Journal of Accounting Research
Vol. 18 No. 1 Spring 1980
Printed in U.S.A.

Financial Ratios and the Probabilistic Prediction of Bankruptcy

JAMES A. OHLSON*

1. *Introduction*

This paper presents some empirical results of a study predicting corporate failure as evidenced by the event of bankruptcy. There have been a fair number of previous studies in this field of research; the more notable published contributions are Beaver [1966; 1968*a*; 1968*b*], Altman [1968; 1973], Altman and Lorris [1976], Altman and McGough [1974], Altman, Haldeman, and Narayanan [1977], Deakin [1972], Libby [1975], Blum [1974], Edmister [1972], Wilcox [1973], Moyer [1977], and Lev [1971]. Two unpublished papers by White and Turnbull [1975*a*; 1975*b*] and a paper by Santomero and Vinso [1977] are of particular interest as they appear to be the first studies which logically and systematically develop probabilistic estimates of failure. The present study is similar to the latter studies, in that the methodology is one of maximum likelihood estimation of the so-called conditional logit model.

The data set used in this study is from the seventies (1970–76). I know of only three corporate failure research studies which have examined data from this period. One is a limited study by Altman and McGough [1974] in which only failed firms were drawn from the period 1970–73 and only one type of classification error (misclassification of failed firms) was analyzed. Moyer [1977] considered the period 1965–75, but the sample of bankrupt firms was unusually small (twenty-seven firms). The

* Associate Professor, University of California, Berkeley. I gratefully acknowledge the financial support from the Wells Fargo Bank. My thanks are also due to R. Wagner and R. Benin, who provided able and valuable assistance in the course of the project. G. Feltham, R. Hamilton, V. Anderson, W. Beaver, and R. Holland supplied valuable comments on earlier versions of this paper. [Accepted for publication March 1979.]

third study, by Altman, Haldeman, and Narayanan [1977], which "updates" the original Altman [1968] study, basically considers data from the period 1969 to 1975. Their sample was based on fifty-three failed firms and about the same number of nonfailed firms. In contrast, my study relies on observations from 105 bankrupt firms and 2,058 nonbankrupt firms. Although the other three studies differ from the present one so far as methodology and objectives are concerned, it is, nevertheless, interesting and useful to compare their results with those presented in this paper.

Another distinguishing feature of the present study which I should stress is that, contrary to almost all previous studies, the data for failed firms were not derived from *Moody's Manual.*[1] The data were obtained instead from 10-K financial statements as reported at the time. This procedure has one important advantage: the reports indicate at what point in time they were released to the public, and one can therefore check whether the company entered bankruptcy prior to or after the date of release. Previous studies have not explicitly considered this timing issue. Some studies, but by no means all, seem implicitly to presume that a report is available at the fiscal year-end date. The latter may or may not be appropriate, depending on the purpose of the study. However, if the purpose is one of investigating pure forecasting relationships, as is the case in this study, then the latter procedure is inadequate. This follows because it is possible that a company files for bankruptcy at some point in time after the fiscal year date, but prior to releasing the financial statements. This is not a trivial problem and neglecting this possibility may lead to "back-casting" for many of the failed firms.

The major findings of the study can be summarized briefly. First, it was possible to identify four basic factors as being statistically significant in affecting the probability of failure (within one year). These are: (*i*) the size of the company; (*ii*) a measure(s) of the financial structure; (*iii*) a measure(s) of performance; (*iv*) a measure(s) of current liquidity (the evidence regarding this factor is not as clear as compared to cases (*i*)–(*iii*)). Second, previous studies appear to have overstated the predictive (in the sense of forecasting) power of models developed and tested. The point of concern is the one alluded to above, that is, if one employs predictors derived from statements which were released after the date of bankruptcy, then the evidence indicates that it will be easier to "predict" bankruptcy. However, even if one allows for this factor, for the sample of firms used in this study, the prediction error-rate is larger in comparison to the rate reported in the original Altman [1968] study as well as most other studies using data drawn from periods prior to 1970. More important, the prediction error-rate is also larger than the one reported in Altman et al. [1977]. On the other hand, the Altman and McGough [1974]

[1] The only exception appears to be the Altman and McGough [1974] study. Altman et al. [1977] do not describe how they derived their data.

and Moyer [1977] studies report significantly larger error-rates, which are comparable to those found in this study. I have not been able completely to account for this most significant difference in the error-rates reported here, in Altman and McGough [1974], and in Moyer [1977], as compared to Altman et al. [1977]. (Any period dependence should after all be relatively minor.)

The model(s) are relatively simple to apply and may be of use in practical applications. The data requirements are such that all of the predictors are easily retrieved from the *Compustat* file. A potential disadvantage is that the model does not utilize any market transactions (price) data of the firms. One may, of course, expect that the predictive power of the model could be enhanced by incorporating such data.[2]

However, one might ask a basic and possibly embarrassing question: why forecast bankruptcy? This is a difficult question, and no answer or justification is given here. It could, perhaps, be argued that we are dealing with a problem of "obvious" practical interest. This is questionable since real-world problems concern themselves with choices which have a richer set of possible outcomes. No decision problem I can think of has a payoff space which is partitioned naturally into the binary status bankruptcy versus nonbankruptcy. (Even in the case of a "simple" loan decision, the payoff configuration is much more complex.) Existing empirical studies reflect this problem in that there is no consensus on what constitutes "failure," with definitions varying significantly and arbitrarily across studies. In other words, the dichotomy bankruptcy versus no bankruptcy is, at the most, a very crude approximation of the payoff space of some hypothetical decision problem. It follows that it is essentially a futile exercise to try to establish the relative decision usefulness of alternative predictive systems. Accordingly, I have not concerned myself with how bankruptcy (and/or failure) "ought" to be defined; I also have refrained from making inferences regarding the relative usefulness of alternative models, ratios, and predictive systems (e.g., univariate versus multivariate). Most of the analysis should simply be viewed as descriptive statistics—which may, to some extent, include estimated prediction error-rates—and no "theories" of bankruptcy or usefulness of financial ratios are tested. Even so, there are a large number of difficult statistical and methodological problems which need to be discussed. Many important problems pertaining to the development of data for bankrupt firms have gone mostly unnoticed in the literature.

2. *Some Comments Regarding Methodology and Data Collection*

The econometric methodology of conditional logit analysis was chosen to avoid some fairly well known problems associated with Multivariate

[2] I am currently undertaking work in this direction. I should note further that the use of price data implicitly is another way of using more information. Hence, it can also be viewed as another way of indirect use of accounting data.

