groups, and 94 percent of the firms failed during the period 1969–75. The average asset size of our failed group is almost $100 million indicative of the increasing size of failures. The bankrupt firms represent all publicly held industrial failures which had at least $20 million in assets, with no known fraud involved and where sufficient data was available. Five non-bankruptcy petition companies were included due to either substantial government support, or a forced merger, or, the banks taking over the business or accepting a distressed restructuring rather than forcing the Chapter 11 petition.

VARIABLES ANALYZED

In other studies, a number of financial ratios and other measures have been found to be helpful in providing statistical evidence of impending failures. We have assembled data to calculate these variables and, in addition, have included several "new" measures that were thought to be potentially helpful as well. The 27 variables are listed in table 1.8, along with certain relevant statistics which will be discussed shortly. Note that in a few cases – e.g., nos. 7 and 9, tangible assets and interest coverage – the variables are expressed in logarithmic form so as to reduce outlier possibilities and to adhere to statistical assumptions. The variables can be classified as profitability (1–6), coverage and other earnings relative to leverage measures (8–14), liquidity (15–18), capitalization ratios (19–23), earnings variability (24–26) and a few miscellaneous measures (7 and 27).

REPORTING ADJUSTMENTS

As noted earlier, we have adjusted the basic data of our sample to consider explicitly several of the most recent and, in our opinion, the most important accounting modifications:

1 *Capitalization of leases*
Without doubt, the most important and pervasive adjustment made was to capitalize all noncancelable operating and finance leases. The resulting capitalized lease amount was added to the firms' assets and liabilities and also we imputed an interest cost to the "new" liability. The procedure involved preparation of schedules of current and expected lease payment obligations from information found in footnotes to the financial statements. The discount rate used to capitalize leases was the average interest rate for new issue, high grade corporate bonds in the year being analyzed plus a risk premium of 10 percent of the interest rate. An amount equal to the interest rate used in the capitalization process times the capitalized lease amount was added to interest costs. Subsequent to our analysis, the Financial Accounting Standards Board (FASB 13, 1980) stipulated that the appropriate discount rate to use is the lessee's cost of debt capital (before taxes) or the internal rate of return on the lease to the lessor, whichever is lower.

2 *Reserves*
If the firms' reserves were of a contingency nature, they were included in equity, and income was adjusted for the net change in the reserve for the year. If the reserve was related to the valuation of certain assets, it was netted against those assets. If the reserve was for contingent liabilities, e.g., law-suits, then it was added to the liabilities. This was the case for Johns Manville (bankruptcy filed in 1982) and A. H. Robins (in 1985) and several other healthcare lawsuits.

Table 1.8 Listing of all variables, group mean, and F-tests based on one period prior to bankruptcy data (ZETA model sample)

Variable		*Population means*		*Univariate*
No.	*Name*	*Failed*	*Non-failed*	*F-test*
(1)	EBIT/TA	–0.0055	0.1117	54.3
(2)	NATC/TC	–0.0297	0.0742	36.6
(3)	Sales/TA	1.3120	1.6200	3.3
(4)	Sales/TC	2.1070	2.1600	0.0
(5)	EBIT/Sales	0.0020	0.0070	30.2
(6)	NATC/Sales	–0.0153	0.0400	33.1
(7)	Log tang. Assets	1.9854	2.2220	5.5
(8)	Interest coverage	–0.5995	5.3410	26.1
(9)	Log no. (8)	0.9625	1.1620	26.1
(10)	Fixed charge coverage	0.2992	2.1839	15.7
(11)	Earnings/debt	–0.0792	0.1806	32.8
(12)	Earnings 5 yr Maturities	–0.1491	0.6976	8.8
(13)	Cash/flow fixed charges	0.1513	2.9512	20.9
(14)	Cash flow/TD	–0.0173	0.3136	31.4
(15)	WC/LTD	0.3532	2.4433	6.0
(16)	Current ratio	1.5757	2.6040	38.2
(17)	WC/total assets	0.1498	0.3086	40.6
(18)	WC/cash expenses	0.1640	0.2467	5.2
(19)	Ret. earn/total assets	–0.0006	0.2935	114.6
(20)	Book equity/TC	0.2020	0.5260	64.5
(21)	MV equity/TC	0.3423	0.6022	32.1
(22)	5 yr MV equity/TC	0.4063	0.6210	31.0
(23)	MV equity/total liabilities	0.6113	1.8449	11.6
(24)	S. e. of estimate of EBIT/TA (norm)	1.6870	5.784	33.8
(25)	BEIT drop	–3.2272	3.179	9.9
(26)	Margin drop	–0.2173	0.179	15.6
(27)	Capital lease/assets	0.2514	0.178	4.2
(28)	Sales/fixed assets	3.1723	4.179	3.5

