0

RAND

Statistical Forecasting of Bankruptcy of Defense Contractors

Problems and Prospects

Anthony G. Bower, Steven Garber

Project AIR FORCE

The research reported here was sponsored by the United States Air Force under Contract F49620-91-C-0003. Further information may be obtained from the Strategic Planning Division, Directorate of Plans, Hq USAF.

Library of Congress Cataloging in Publication Data

Bower, Anthony G.
Statistical forecasting of bankruptcy of defense contractors : problems and prospects / Anthony G. Bower, Steven Garber.
p. cm.
"MR-410-AF."
"Prepared for the U.S. Air Force."
Includes bibliographical references.
ISBN 0-8330-1524-9
1. Defense industries—United States—Finance—Forecasting. 2. Disarmament—Economic aspects—United States. 3. Bankruptcy—United States. I. Garber, Steven, 1950– . II. RAND. III. Title.
HD9743.U6B63 1994
338.4'36233'0973—dc20 94-5047
CIP

Published 1994 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Internet: order@rand.org.

RAND

Statistical Forecasting of Bankruptcy of Defense Contractors

Problems and Prospects

Anthony G. Bower, Steven Garber

Prepared for the United States Air Force

Project AIR FORCE

Preface

This monograph report presents the research results on the question of whether a credible objective model can be developed to predict defense contractor bankruptcy. This research was conducted in the Acquisition Project within the Resource Management and Systems Acquisition Program of RAND's Project AIR FORCE, the Air Force's federally funded research and development center (FFRDC) for studies and analyses.

The specific study supporting this research has been "Management Challenges in the New Procurement Environment: Financial Distress Among Defense Contractors and the Costs of Limited-Production Developments." This study was co-sponsored by Mr. LeRoy T. Baseman, Deputy Assistant Secretary for Cost and Economics, SAF/FMC, and by Major General Stephen P. Condon, Deputy Assistant Secretary for Management Policy and Program Integration, SAF/AQX. This report should be of interest to financial professionals within the military who have direct responsibility for reporting on contractor financial viability, as well as to others within the acquisition community whose decisions are affected by the possibility of contractor bankruptcy.

Contents

Figures

Summary

This report considers whether a model can be found or developed that would reliably predict whether a defense contractor will go bankrupt.[1] The term "model" refers to a mechanical procedure in which data about the firm are used as inputs, and a prediction regarding bankruptcy is generated. This is in contrast to subjective "frameworks," in which the user is expected to employ judgment, given a set of relevant variables.

There is a substantial literature presenting models to assess the bankruptcy prospects of firms. The early attempts were based on multivariate discriminant analysis (MDA), in which a sample of firms, some (usually half) known to have gone bankrupt, is chosen. Financial data are used to construct variables characterizing each firm in the sample. MDA procedures are used to choose the variables (and their weights) that minimize a weighted number of classification errors of the two types (i.e., classifying a nonbankrupt firm as bankrupt and vice versa). Graphically, this corresponds to choosing a line (or plane) in variable space that best discriminates between bankrupt and nonbankrupt companies. This method is commonly used; one version is referred to as the "Z-score." It provides a classification scheme but does not typically generate an implied probability of bankruptcy. Many more recent papers employ techniques known as "logit" or "probit," which are multiple linear regression techniques adapted to situations in which the dependent variable is either a zero or a one. These techniques provide an estimated probability for each firm.

In nearly all of these papers, the authors begin with a long list of candidate variables, most of which are financial ratios computed from the firm's most recent balance sheet. In some, variables are added that relate to the firm's securities, such as the ratio of market value to total assets. Usually the list of variables is narrowed down to a manageable number (typically, five to ten) on the basis of goodness of fit. In the statistics community, such a process is often (pejoratively) termed "data mining" or a "fishing expedition." The result of such a process is virtually certain to be an equation whose goodness of fit is deceptively high; it is a misleading indicator of the predictive reliability of the equation.

[1]The analysis here is applicable to all defense contractors, not just USAF contractors. In this report, USAF and DoD are used interchangeably.