Discriminant Analysis (*MDA*, for short). The *MDA* approach has been the most popular technique for bankruptcy studies using vectors of predictors. Among some of the problems with these studies are:[3] (*i*) There are certain statistical requirements imposed on the distributional properties of the predictors. For example, the variance–covariance matrices of the predictors should be the same for both groups (failed and nonfailed firms); moreover, a requirement of normally distributed predictors certainly mitigates against the use of dummy independent variables. A violation of these conditions, it could perhaps be argued, is unimportant (or simply irrelevant) if the only purpose of the model is to develop a discriminating device. Although this may be a valid point, it is nevertheless clear that this perspective limits the scope of the investigation. Under many circumstances, it is of interest to go through more traditional econometric analysis and test variables for statistical significance, etc. (*ii*) The output of the application of an *MDA* model is a score which has little intuitive interpretation, since it is basically an ordinal ranking (discriminatory) device. For decision problems such that a misclassification structure is an inadequate description of the payoff partition, the score is not directly relevant.[4] If, however, prior probabilistics of the two groups are specified, then it is possible to derive posterior probabilities of failure. But, this Bayesian revision process will be invalid or lead to poor approximations unless the assumptions of normality, etc. are satisfied. (*iii*) There are also certain problems related to the "matching" procedures which have typically been used in *MDA*. Failed and nonfailed firms are matched according to criteria such as size and industry, and these tend to be somewhat arbitrary. It is by no means obvious what is really gained or lost by different matching procedures, including no matching at all. At the very least, it would seem to be more fruitful actually to include variables as predictors rather than to use them for matching purposes.

The use of conditional logit analysis, on the other hand, essentially avoids all of the problems discussed with respect to *MDA*. The fundamental estimation problem can be reduced simply to the following statement: given that a firm belongs to some prespecified population, what is the probability that the firm fails within some prespecified time period? No assumptions have to be made regarding prior probabilities of bankruptcy and/or the distribution of predictors. These are the major advantages. The statistical significance of the different predictors are obtained from asymptotic (large sample) theory. To be sure, as is the case in any parametric analysis, a model must be specified, so there is always room for misspecification of the basic probability model. (Al-

[3] See also Eisenbeis [1977] and Joy and Tollefson [1975] for extensive discussions.

[4] The payoff partition will be inadequate if it is not feasible to define a utility function over the two types of classification errors. Any economic decision problem would typically require a richer state partition.

though it is possible to test for misspecification, it is beyond the confines of this paper to discuss and report on the results of such tests.)

Regardless of the virtues of probabilistic prediction over *MDA*, there are important problems with respect to data collection of bankrupt firms which deserve preliminary discussion. This matter was alluded to in the introduction. Realistic evaluation of a model's predictive relationships requires that the predictors are (would have been) available for use prior to the event of failure. Now, it is of course true that annual reports are not publicly available at the end of the fiscal year, since the financial statements must be audited. Previous studies have not mentioned this problem, at least not explicitly. This is not surprising since most previous studies have used *Moody's Manual* to derive the pertinent financial ratios, and the *Manual* does not indicate at what point in time the data were made available. Another reason is that not all studies have been concerned with strict forecasting relationships. That is, whether accounting statements were publicly available or not had no direct bearing upon the subject at hand. One such case is Beaver [1968*a*; 1968*b*], who studied whether financial ratios will reflect impending failure. The timing issue can be expected to be serious for firms which have a large probability of failure in the first place. Such firms are in poor shape and the auditing process could be particularly problematic and time-consuming. Thus, it is somewhat risky to assume that financial reports were available, say, three months after the end of the fiscal year. There are other disadvantages associated with *Moody's Manual.* The data are often highly condensed, and it is generally complicated, if not impossible, to reconstruct actual balance sheets and income statements.[5] Again, firms which are in poor shape are particularly difficult, since one can never be sure whether some of the many possible special items have been given special treatment in *Moody's* tabulation.[6] Moreover, it should be noted that the comparative schedules over the different years are ex-post reconstructions, and items from previous years may have been restated and may differ from the amounts originally reported. At a nontrivial cost, this problem can be circumvented if one uses several annual editions of *Moody's* for the same firm.

Clearly, much can be gained by improving the data base. The evaluation of the predictive classification power of a model should be more realistic, and, more important here, the same should apply for standard tests of statistical significance. This is not to suggest that it is important to have "super accurate" data for purposes of developing (as opposed to evaluating) a discriminatory device. It might well be that the predictive quality of any model is reasonably robust across a variety of data-gathering and estimating procedures.

[5] The conclusion is based on a few "case studies."

[6] The summaries of taxes (loss carry-forward, in particular) and measures of operating performance appear to be the most difficult items to deal with.

3. *Collection of Financial Statement Data*

The collection of data for bankrupt firms requires a definition of failure and specification of the population from which firms are drawn. In this study, the definition is purely legalistic. The failed firms must have filed for bankruptcy in the sense of Chapter X, Chapter XI, or some other notification indicating bankruptcy proceedings.[7] The population boundaries are restricted by: (*i*) the period from 1970 to 1976; (*ii*) the equity of the company had to have been traded on some stock exchange or over-the-counter (*OTC*) market; (*iii*) the company must be classified as an industrial. The first criterion was chosen simply because it is the most recent period, and the cutoff point of 1970 was selected as a matter of practicality. The second criterion excludes small or privately held corporations. This is crucial, since otherwise the use of *Compustat* firms as a source of nonbankrupt firms would be precluded. The third criterion excludes utilities, transportation companies, and financial services companies (banks, insurance, brokerage, REITs, etc.). Companies in these industries are structurally different, have a different bankruptcy environment, and appropriate data are, in some cases, difficult to obtain.

The following procedures were used to generate a list of failed firms satisfying the inclusion criteria. (1) A primary listing of failed firms was tabulated from the *Wall Street Journal Index* (*W.S.J.I.*). Type and date of bankruptcy were recorded. If the name of the company indicated that the firm in question was a nonindustrial, then it was excluded. (2) A secondary listing of firms was tabulated by excluding all firms on the primary listing which had not been traded on one of the stock exchanges (or *OTC*) during the three-year period prior to the date of bankruptcy. If a company could be traced to one of the exchanges, then the exchange was recorded. This kind of information was derived from various stock guides issued by *Moody's* and *Standard and Poor's*. Of course, as a practical matter, it was assumed that a stock had not been traded if no evidence could be found to that effect. (3) Attempts were made to augment the secondary listing by examining other miscellaneous sources of data. This led to some relatively minor additions to the listing of bankrupt firms. However, in tracing bankrupt firms it seemed to me that very few firms were omitted from the *W.S.J.I.*, so long as the firms satisfied the inclusion criteria.

The next phase was one of actually collecting financial data for the bankrupt firms. The objective was to obtain three years of data prior to the date of bankruptcy. Each report had to include the balance sheet, income statement, funds statement, and the accountants' report. In case the last available accountants' report explicitly stated that the company had filed for bankruptcy, then a fourth report was collected. All reports were retrieved from the Stanford University Business School Library,

[7] See Altman [1971] for a discussion of the difference between different types of bankruptcy proceedings.