Notation:
EBIT = *earnings before interest and taxes*
NATC = *net available for total capital*
TA = *total tangible assets*
LTD = *long term debt*
MV = *market value of equity*
TC = *total capital*
TD = *total debt*
WC = *working capital*
CF = *cash flow (before interest #13, after interest #14)*

3 *Minority interests and other liabilities on the balance sheet*
These items were netted against other assets. This allowed for a truer comparison of earnings with the assets generating the earning.

4 *Captive finance companies and other nonconsolidated subsidiaries*
These were consolidated with the parent company accounts as well as the information would allow. The pooling of interest method was used. This was made mandatory by the FASF in 1987.

5 *Goodwill and intangibles*
These were deducted from assets and equity because of the difficulty in assigning economic value to them.
6 *Capitalized R&D costs, capitalized interest and certain other deferred charges*
These costs were expensed rather than capitalized. This was done to improve comparability and to give a better picture of actual funds flows.

STATISTICAL METHODOLOGY

Distress classification is again attempted via the use of a multivariate statistical technique known as discriminant analysis. In this study, the results using both linear and quadratic structure are analyzed. The test for assessing whether a linear or quadratic structure is appropriate – sometimes referred to as the H_1 test, provides the proper guidance when analyzing a particular sample's classification characteristics. Essentially, if it is assessed that the variance–covariance matrices of the G groups are statistically identical, then the linear format which pools all observations is appropriate. If, however, the dispersion matrices are not identical, then the quadratic structure will provide the more efficient model since each group's characteristics can be assessed independently as well as between groups. Efficiency will result in more significant multivariate measures of group differences and greater classification accuracy of that particular sample. What has not been assessed up to this point, is the relative efficiency of the linear vs. quadratic structures when the sample data are not the same as that used to construct the model, i.e., holdout or secondary samples. This point is analyzed in the next section.

Empirical Results: The 7-variable Model

After an iterative process of reducing the number of variables, we selected a 7-variable model which not only classified our test sample well, but also proved the most reliable in various validation procedures. That is, we could not significantly improve on our results by adding more variables, and no model with fewer variables performed as well.

- X_1, return on assets, measured by the earnings before interest and taxes/total assets
  This variable has proven to be extremely helpful in assessing firm performance in several past multivariate studies.
- X_2, stability of earnings, measured by a normalized measure of the standard error of estimate around a 5–10-year trend in X_1
  Business risk is often expressed in terms of earnings fluctuations and this measure proved to be particularly effective. We did assess the information content of several similar variables which attempted to measure the potential susceptibility of a firm's earnings level to decline which could jeopardize its ability to meet its financial commitments. These variables were quite significant on a univariate level but did not enter into our final multivariate model.
- X_3, debt service, measured by the familiar interest coverage ratio, i.e., earnings before interest and taxes/total interest payments (including that amount imputed from the capitalized lease liability)

We have transposed this measure by taking the log to the base 10 so as to improve the normality and homoscedasticity of this measure.