Since most such models are derived through data mining, it is not surprising that the "in sample" fit tends to be very good. When authors leave part of the potential estimation sample out during the estimation process (i.e., use a "hold-out sample"), the predictive performance for these firms is still reasonable, but not as good. When extended to "out-of-sample" prediction (i.e., different time periods or industries from those characterizing the estimation sample), however, the predictive reliability of these models falls dramatically.

The business environment of defense firms is markedly different from those firms, usually a sample of manufacturing or retail firms, for which the models in the literature have been developed. The existence of progress payments, "GOCO" plants (Government Owned, Contractor Operated), a monopsonistic buyer, and the significance of the product for the security of the nation, among other factors, creates a very different business environment. Loosely, defense contractors are "far out of sample," and the models in the literature cannot be expected to predict reliably when applied to defense contractors.

Furthermore, the relevant problem for the Air Force is to predict bankruptcy over the next decade or so—a period when the defense industry is expected to suffer a prolonged and deep drawdown. One cannot expect the models in the literature to predict well, since most of the observations are for firms within industries that are not suffering a severe general decline.

Could the general methodology be applied, using an appropriate sample of defense firms to fit an equation? Clearly, the new equation must apply to defense firms during the current drawdown. One might begin by collecting data on companies that are more relevant for defense and for the current circumstances. For example, firms included might be defense firms during other drawdowns.

However, this approach is not promising for at least four reasons: First, with the demise of the Soviet threat, the security environment has fundamentally changed. Second, the size of the drawdown is large and does not appear to be reversible. Third, the defense business has changed substantially, partly because of the enormous level of resources required for a new system and partly because of the drawdown, which means that fewer rescuers are available for failing firms. The above considerations make drawing inferences about bankruptcy from other drawdowns perilous. Fourth, the small number of bankruptcies among publicly traded defense firms (i.e., those for which financial data are available) makes the construction of an adequate sample problematic. Even studies that consider the entire manufacturing sector are based on a relatively small number of companies.

Thus, the Air Force should avoid statistical approaches that use historical financial data. Might there be a more promising approach for predicting bankruptcy? The market value of the securities of the firm do reflect the firm's business prospects, as well as the market's perceptions concerning intangibles (such as the quality of management, the soundness of the strategic plan, relationships with the clients, etc.), which determine whether capital markets are likely to come to the rescue of the firm if bankruptcy looms. An alternative approach, therefore, is to use securities-market information to attempt to estimate the market's perception of the likelihood that the firm will go bankrupt.

Such an approach seems conceptually feasible, at least in the case of a simple bond (e.g., a bond that either pays interest and principal or defaults). The risk premium associated with a simple bond should relate directly to the perceived risk of default. The literature applying this idea is limited and disappointing, however. The leading study attempts to examine the relationship between default risk and bond prices. When used to solve for the implied probability of bankruptcy, it leads to implausibly pessimistic predictions. This inaccuracy is related to a well-known phenomenon, examined by Michael Milken and others: the market appears to undervalue low-grade bonds; there is an apparent "junk-bond premium." Thus inferring the probability of bankruptcy from the value of a firm's bonds has not produced sensible results, for reasons that are not completely understood.

Even if the current methodology is improved and plausible probabilities could be derived from securities prices, another difficulty would remain. If the Air Force wishes to predict the probability of a bankruptcy in the context of a source selection, then it is interested in a conditional probability (i.e., what is the probability of bankruptcy, *given* that the contract is awarded to the firm?). But securities prices reflect unconditional probabilities (e.g., probability of bankruptcy, given the market's beliefs about the likelihood that a firm will be awarded a contract). The unconditional probability puts a bound on the conditional one, but to compute a relevant probability for a source-selection decision, the Air Force would have to take account of how likely the market thinks it is that the firm in question will be awarded the contract. Analysis suggests that the Air Force's inference about the probability of bankruptcy would be highly sensitive to such guesses.