TABLE 1

Bankrupt Firms: Year, Type of Bankruptcy, and Exchange Listing

Type	Year							Totals
	1970	1971	1972	1973	1974	1975	1976	
Chapter X	0	2	2	1	1	0	0	6
Chapter XI	1	4	14	20	18	14	14	85
Other or unknown	0	0	5	6	0	1	2	14
Totals	1	6	21	27	19	15	16	105
New York Stock Exchange								8
American Stock Exchange								43
Other*								54

* Over-the-counter market or regional exchange.

which has an extensive microfilm file of 10-K reports. The relevant parts of the 10-K reports were photocopied and subsequently coded. Some firms had to be deleted from the sample because no report whatsoever was available, but these were few. Other firms, again very few, were deleted because they were corporate shells and had no sales. On the other hand, no firm was deleted because of its young (exchange) age, and some firms had only one set of reports.

In the process of coding items from the annual reports, I noted that all but one firm which went bankrupt in 1970, and some of the 1971 firms, had no funds statement in their annual report. This was not true for firms which filed for bankruptcy in subsequent years. Similar observations are applicable for firms on the *Compustat* file, although omissions were much less frequent. The SEC did not require a funds statement until the early seventies.[8] I decided that firms without a funds statement should be deleted, since it would have been impossible otherwise to use ratios derived from the funds statement.

The final sample was made up of 105 bankrupt firms. Basic information regarding year, exchange, and type of bankruptcy is given in table 1. As one would suspect, relatively few firms were listed on the NYSE, compared to the other two categorizations. Furthermore, Chapter XI bankruptcy was apparently much more frequent than Chapter X.

I noted that while eighteen of the 105 firms (17 percent) had accountants' reports which disclosed that the company had entered bankruptcy, the fiscal year-end was prior to the date of bankruptcy. These reports were deleted and reports from the previous fiscal year were substituted. As a consequence, the average lead time between the date of the fiscal year of the last relevant report and bankruptcy is quite long, approximately thirteen months. Table 2 shows the entire frequency distribution. This lead time is quite a bit longer, compared to what has been reported in previous studies. A cursory review of the data indicated that the time lag between the fiscal year-end and the date of the accountants' report

[8] Funds statements have been required since September 30, 1971; see *APB opinion No. 19*.

TABLE 2

*Lead Times of Last Set of Annual Reports (Not Indicating Bankruptcy) Prior to Bankruptcy**

Months	3	3.5	4.5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5	13	13.5	14	14.5	15	15.5
Number of Reports	2	1	1	1	3	4	5	2	2	4	2	4	3	4	4	7	3	2	2	2	3	8	5	4
Months	16	16.5	17	17.5	18	18.5	19.5	20	20.5	21.5	23	23.5	25	26	27.5	28	33.5							
Number of Reports	4	4	3	1	1	1	1	1	2	1	1	1	1	2	1	1	1							

* Mean: 13 months; mode: 14.5 months; median: 12.5 months.

can be quite long. In fact, for the group of eighteen firms, the lag was always longer in the year which was closer to bankruptcy. The following example is reasonably representative. Hers Apparel Ind. filed for bankruptcy May 31, 1974; the accountants' report for the fiscal year-end February 28, 1974 is dated July 19, 1974. In the previous year, the report was dated April 24, 1973. Note that the lead time between fiscal year date of last "relevant" report and date of bankruptcy is approximately thirteen months in this case (i.e., April 24, 1973 to May 31, 1974). In 1974, it took approximately four and one-half months to complete the audit. (I found many cases which exceeded four and one-half months.) There were also a number of firms for which additional relevant reports could have existed. Under such circumstances, search procedures were attempted, but with little success. For most of these cases, it appears as if the firms simply never filed any additional reports with the S.E.C. This is by no means implausible, since firms can apply for extension of their deadline, and after bankruptcy has actually occurred there may simply be no point in going through an audit and preparing a standard annual report. Of course, it is also possible that additional reports did exist, but never got to the Stanford University Library. In order to play it "safe," I decided that no firm was to be deleted because reports were potentially missing. As a consequence, any evaluation of a model based on this data set probably understates the predictive classification performance.

A sample of nonbankrupt firms was obtained from the *Compustat* tape. Ideally, all reports for all firms satisfying the population constraints should have been included as a control sample. However, this was deemed to be too costly and impractical (due to core memory constraints). I decided instead that every firm on the *Compustat* tape (excluding utilities, etc.) should contribute with only one vector of data points; the year of any given firm's report was obtained by random procedure. This led to 2,058 vectors of data points for nonbankrupt firms. The breakdown into exchange listings was as follows: New York Stock Exchange = 42%, American Stock Exchange = 32%, Other = 26%.

4. *A Probabilistic Model of Bankruptcy*

Let $\mathbf{X}_i$ denote a vector of predictors for the ith observation; let $\boldsymbol{\beta}$ be a vector of unknown parameters, and let $P(\mathbf{X}_i, \boldsymbol{\beta})$ denote the probability of bankruptcy for any given $\mathbf{X}_i$ and $\boldsymbol{\beta}$. P is some probability function, $0 \leq P \leq 1$. The logarithm of the likelihood of any specific outcomes, as reflected by the binary sample space of bankruptcy versus nonbankruptcy, is then given by:

$$l(\boldsymbol{\beta}) \equiv \sum_{i \in S_1} \log P(\mathbf{X}_i, \boldsymbol{\beta}) + \sum_{i \in S_2} \log(1 - P(\mathbf{X}_i, \boldsymbol{\beta})),$$

where S_1 is the (index) set of bankrupt firms and S_2 is the set of nonbankrupt firms. For any specified function P, the maximum likelihood estimates of $\beta_1, \beta_2 \cdots$, are obtained by solving:

$$\max_{\beta} l(\boldsymbol{\beta}).$$

In the absence of a positive theory of bankruptcy, there is no easy solution to the problem of selecting an appropriate class of functions P. As a practical matter, all one can do is to choose on the basis of computational and interpretative simplicity. One such function is the logistic function:

$$P = (1 + \exp\{-y_i\}^{-1}), \qquad \text{where } y_i \equiv \sum_j \beta_j X_{ij} = \boldsymbol{\beta}'\mathbf{X}_i.$$

There are two implications which should be mentioned. First, P is increasing in y; second, y is equal to $\log[P/(1 - P)]$. The model is thus relatively easy to interpret, and this is its main (and perhaps only) virtue.[9]

5. *Ratios and Basic Results*

For purposes of the present report, no attempt was made to develop any "new or exotic" ratios. The criterion for choosing among different predictors was simplicity. The first three models estimated, Models 1–3, were composed of an intercept and the following nine independent variables:[10]

1. *SIZE* = log(total assets/GNP price-level index). The index assumes a base value of 100 for 1968. Total assets are as reported in dollars. The index year is as of the year prior to the year of the balance sheet date. The procedure assures a real-time implementation of the model. The log transform has an important implication. Suppose two firms, A and B, have a balance sheet date in the same year, then the sign of $P_A - P_B$ is independent of the price-level index. (This will not follow unless the log transform is applied.) The latter is, of course, a desirable property.
2. *TLTA* = Total liabilities divided by total assets.
3. *WCTA* = Working capital divided by total assets.
4. *CLCA* = Current liabilities divided by current assets.
5. *OENEG* = One if total liabilities exceeds total assets, zero otherwise.
6. *NITA* = Net income divided by total assets.
7. *FUTL* = Funds provided by operations divided by total liabilities.