- X_4, cumulative profitability, measured by the firm's retained earnings (balance sheet)/total assets
This ratio, which imputes such factors as the age of the firm, debt and dividend policy as well as its profitability record over time, was found to be quite helpful in the Z-Score model, discussed earlier. As our results show, this cumulative profitability measure is unquestionably the most important variable-measured univariately and multivariately.
- X_5, liquidity, measured by the familiar current ratio
Despite previous findings that the current ratio was not as effective in identifying failures as some other liquidity measures, we now find it slightly more informative than others, such as the working capital/total assets ratio.
- X_6, capitalization, measured by common equity/total capital
In both the numerator and the denominator, the common equity is measured by a five-year average of the total market value, rather than book value. The denominator also includes preferred stock at liquidating value, long-term debt and capitalized leases. We have utilized a 5-year average to smooth out possible severe, temporary market fluctuations and to add a trend component (along with X_2 above) to the study.
- X_7, size, measured by the firms' total assets
This variable, as is the case with the others, was adjusted for financial reporting changes. No doubt, the capitalization of leasehold rights has added to the average asset size of both the bankrupt and non-bankrupt groups. We have also transformed the size variable to help to normalize the distribution of the variable due to outlier observations. Again, a logarithmic transformation was applied.

RELATIVE IMPORTANCE OF DISCRIMINANT VARIABLES

The procedure of reducing a variable set to an acceptable number is closely related to an attempt to determine the relative importance within a given variable set. Several of the prescribed procedures for attaining the "best" set of variables, e.g., stepwise analysis, can also be used as a criterion for ranking importance. Unfortunately, there is no one best method for establishing a relative ranking of variable importance. Hence, we have assessed this characteristic by analyzing the ranks suggested by five different tests:

1. forward stepwise
2. backward stepwise
3. scaled vector (multiplication of the discriminant coefficient by the appropriate variance–covariance matrix item)
4. separation of means test
5. the conditional deletion test, which measures the additional contribution of the variable to the multivariate F-test given that the other variables have already been included.

In several studies that we have observed, the rankings across these tests are inconsistent and the researcher is left with a somewhat ambiguous answer. This was definitely not the case in our study.

Regardless of which test statistic is observed, the most important variable is the cumulative profitability ratio, X_4. In fact, our scaled vector analysis indicates that this single ratio contributes 25 percent of the total discrimination. Second in importance is the stability of earnings ratio (X_2) and, except for the univariate test of significance, it too has a consistent across tests.

LINEAR VS. QUADRATIC ANALYSIS

The H_1 test of the original sample characteristics clearly rejects the hypothesis that the group dispersion matrices are equal. Therefore, the linear structure classification rule (excluding error costs), is not appropriate and the quadratic structure appears to be the more efficient one.

As can be observed in table 1.9, the quadratic and linear models yield essentially equal total sample accuracy results for the original sample classifications, but the holdout sample tests indicate a clear superiority for the linear framework. This creates a dilemma and we have chosen to concentrate on the linear test due to

1 the possible high sensitivity to individual sample observaions of the quadratic parameters (that is, we observe 35 different parameters in the quadratic model compared with only 7 in the linear case, not including the intercept), and
2 the fact that all of the relative tests of importance are based on the linear model.

Table 1.9 Overall classification accuracy

Years prior to bankruptcy	*Bankrupt firms*		*Non-bankrupt firms*		*Total*	
	% Linear	*% Quadratic*	*% Linear*	*% Quadratic*	*% Linear*	*% Quadratic*
1 Original sample	96.2	94.3	89.7	91.4	92.8	92.8
1 (Lachenbruch validation test)	(92.5)	(85.0)	(89.7)	(87.9)	(91.0)	(86.5)
2 Holdout	84.9	77.4	93.1	91.9	89.0	84.7
3 Holdout	74.5	62.7	91.4	92.1	83.5	78.9
4 Holdout	68.1	57.4	89.5	87.8	79.8	74.0
5 Holdout	69.8	46.5	82.1	87.5	76.8	69.7

CLASSIFICATION ACCURACY – ORIGINAL AND HOLDOUT SAMPLES

Table 1.10 presents classification and holdout sample accuracy of the original sample based on data from one year prior to bankruptcy. Lachenbruch (1967) suggests an almost unbiased validation test of original sample results by means of a type of jack-knife, or one isolated observation at a time approach. The individual observations' classification accuracy is then cumulated over the entire sample. Years 2–5 "holdout" sample results are also presented. These results are listed for both the linear and quad-ratic structures of the 7-variable model.

The linear model's accuracy, based on one year prior data, is 96.2 percent for the bankrupt group and 89.7 percent for the non-bankrupt. The upward bias in these results appears to be slight since the Lachenbruch results are only 3 percent less for the

Table 1.10 Classification accuracy between the ZETA® model and various forms of the Z-Score model

	ZETA model		*Altman's 1968 model*		*1968 model, ZETA sample*		*1968 variables, ZETA parameters*	
Years prior to bankruptcy (1)	*Bankrupt (2)*	*Non-bankrupt (3)*	*Bankrupt (4)*	*Non-bankrupt (5)*	*Bankrupt (6)*	*Non-bankrupt (7)*	*Bankrupt (8)*	*Non-bankrupt (9)*
1	96.2	89.7	93.9	97.0	86.8	82.4	92.5	84.5
2	84.9	93.1	71.9	93.9	83.0	89.3	83.0	86.2
3	74.5	91.4	48.3	n.a.	70.6	91.4	72.7	89.7
4	68.1	89.5	28.6	n.a.	61.7	86.0	57.5	83.0
5	69.8	82.1	36.0	n.a.	55.8	86.2	44.2	82.1

Note: All figures given as percentages.

failed group and identical for the nonfailed group. As expected, the failed group's classification accuracy is lower as the data become more remote from bankruptcy, but still quite high. In fact, we observe 70 percent accuracy as far back as five years prior to failure. This compares very favorably to the results recorded by the Z-Score model, where the accuracy dropped precipitously after two years prior.

An interesting result was observed by comparing the quadratic structure's results for that of the linear (table 1.9). As noted earlier, the total samples' classification accuracy is identical for the two structures in period 1, with the linear showing a slight edge in the bankrupt group and the quadratic in the non-bankrupt group. The most obvious and important differences, however, are in the validation and "holdout" tests of the bankrupt group. Here, the linear model is clearly superior, with the quadratic misclassifying over 50 percent of the future bankrupts five years prior. The Lachenbruch validation test also shows a large bankrupt classification accuracy difference (over 7 percent favoring the linear model). Subsequent analysis will report only the linear results.

COMPARISON WITH THE Z-SCORE MODEL

Table 1.10 compares the original sample classification accuracy and also the accuracy for up to five years prior to financial distress of the Z-Score and ZETA® models. Note that the one-year prior classification accuracy of bankrupt firms is quite similar for both models (96.2 percent for ZETA® and 93.9 percent for Z-Score) but that the accuracy is consistently higher for the ZETA® model in years 2–5 prior to the distress date. Indeed, by the fifth year, the ZETA® model is still about 70 percent accurate but the Z-Score's accuracy falls to 36 percent. Note also that the Z-Score's accuracy on the ZETA® sample (columns 6 and 7) is actually considerably higher in years 2–5 than on the original sample. Finally, when we recalibrate the Z-Score model's coefficients based on the ZETA® sample, the classification results (column 8) are much better than the originals (column 4) in all but the first year prior.

Group Prior Probabilities, Error Costs and Model Efficiency

Earlier, we showed the classification rules for both linear and quadratic analyses. If one assumes equal prior probabilities of group membership, the linear model will result in a cutoff or critical score of zero. This is due to the constant term in the ZETA® model. All firms scoring above zero are classified as having characteristics similar to the non-bankrupt group and those with negative scores similar to bankrupts. The same zero cutoff score will result if one desired to minimize the total cost of misclassification. That is, assuming multi-normal populations and a common convariance matrix, the optimal cutoff score ZETA, is equal to

$$\text{ZETA}_c = \ln \frac{q_1 c_{\text{I}}}{q_2 c_{\text{II}}}$$

where q_1, q_2 is the prior probability of bankrupt (q_1) or non-bankrupt (q_2), and c_{I}, c_{II} are the costs of Type I and Type II errors, respectively.

Further, if one wanted to compare the efficiency of the ZETA® bankruptcy classification model with alternative strategies, the following cost function is appropriate for the expected cost of ZETA (EC_{ZETA}).

$$EC_{ZETA} = q_1 \frac{M_{12}}{N_1} c_{I} + q_2 \frac{M_{21}}{N_2} c_{II}$$

where M_{12}, M_{21} are the observed Type I and Type II errors (misses) respectively, and N_1, N_2 are the number of observations in the bankrupt (N_1) and non-bankrupt (N_2) groups.