[9] See McFadden [1973] for a comprehensive analysis of the logit model.

[10] No attempt was made to select predictors on the basis of rigorous theory. To put it mildly, the state of the art seems to preclude such an approach. (The first six predictors were partially selected simply because they appear to be the ones most frequently mentioned in the literature.)

8. *INTWO* = One if net income was negative for the last two years, zero otherwise.
9. *CHIN* = $(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$, where NI_t is net income for the most recent period. The denominator acts as a level indicator. The variable is thus intended to measure change in net income. (The measure appears to be due to McKibben [1972]).

Previous studies, "common sense," and perhaps even theory, would suggest that the sign of the coefficients of the different ratios should be as follows:

Positive	Negative	Indeterminate
TLTA	*SIZE*	*OENEG*
CLCA	*WCTA*	
INTWO	*NITA*	
	FUTL	
	CHIN	

OENEG serves as a discontinuity correction for *TLTA*. A corporation which has a negative book value is a special case. Survival would tend to depend upon many complicated factors, and the effect of the extreme leverage position needs to be corrected. A positive sign would suggest almost certain bankruptcy, while a negative sign suggests that the situation is very bad indeed (due to *TLTA*), but not *that* bad. (Granted, this is a very heuristic procedure to capture something very complicated.)

A "profile" analysis of the data supports the hypotheses regarding the signs. Table 3 shows the means and standard deviations of the predictors for three sets of data: one year prior to bankruptcy, nonbankruptcy firms, and two years prior to bankruptcy. The results are hardly surprising. The ratios deteriorate as one moves from nonbankrupt firms to two years prior to bankruptcy to one year prior to bankruptcy. Although the data

TABLE 3
Profile Analysis

	One Year Prior to Bankruptcy		Nonbankrupt Firms		Two Years Prior to Bankruptcy	
Variable	*mean*	*stdv*	*mean*	*stdv*	*mean*	*stdv*
SIZE	12.134	1.38	13.26	1.570	12.234	1.414
TLTA	0.905	0.637	0.488	0.181	0.718	0.311
WCTA	0.041	0.608	0.310	0.182	0.157	0.320
CLCA	1.32	2.52	0.525	0.740	0.814	0.671
NITA	−0.208	0.411	0.0526	0.0756	−0.052	0.155
FUTL	−0.117	0.421	0.2806	0.360	−0.0096	0.332
INTWO	0.390	0.488	0.0432	0.2034	0.180	0.384
OENEG	0.18	0.385	0.0044	0.0660	0.060	0.237
CHIN	−0.322	0.644	0.0379	0.458	0.00308	0.8673
N	105		2,058		100	

and ratios are not quite comparable with those of Beaver [1966], the results here are quite similar to the profiles he presented [1966, p. 82]. It should also be noted that the standard deviations of the predictors (except for size) are larger for year-1 firms, compared to nonbankrupt firms. These differences are significant at a 5-percent level or better. Hence, as discussed in Section 2, standard assumptions of *MDA* are unlikely to be valid.

Three sets of estimates were computed for the logit model using the predictors previously described. Model 1 predicts bankruptcy within one year; Model 2 predicts bankruptcy within two years, given that the company did not fail within the subsequent year; Model 3 predicts bankruptcy within one or two years. A summary of the results are shown in table 4. This table indicates that all of the signs were as predicted for Model 1. Only three of the coefficients (*WCTA*, *CLCA*, and *INTWO*) have t-statistics less than two, so the others are all statistically significant at a respectable level. This includes *SIZE*, which has a relatively large t-statistic. An overall measure of goodness-of-fit is given by the likelihood ratio index. The index is similar to a R^2 in the sense that it equals one in case of a perfect fit, and zero if the estimated coefficients are zero.[11] For Model 1, the ratio is 84 percent, and this is significant at an extremely low α-level. The statistic "Percent Correctly Predicted" equals 96.12 percent; it is tabulated on the basis of a cutoff point of .5. That is, classify the company if and only if $P(\mathbf{X}_i, \hat{\boldsymbol{\beta}}) > 0.5$. Whether this is a "good" or "bad" result is not easy to answer at this stage, so further discussion regarding the model's predictive power is postponed until the next section. At this point, we can note that if all firms were classified as nonbankrupt, then 91.15 percent would be correctly classified (2,058/(105 + 2,058)). Thus the marginal (unconditional, prior) probability of bankruptcy is an important quantity in the above type of statistic. Further, there is no apparent reason why .5 is an appropriate cutoff point, since it presumes implicitly that the loss function is symmetric across the two types of classification errors.

Table 5 shows the correlation coefficients of the estimation errors in Model 1. The coefficients of the financial state variables (variables 1–4 in the table) are uncorrelated with those of the performance variables (variables 5–9). Hence, both sets of variables contribute significantly and independently of each other to the likelihood function. This strongly supports the contention that both sets of variables are important in establishing the predictive relationship.

Models 2 and 3 have somewhat weaker goodness-of-fit statistics, which

[11] The likelihood ratio index is defined to be:

$$1 - \text{log likelihood at convergence/log likelihood at zero.}$$

The index will take on the value of one in case of a perfect fit, since log likelihood at convergence then equals zero. If there is no fit, then obviously the index equals zero. See McFadden [1973] for further details.

TABLE 4

Prediction Results

	Variable									
	SIZE	*TLTA*	*WCTA*	*CLCA*	*NITA*	*FUTL*	*INTWO*	*OENEG*	*CHIN*	*CONST*
Model 1										
Estimates	−.407	6.03	−1.43	.0757	−2.37	−1.83	0.285	−1.72	−.521	−1.32
t-statistics	−3.78	6.61	−1.89	.761	−1.85	−2.36	.812	−2.450	−2.21	−.970
Model 2										
Estimates	−.519	4.76	−1.71	−.297	−2.74	−2.18	−.780	−1.98	.4218	1.84
t-statistics	−5.34	5.46	−1.78	−.733	−1.80	−2.73	−1.92	−2.42	2.10	1.38
Model 3										
Estimates	−.478	5.29	−.990	0.062	−4.62	−2.25	−.521	−1.91	.212	1.13
t-statistics	−6.23	7.72	−1.74	.738	−3.60	−3.42	−1.73	−3.11	1.30	1.15

	Likelihood Ratio Index	Percent Correctly Predicted
Model 1	0.8388	96.12
Model 2	0.7970	95.55
Model 3	0.719	92.84

TABLE 5

Correlation Matrix of (Estimated) Estimation Errors in Model 1

	SIZE	*TLTA*	*WCTA*	*CLCA*	*OENEG*	*NITA*	*FUTL*	*INTWO*	*CHIN*
SIZE	1	−.28	*	*	*	*	*	*	*
TLTA		1	.32	*	−.49	*	*	*	*
WCTA			1	.46	*	*	*	*	*
CLCA				1	*	*	*	*	*
OENEG					1	*	*	*	*
NITA						1	−.41	.40	−.44
FUTL							1	*	*
INTWO								1	−.32
CHIN									1

* = Absolute value of coefficient less than .20.

is exactly what one would expect in view of the profile analysis. Note also that the signs of *INTWO* and *CHIN* differ from those of Model 1. The positive and significant coefficient for *CHIN* in Model 2 can perhaps be explained by a scenario proposed by Deakin [1972]. Firms with a positive change in earnings may be particularly tempted to raise external capital through borrowing, and this will then imply that they become higher-risk firms at a subsequent point. Of course, this is only one possible explanation and the evidence is far too weak even vaguely to suggest that it is in fact the case.

In all three models, size appears as an important predictor. This finding is consistent with Horrigan's [1968] study of bond ratings, wherein he too found that size was an important determinant. One could perhaps argue that the conclusion is invalid because the bankrupt and nonbankrupt firms are drawn from different populations. Specifically, one cannot be sure that all the nonbankrupt firms would have been on the *Compustat* tape if they had not failed.[12] No direct test of the problem is therefore feasible. The *Compustat* tape is heavily biased toward the relatively large firms listed on the two major exchanges. If size is a spurious variable, then it is likely that dummy variables reflecting exchange listings are more important than size. This test was implemented, and the results are shown in table 6. The t-statistic for *SIZE* is larger than two, whereas the two exchange dummies NYSE and AMSE are essentially insignificant. These results again support the contention that size is an important predictor of bankruptcy. Even so, the test and conclusion must be viewed with great caution, since *Compustat* firms are different from non-*Compustat* firms on a number of complex dimensions. Therefore, we cannot be sure that size is a surrogate which is superior to exchange listing. If size, in fact, is a superior surrogate, then the statistical significance of size may simply reflect a general *Compustat* bias.[13]

[12] The *Compustat* file does not include all firms which have (or have had) their equity traded.

[13] It would have been preferable to obtain data for nonbankrupt firms by sampling 10–K reports from the Stanford Library. However, this would have been a very costly procedure.

TABLE 6
Model 4

Variable	Estimates	*t*-Statistics
SIZE	−.267	−2.02
TLTA	5.63	6.04
WCTA	−1.43	−1.91
CLCA	.0585	.595
NITA	−2.35	−1.82
FUTL	−1.99	−2.53
INTWO	.307	.877
OENEG	−1.56	−2.20
CHIN	−.5092	−2.15
NYSE	−.854	−1.71
AMSW	−.0513	−.186
CONST	−2.63	−1.70
Percent Correctly Predicted		96.30%
Likelihood Ratio Index		.8399

Subject to the qualification above, the results indicate that the four factors derived from financial statements which are statistically significant for purposes of assessing the probability of bankruptcy are: (*i*) size (*SIZE*); (*ii*) the financial structure as reflected by a measure of leverage (*TLTA*); (*iii*) some performance measure or combination of performance measures (*NITA* and/or *FUTL*); (*iv*) some measure(s) of current liquidity (*WCTA* or *WCTA* and *CLCA* jointly).[14] In an attempt to determine whether other factors can be obtained from financial statements which could increase the predictability of failure, I estimated an additional model. This model was Model 1 supplemented by a measure of profit margin, computed as funds from operations divided by sales, and a ratio of assets with little or no cash value (intangibles plus deferred assets) divided by total assets. The estimation results showed not only that both of these variables were completely insignificant, but also that the estimated coefficients had "incorrect" signs. (The *t*-statistics were .14 and −.42, respectively.) Whether other accounting predictors could have done a better job in improving upon the likelihood function is not clear, but, in my view, it is not overly likely.[15] On the other hand, nonaccounting data, such as information based on equity prices and changes in prices might prove to be most useful. I intend to test these kinds of models once adequate data have been gathered.

[14] If *CLCA* is deleted in the equation estimated, then the *t*-statistic of *WCTA* is slightly greater than two. For all practical purposes, this will be the only difference.

[15] One possible exception would be measures of "trends" of ratios and/or volatility of ratios (and performance measures). Due to the considerable costs of gathering and organizing data, such a refinement was judged to be cost-benefit inefficient.

6. *Evaluation of Predictive Performance*

There is no way one can completely order the predictive power of a set of models used for predictive (decision) purposes. As a minimum, this requires a complete specification of the decision problem, including a preference structure defined over the appropriate state-space. Previous work in the area of bankruptcy prediction has generally been based on two highly specific and restrictive assumptions when predictive performance is evaluated. First, a (mis)classification matrix is assumed to be an adequate partition of the payoff structure. Second, the two types of classification errors have an additive property, and the "best" model is one which minimizes the sums of percentage errors. Both of these assumptions are arbitrary, although it must be admitted that the first assumption is of some value if one is to describe at least one implication of using a model. Much of this discussion will therefore focus on such a (mis)classification description. Nevertheless, the second assumption will also be used at some points, since it would otherwise be impossible to compare the results here with those of previous studies. The comparison cannot be across models because the time periods, predictors, and data sets are different. Rather, the question of interest is one of finding to what extent the results conform with each other.