In our tests, we have implicitly assumed equal prior probabilities and equal costs of errors, resulting in a zero cutoff score. We are actually aware, however, of the potential bias involved in doing so. Instead of attempting earlier to integrate probability priors and error costs, we have assumed equal estimates for each parameter, because to a great extent the two parameters neutralize each other, and it was much easier than attempting to state them precisely. The following is our reasoning.

The "correct" estimate of q_1 is probably in the range 0.01–0.05. That is, the prior probability that a firm will go bankrupt within a year or two is probably in this 0.01–0.05 range. Although the ZETA® model's parameters are based on data from one year prior to bankruptcy, it is not specifically a one-year prediction model. The procedure, in this sense, is attemporal. It is, in our opinion, incorrect to base one's prior probability estimates on a single year's reported statistics. In addition, there are many definitions of financial distress which economically approximate bankruptcy. These include non-judicial arrangements, extreme liquidity problems which require the firm's creditors or other external forces to take over the business or agree to a distressed restructuring (composition or extension of claims), bond default, etc. In the final analysis, we simply do not know the precise estimate of bankruptcy priors, but at the same time assert that one must assume the estimate is greater than a single year's reported data. Hence, we believe the prior probability estimate is in the 1–5 percent range. In the subsequent analysis, we utilize 2 percent.

COST OF CLASSIFICATION ERRORS

Another input that is imperative to the specification of an alternative to the zero cutoff score is the cost of error in classification. No prior study to the ZETA® analysis (Altman et al., 1977) had explicitly included this element analysis. To attempt to precise the cost component into an analysis of model efficiency, it is necessary to specify the decision maker's role. In this study, we utilize the commercial bank loan function as the framework of analysis. The Type I bankruptcy classification is analogous to that of an accepted loan that defaults and the Type II error to a rejected loan that would have resulted in a successful payoff. Many of the commercial factors involved in assessing these error costs were first noted in an excellent discussion [following Beaver's, (1967) paper] by Neter. It should be noted that even in 1999, commercial bankers are still struggling with a credible assumption of the total cost of lending errors.

An empirical study was performed to assess the costs of these lending errors with the following specification for the equivalent Type I (c_{I}) and Type II (c_{II}) error costs.

$$c_{\mathrm{I}} = 1 - \frac{LLR}{GLL} \qquad c_{\mathrm{II}} = r - i$$

where

LLR = amount of loan losses recovered
GLL = gross loan losses (charged-off)
r = effective interest rate on the loan
i = effective opportunity cost for the bank.

The commercial bank takes the risk of losing all or a portion of the loan should the applicant eventually default. The exact amount is a function of the success the bank has in recovering the loan principal. We are quite aware that there are additional costs involved in the recovery process, including legal, transaction, and loan charge-off officer opportunity costs. These costs are not reported but obviously increase the Type I error cost. In addition, if the Type II error (c_{II}) is positive, i.e., $r > 1$, then there will be an added cost element in c_{I}. This added element involves the lost interest on that remaining part of the loan which is not recovered (GLL – LLR) for the duration of the defaulted loan. We will examine c_{II} below, but will not include this added element in our calculation of c_{I}. Again, however, it is clear that we are underestimating c_{I} somewhat.

Recoveries in the Public Bond Market

While there has been almost no rigorous studies published which quantify the effective costs of lending errors for loans and other private placements, a number of recent studies have documented losses in the public bond markets (Altman and Eberhart, 1994; Moody's, 1995; Standard & Poor's, 1995). The former documents recoveries at default and also upon emergence from Chapter 11. These public bond market studies observe recoveries stratified by bond seniority. For commercial loans, the most likely equivalents to the public bond market are the straight (non-convertible) senior secured and senior unsecured classes. Table 1.11 lists these recoveries at the time of default and upon emergence from Chapter 11.