Without loss of generality, one may regard the estimated probability of failure, $P(\mathbf{X}_i, \hat{\boldsymbol{\beta}})$, as a signal which classifies firm i into one of the two groups. Hence, it is of interest to describe the conditional distributions of these signals. Figure 1 shows the empirical frequency of $P(\mathbf{X}_i, \hat{\boldsymbol{\beta}})$ for the 105 firms which failed within a year; $\hat{\boldsymbol{\beta}}$ is the vector of estimated coefficients obtained from Model 1. Figure 2 shows the frequency for the 2,058 nonfailed firms where $\hat{\boldsymbol{\beta}}$ is again taken from Model 1. Figure 3 shows the probabilities for 100 firms two periods prior to bankruptcy; the β's

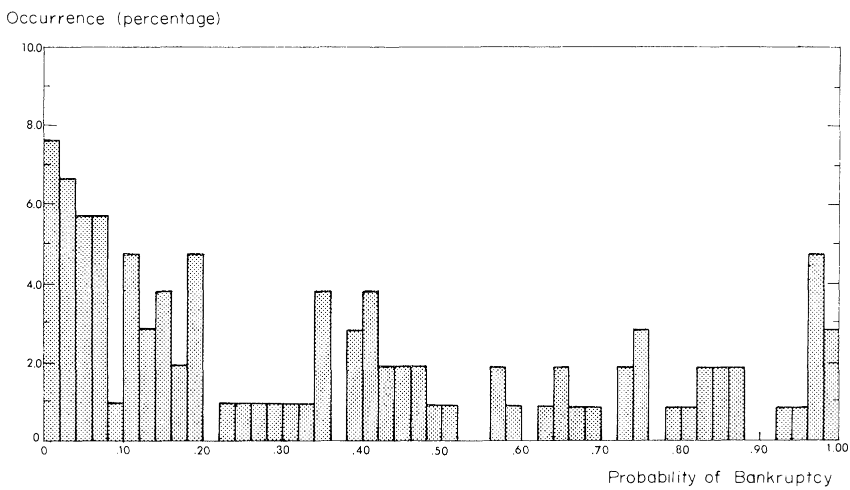

FIG. 1.—Firms one year prior to bankruptcy (105 firms)

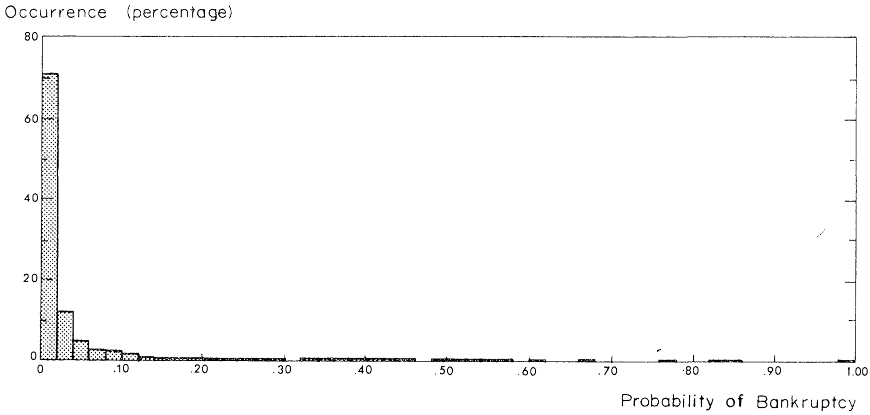

FIG. 2.—Nonbankrupt firms (2,058 firms)

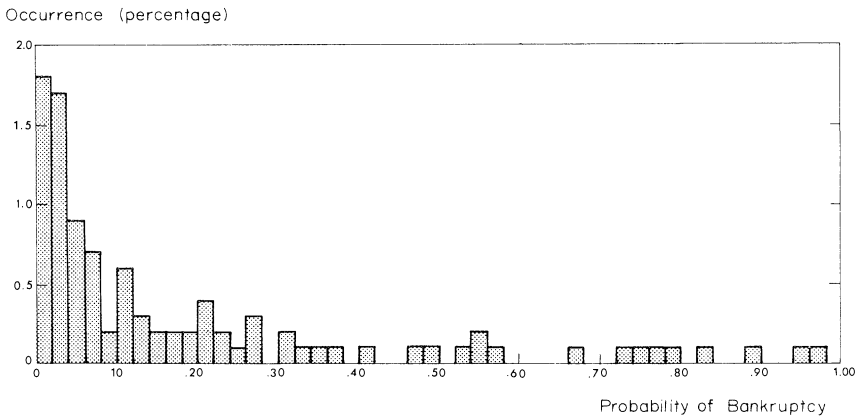

FIG. 3.—Firms two years prior to bankruptcy (100 firms)

were taken from Model 2. The mean probabilities are .39, .03, and .20, respectively. This is, of course, in accordance with what one would expect on the basis of prior reasoning.

Using the data which underlie figures 1–3, one can readily perform analysis of classification errors for different cutoff points. The focus will be on a prediction of bankruptcy within one year. A Type-I error will be said to occur if $P(\mathbf{X}_i, \hat{\boldsymbol{\beta}})$ is greater than the cutoff point and the firm is nonbankrupt; in a similar fashion, one defines a Type-II error for bankrupt firms if the probability is less than the cutoff point. It would have been preferable to perform the error analysis on a "fresh" data set and thereby (in)validate the models estimated.[16] Due to the lack of data beyond 1976, this was not possible at the time of the study. This should not be a serious problem, however, for the following reasons. First, I have

[16] It would, of course, have been possible to cut the sample in half and then go through the usual kind of procedures. However, the primary purpose of this paper is not one of getting a precise evaluation of a predictive model. Hence, it was decided that the full sample would be used in order to produce the smallest possible errors of the estimated coefficients.

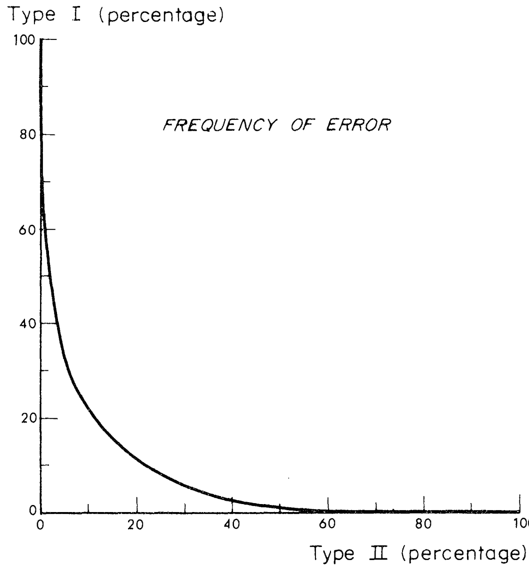

FIG. 4.

not indulged in any "data dredging" and no attempt was made to find a "best" model or even a model which is "superior" to Model 1. Second, a logit analysis is not an econometric method which is designed to find an "optimal" frontier, trading off one type of error against another. This is in contrast to *MDA* models which satisfy optimality conditions under appropriate assumptions. Third, as it turns out, the sum of the percentage of errors is relatively robust across a wide range of cutoff points. Finally, the sample size is relatively large, so the estimates should not be too sensitive to the particular sample used.[17]

Figure 4 depicts the frontier trading of one error against another, when the errors are expressed as percentages. Figure 5 shows the mapping from cutoff points to the two different types of errors. The cutoff point which minimizes the sum of errors is .038. At that point, 17.4 percent of the nonbankrupt firms and 12.4 percent of the 105 bankrupt firms are misclassified. Given an (infinite) population, in which half of the firms are failed and the other half nonfailed, the expected error rate would then be 14.9 percent, provided that the cutoff point is .038, and provided that the distribution of the predictors is representative of the population. I might also note that if one selects a cutoff point equal to .0095, then no Type-II error occurs and Type-I errors equal .47.