We have measured c_{I} based on annual report data from 26 of the largest US commercial banks and questionnaire returns from a sample of smaller, regional banks in the Southeast USA. A questionnaire was sent to approximately 100 Southeast banks

Table 1.11 Bond recoveries (% of par value), by seniority, at default and upon emergence from Chapter 11

Bond priority	N	*Recovery at default (%)*	*Recovery upon emergence (%)*
Senior secured	24	60.51	100.91
Senior unsecured	71	52.28	81.05
Senior subordinated	35	30.70	23.38
Subordinated	54	27.96	32.41

Source: Altman and Eberhart (1994), Altman (1993).

with 33 usable responses. The range of commercial bank asset sizes in this small-bank sample was between \$12 million and \$3 billion, with the average equal to \$311 million and the median equal to \$110 million. The large-bank sample's asset size averaged \$13.4 billion with a \$10 billion median.

Both the data sources encompass a five-year period, 1971–75 inclusive, and we measure the average loan loss recovery statistics for senior unsecured loans on a contemporary and a one-year lag (recoveries lagging charge-offs) basis. The results of this investigation show that the average c_{I} on a contemporary basis is in the 76.7–83.0 percent range; when measured on a one-year lag basis, the averages are lower (68.6–72.2 percent). The year 1975 was an abnormally high loan charge-off year in the US banking system and since this data is included in the contemporary statistics but not in the one-year lag data, we believe the more representative result for c_{I} is in the vicinity of 70 percent. We use this statistic for c_{I}.

The simple formula for c_{II} specifies that the decision not to lend to an account that would have repaid successfully forgoes the return on that loan, but the loss is mitigated by the alternative use of loanable funds. In its strictest sense, the bank's opportunity cost implies another loan at the same risk which is assumed to pay off. In this case, c_{II} is probably zero or extremely small. Conservatively speaking, however, an account is rejected due to its high risk characteristics and alternative uses probably will carry lower risk attributes. Hence, $r - i$ will be positive but still quite low. Carried to the other extreme, the alternative use would be an investment in a riskless asset, i.e., government securities of the same maturity as the loan, and $r - i$ will be somewhat higher – perhaps 2–4 percent. The relationship between r and i will vary over time and is particularly sensitive to the demand and supply equilibrium relationship for loanable funds. As an approximation, we specify c_{II} as 2 percent; hence $c_{\mathrm{I}}/c_{\mathrm{II}}$ is equal to 35 times (0.70/0.20).

Revised Cutoff Score and Model Efficiency Tests

With respect now to the calculation of the critical or cutoff score $\mathrm{ZETA_c}$, we have,

$$\mathrm{ZETA_c} = \ln \frac{q_1 c_{\mathrm{I}}}{q_2 c_{\mathrm{II}}} = \ln \frac{0.02 \cdot 0.70}{0.98 \cdot 0.02} = \ln 0.714$$

$$\mathrm{ZETA_c} = -0.337.$$

Before comparing the efficiency of the various alternative bankruptcy classification strategies, it should be noted that the observed classification accuracy of a model such as ZETA will change with the new cutoff score. For example, with the cutoff score of –0.337, the number of Type I errors increases from two (3.8 percent) to four (7.6 percent), while the Type II errors decreases from 6 (10.3 percent) to 4 (7.0 percent).

Adjustments to the Cutoff Score and Practical Applications

In addition to the utilization of prior probabilities of group membership and cost estimates of classification errors for comparative model efficiency assessment, these inputs could prove valuable for practical application purposes. For instance, the bank lending-officer or loan-review analyst may wish to be able to logically adjust the critical cutoff score to consider his own estimates of group priors and error costs and/or to reflect current economic conditions in the analysis. One could imagine the cutoff score falling (thereby lowering the acceptance criterion) as business conditions improve and the banker's prior probability of bankruptcy estimate falls from say 0.02 to 0.015. Or, a rise in cutoff scores could result from a change (rise) in the estimate of the Type I error cost *vis-à-vis* the Type II error cost. The latter condition possibly will occur for different decision makers. For instance, the cost to a portfolio manager of not selling a security destined for failure is likely to be extremely high relative to his cost of not investing in a stock (which does not fail) due to its relatively low $ZETA_c$. The portfolio manager may indeed want to raise the cutoff or threshold level to reduce the possibility of intangible (law suit costs) as well as tangible (lower prices) costs involved with holding a failed company's stock.