It is not easy to compare the results above with those reported by others. First, the lead time from last fiscal year-end to the date of

[17] The points just made will be invalid in a real-world application of the model if the β-parameters are significantly different in future periods. This is a fairly obvious observation which has been made by others (see, e.g., Moyer [1977]).

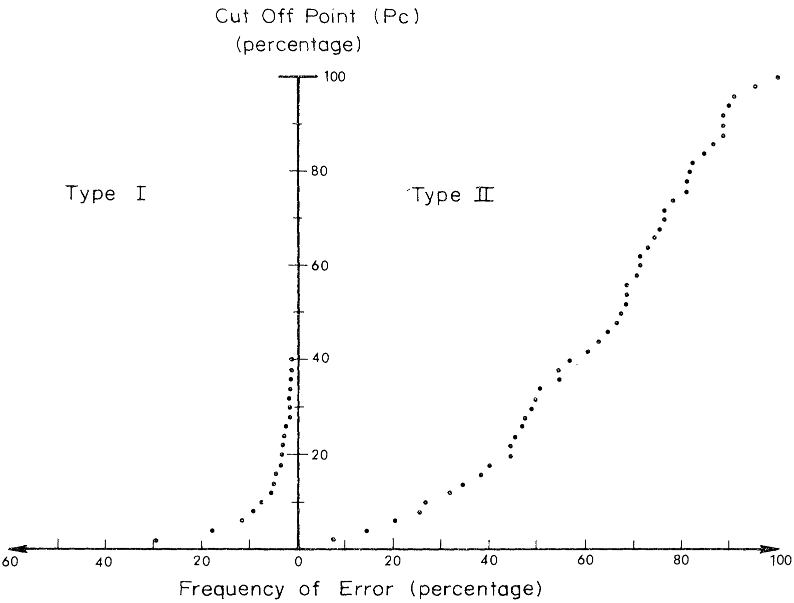

FIG. 5.

bankruptcy is much longer in the present study. Second, most studies, except for the ones by Altman and McGough [1974], Moyer [1977], and Altman et al. [1977], have not considered data from the seventies. Third, studies have differed in their selection of predictors. However, it is hard to believe that the latter is of great significance once a fair number of "basic" ratios and predictors are incorporated. One potentially important exception to the latter is the use of nonaccounting-based data such as market-price data. Such data have not been used here. Fourth, it is possible that the predictive results could be sensitive to choice of estimation procedures (e.g., the logit model versus *MDA*).

A review of previous studies indicates that most of these reported error rates which are less than those given here, with several studies reporting rates in the vicinity of 5 percent. Some of the potential sources that may account for this differential are considered below.

The first point mentioned (the lead time) can be investigated to some extent. If it is true that this is an important consideration, then the average lead time should be larger for misclassified bankrupt firms, compared to correctly classified bankrupt firms. The difference is approximately 1.75 months, so the factor may be of some importance. The difference is significant at the 20-percent level (approximately). Indeed, three of the misclassified firms would have been "correctly" classified if I had used data from the subsequent annual report (i.e., these firms belong to the group of eighteen firms discussed in a previous section). Under these circumstances the minimum of the average error is somewhat less than 13 percent, the cutoff point is somewhat larger, and both types of errors are reduced. Hence, the lead time accounts for some of the differences. But it is hard to believe that this is the entire explanation.

As previously indicated, evidence exists regarding the second issue. Altman and McGough [1974] applied the model which Altman [1968] developed in 1968 to twenty-eight firms that failed during the period 1970–73. The predictors were computed from data one year prior to bankruptcy; the definition of "one year prior to bankruptcy" appears to be completely equivalent to that of the present study. In the 1968 study Altman reported a misclassification rate of approximately 5 percent for failed firms. The same model applied to the second group of firms yielded a misclassification rate of 18 percent (5/28). This is a substantial and significant increase, since the probability of getting five or more misclassifications out of twenty-eight observations is approximately one in one thousand if the true parameter is 5 percent. The study lacks a systematic analysis of the other type of error, so the evidence is not conclusive. More important, Altman et al. [1977] reworked much of what Altman did in 1968 using data from 1969–75 (94 percent of the total sample was from this period). The model development included a number of refinements in the utilization of discriminant analysis, as well as in the computation of financial ratios. The authors report: "... bankruptcy classification accuracy ranges from over 96 (93% holdout) one period prior to bankruptcy" (Altman et al. [1977, p. 50]). Needless to say, such results are not in accordance with those of the present study or, for that matter, the two other studies which used data from the seventies. I am unable to account for this difference. Unfortunately, Altman et al. [1977] do not report on the average lead time, so it is impossible to evaluate the importance of this factor. Also, they did not apply (or report) the predictive performance of their recent *ZETA* model on the 1974 sample. However, the authors do seem to suggest some sample dependence (see Altman et al. [1977, n. 16]). There are also differences in the definition of bankruptcy.

In sum, differences in results are most difficult to reconcile. Moyer [1977] recently reexamined the Altman model using data from the 1965–75 period. (The Altman [1968] sample was from the 1946–65 period.) The error rate reported by Moyer for the Altman model was no less than 25 percent! Reestimation of the parameters of the Altman model (using 1965–75 data and using the estimated model to classify firms) yielded an error rate of 10 percent. The latter result must be qualified (downward bias) due to the small sample size (twenty-seven bankrupt and twenty-seven nonbankrupt firms). The data were derived from *Moody's Manual* leading to additional downward bias.

Would the use of other predictors have affected the predictive power? The question cannot really be answered unless one tries out a large number of models and thereafter, necessarily, cross-validates the "best" model. The results from the model which added two predictors to Model 1 are not very encouraging, so at this point I must be quite skeptical. As previously suggested, a significant improvement in goodness-of-fit is more likely to occur by augmenting the accounting-based data with market-price data.

At this point, I want to emphasize that the reports of the misclassified bankrupt firms seem to lack any "warning signals" of impeding bankruptcy. All but two of the thirteen companies reported a profit. The two losses were relatively minor (*NITA* was −0.022 and −0.044, respectively), and these two companies had strong financial positions (*TLTA* was .23 and .37, respectively). The median *TLTA* ratio is .55, and the range is .23–.70. The median *NITA* is 3.4 percent with a range of −0.44 to 0.156. (The firm which had *TLTA* = .70 had *NITA* = .156.) Other ratios analyzed showed the same "healthy" patterns. It is not surprising that these firms were misclassified, especially if one considers the profile of the nonbankrupt firms shown in table 3.

Moreover, the accountants' reports would have been of little, if any, use. None of the misclassified bankrupt firms had a "going-concern qualification" or disclaimer of opinion. A review of the opinions revealed that eleven of these companies had completely clean opinions, and the two that did not had relatively minor uncertainty exceptions. Curiously, some of the firms even paid dividends in the year prior to bankruptcy. Hence, if any warning signals were present, it is not clear what these actually were.[18]

There is always the possibility that an alternative estimating technique, other than the logit model used, could yield a more powerful discriminatory device. Unfortunately, a priori reasoning appears to be of no use in finding such an "optimal" estimating technique. All one can do is to try some alternatives. One approach I tried, *MDA*, produced results which were somewhat "worse" than those previously reported, in that the minimum average error rate was 16 percent. More generally, I would hypothesize that many "reasonable" procedures will lead to results which will not differ too much. This robustness property can be illustrated as follows. If we use the estimates from Model 2 for the purpose of predicting bankruptcy within one year, the β-estimates from Model 2 will be evaluated in terms of their predictive power with respect to firms one year prior to bankruptcy and the 2,058 *Compustat* firms. Again, different cutoff points yield a trade-off between the two types of errors. Table 7 displays the two types of errors at selected cutoff points for Models 1 and 2. Interestingly enough, if a cutoff point of .08 is selected for Model 2, then the average error is 14.4 percent, and this is slightly better than the minimum attained by Model 1. Model 2 performs better at some other

[18] Ratios other than those used in the estimating equations were also examined. For all of the misclassified firms, I was unable to detect any ratio which was clearly out of line. However, it is quite possible that a time-series analysis of an extended period would indicate that some of the firms had "significant" above-average operating risks. By the use of market data, this problem will be investigated in the future.

TABLE 7

*Type I–Type II Analysis for Selected Cutoff Points**

Estimates from: Cutoff Point	Model 1		Model 2	
	Type I**	Type II	Type I	Type II
0.0	100%	0%	100%	0%
0.02	28.7	7.6	54.3	0%
0.04	16.7	14.3	37.7	0.95
0.06	11.8	20.0	26.8	4.76
0.08	9.3	25.7	20.2	8.6
0.10	7.2	26.7	17.0	12.4
0.20	3.3	44.8	7.2	31.4
0.30	1.75	48.6	3.6	43.8
0.40	1.07	57.1	2.0	50.5
0.42	0.92	61.0	1.75	51.4
0.50	0.63	67.6	1.07	57.1
0.54	0.44	68.6	0.82	61.0
0.60	0.29	71.4	0.68	62.9
0.70	0.19	76.2	0.49	70.5
0.80	0.15	81.9	0.24	74.3
0.90	0.049	88.6	0.19	82.9
1.00	0	1.00	0	1.00

* Data sets: nonbankrupt firms and one year prior to bankruptcy.
** Type I: predict bankruptcy; actual nonbankrupt.

points too; that is, for a fixed level of one type of error for both models, the complementary error is lower for Model 2. A close examination of table 7 will verify this. To be sure, for some error rates Model 1 is better than Model 2. It seems reasonable to suggest that the models are essentially equivalent as predictive tools.

7. Conclusions

There are two conclusions which should be restated. First, the predictive power of any model depends upon when the information (financial report) is assumed to be available. Some previous studies have not been careful in this regard. Second, the predictive powers of linear transforms of a vector of ratios seem to be robust across (large sample) estimation procedures. Hence, more than anything else, significant improvement probably requires additional predictors.

REFERENCES

ALTMAN, E. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance* (September 1968).

———. *Corporate Bankruptcy in America.* Lexington, Mass.: Heath Lexington, 1971.

———. "Predicting Railroad Bankruptcies in America." *Bell Journal of Economics and Management Science* (Spring 1973).

———, R. HALDEMAN, AND P. NARAYANAN. "ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations." *Journal of Banking and Finance* (June 1977).

———, AND B. LORRIS. "A Financial Early Warning System for Over-the-Counter Broker Dealers." *Journal of Finance* (September 1976).

———, AND T. MCGOUGH. "Evaluation of a Company as a Going Concern." *Journal of Accountancy* (December 1974).

BEAVER, W. "Financial Ratios as Predictors of Failure." *Empirical Research in Accounting: Selected Studies, 1966.* Supplement to *Journal of Accounting Research* 4.

———. "Alternative Financial Ratios as Predictions of Failure." *The Accounting Review* (January 1968*a*).

———. "Market Prices, Financial Ratios and the Prediction of Failure." *Journal of Accounting Research* (Autumn 1968*b*).

BLUM, M. "Failing Company Discriminant Analysis." *Journal of Accounting Research* (Spring 1974).

DEAKIN, E. B. "A Discriminant Analysis of Predictors of Business Failure." *Journal of Accounting Research* (Spring 1972).

EDMISTER, R. O. "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction." *Journal of Financial and Quantitative Analysis* (March 1972).

EISENBEIS, R. A. "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics." *Journal of Finance* (June 1977).

HORRIGAN, J. "The Determination of Long-Term Credit Standing with Financial Ratios." *Empirical Research in Accounting: Selected Studies, 1968.* Supplement to *Journal of Accounting Research* 6.

JOY, M. O., AND J. O. TOLLEFSON. "On the Financial Applications of Discriminant Analysis." *Journal of Financial and Quantitative Analysis* (December 1975).

LEV, B. "Financial Failure and Informational Decomposition Measures." In *Accounting in Perspective Contributions to Accounting Thoughts by Other Disciplines*, edited by R. R. Sterling and W. F. Bentz, pp. 102–11. Cincinnati: Southwestern Publishing Co., 1971.

LIBBY, R. "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence." *Journal of Accounting Research* (Spring 1975).

MCFADDEN, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by P. Zarembka. New York: Academic Press, 1973.

MCKIBBEN, W. "Econometric Forecasting of Common Stock Investment Returns: A New Methodology Using Fundamental Operating Data." *Journal of Finance* (May 1972).

MOYER, R. "Forecasting Financial Failure: A Re-Examination." *Financial Management* (Spring 1977).

SANTOMERO, A., AND J. D. VINSO. "Estimating the Probability of Failure for Commercial Banks and the Banking System." *Journal of Banking and Finance* (September 1977).

WHITE, R. W., AND M. TURNBULL. "The Probability of Bankruptcy: American Railroads." Working paper, Institute of Finance and Accounting, London University Graduate School of Business, February 1975*a*.

———. "The Probability of Bankruptcy for American Industrial Firms." Working paper, July 1975*b*.

WILCOX, J. "A Prediction of Business Failure Using Accounting Data." *Empirical Research in Accounting: Selected Studies, 1973.* Supplement to *Journal of Accounting Research* 11.