Another example of a practical application of cutoff score adjustment is the case of an accounting auditor. He might wish to use the model to decide whether a "going concern" qualified opinion should be applied. His expected cost for doing so is likely to be quite high (loss of client) relative to the expected cost of a stockholder law suit. This might lead to a fairly low cutoff score. On the other hand, the environment may be such that the law suit expected cost is prohibitive.

Conclusions

The ZETA® model for assessing bankruptcy risk of corporations demonstrates improved accuracy over existing failure classification model (Z-Score) and, perhaps more importantly, is based on data more relevant to current conditions and to a larger number of industrial firms. Recall, however, our use of the Z″ model for non-manufacturers. We are concerned with refining existing distress classification techniques by the use of the most relevant data, combined with developments in the application of discriminant analysis to finance. The ZETA® model's bankruptcy classification accuracy ranges from over 96 percent (93 percent holdout) one period prior to bankruptcy to 70 percent five annual reporting periods prior. We have assessed the effect of several elements involved with the application of discriminant analysis to financial problems. These include linear vs. quadratic analysis for the original and holdout samples, introduction of prior probabilities of group membership and costs of error estimates into the classification rule, and comparison of the model's results with naïve bankruptcy classification strategies.

The potential applications of the ZETA® bankruptcy identification model are in the same spirit as previously developed models. These include creditworthiness analysis of firms for financial and non-financial institutions, identification of undesirable investment risk for portfolio managers and individual investors, and to aid in more effective internal and external audits of firms with respect to going-concern considerations, among others.

REFERENCES

Altman, E., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance* 23, September 1968, 589–609.

Altman, E., "Predicting Railroad Bankruptcies in America," *Bell Journal of Economics and Management Service*, Spring 1973.

Altman, E., *Corporate Financial Distress and Bankruptcy*, 2nd ed., John Wiley & Sons, New York, 1993.

Altman, E., and A. C. Eberhart, "Do Seniority Provisions Protect Bondholders' Investments?" *Journal of Portfolio Management*, Summer 1994.

Altman, E., and J. LaFleur, "Managing a Return to Financial Health," *Journal of Business Strategy*, Summer 1981.

Altman, E., R. Eisenbeis, and J. Sinkey, *Applications of Classification Procedures in Business, Banking and Finance*, JAI Press, Greenwich, CT, 1981.

Altman, E., R. Haldeman, and P. Narayanan, "ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations," *Journal of Banking and Finance*, June 1977.

Altman, E., J. Hartzell, and M. Peck, "*Emerging Markets Corporate Bonds: A Scoring System*," Salomon Brothers Inc, New York, 1995.

Beaver, W., "Financial Ratios as Predictors of Failures," in *Empirical Research in Accounting*, selected studies, 1966 in supplement to the *Journal of Accounting Research*, January 1967.

Deakin, E. B., "A Discriminant Analysis of Predictors of Business Failure," *Journal of Accounting Research*, March 1972.

Dun & Bradstreet, "The Failure Record," 1994, and annually.

Fisher, L., "Determinants of Risk Premiums on Corporate Bonds," *Journal of Political Economy*, June 1959.

Lachenbruch, P. A., "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant analysis," *Biometrics*, 23, 1967.

Moody's, "Corporate Bond Defaults and Default Rates," L. Carty, D. Lieberman, and J. Fons, January 1995.

Moody's, E. Falkensten, A. Boral and L. Carty, "RiskCalc™ Private Model: Moody's Default Model for Private Firms," Global Credit Research, May 2000.

Osler, C. and G. Hong, "Rapidly Rising Corporate Debt: Are Firms Now Vulnerable to an Economic Slowdown?" *Current Issues in Economics & Finance*, Federal Reserve Bank of New York, June 2000.

Standard & Poor's, "Corporate Defaults Level Off in 1994," L. Brand, T. K. Ho, and R. Bahar, *CreditWeek*, May 1, 1995.

ZETA®, "ZETA Analysis, A New Model for Bankruptcy Classification," *Journal of Banking and Finance*, June 1977.