We conclude that subjective methods of pro forma analysis and expert financial judgment are currently more promising than any mechanical model that could be devised for evaluating the probability that a defense contractor will go bankrupt in the current environment. However, there is some probability of building a reliable objective model. Our review provides guidance for those attempting to develop more useful objective forecasting tools. We offer the following recommendations to those who might undertake this task:

- If sufficient bankruptcies are available historically, future empirical effort should concentrate on estimation of models using data only from defense contractors.
- Methods that may be quite unreliable for estimation of bankruptcy probabilities might be much more reliable for *ranking* companies according to these probabilities. For some DoD decisions, being able to rank companies may be very useful.
- Bankruptcy is fundamentally a sequential or interactive process. Companies first fall into a vulnerable position from which poor performance can push them over the edge. Use of statistical methods that reflect such characteristics warrant serious consideration in future efforts.
- Regarding choice of predictor variables, it seems very worthwhile to consider use of bond yields (or ratings) in conjunction with accounting data. Use of book values of assets should be avoided. In addition, it could be very useful—albeit very challenging—to develop variables to reflect interdependencies among defense contractors. At least two sources of interdependency warrant attention: those between competitors and those between prime contractors and their subcontractors.

Acknowledgments

We thank Lieutenant Colonel Terry Raney, Chairman of the Contractor Finance Department of the Defense Systems Management College, for his very helpful advice in the early stages of this research. We also wish to thank Mr. Walt Hosey, USAF, for his helpful comments in the initial stages of this research and Major Clayton Chun and Major Tim Sebring (both of SAF/FMC) for their comments on the interim briefing that we gave in Washington. John Adams and Bridger Mitchell of RAND provided very useful technical reviews. Ed Keating, Michael Kennedy, Ellen Pint, Rachel Schmidt, and especially Dennis Smallwood of RAND also contributed useful insights and comments.

Acronyms

CAPM	Capital Asset Pricing Model
cdf	cumulative distribution function
GOCO	Government Owned, Contractor Operated
MDA	Multivariate Discriminate Analysis
MLA	Multinomial Logit Analysis
RP	Recursive Partitioning
YTM	Yield to Maturity

1. Introduction

The Problem

The defense industry has experienced declining revenues (in real terms) since the mid-1980s, and the defense drawdown is expected to continue for the foreseeable future. This situation will have severe financial effects on many aerospace and other USAF contractors. Some of these effects have already manifested themselves in the form of layoffs and plant closures. Repercussions such as bankruptcies can be expected in the future.

Deterioration of the financial condition of a defense contractor—i.e., an increase in the likelihood that it will not be able to meet its future financial obligations—may threaten its performance of military contracts. A contractor in a weak financial position—which we refer to generally as being in a state of "financial distress"—may not have the resources to absorb large cost overruns, and its managers may become preoccupied with short-term financial concerns.

This report discusses various issues relevant to improving USAF's capability to predict financial distress of its contractors. As detailed below, the form of financial distress focused on here is (Chapter 11) bankruptcy. The goal is to forecast bankruptcy prospects of a very specific population: defense contractors operating in the current drawdown. This population has several salient features. First, the defense business has some unique characteristics that will affect prediction. Second, defense contractors operate within a declining industry. Third, the appropriate forecasting horizon for Air Force purposes may be quite short (say, 3 months) or quite long (in the case of a source selection, 10 years or more). Such features of the population are relevant to the prediction task and are discussed in more detail.

Contract performance may be threatened by forms of financial distress less severe than bankruptcy. Firms that reduce or eliminate their dividends or default on a bond or loan may be expected to have difficulty performing on their contracts because of increased borrowing costs, a shortage of working capital, loss of key personnel, or distractions to management. In addition, financial distress less severe than bankruptcy is a warning sign of future bankruptcy. It is useful for the USAF to be able to anticipate such developments—at least to be able to anticipate problems before they arise and, perhaps, to support more proactive strategies.

Definitions of Bankruptcy

Of the many types of bankruptcy only two affect corporations of the size considered here. Chapter 7 bankruptcy (liquidation) is more prevalent among small businesses. Chapter 11 (reorganization) is the most common among major corporations. In Chapter 11, the company is shielded temporarily from its creditors, and creditors and shareholders attempt to reach a reorganization agreement to avoid liquidation. Often, a portion of the firms' debts are forgiven by creditors in exchange for stock. Also, bankruptcy laws allow the firm to shed unprofitable contracts (e.g., labor agreements). Frequently, the firm changes management. If the attempted reorganization fails, then the firm will enter Chapter 7 liquidation. Almost all firms of the size of prime defense contractors "successfully" emerge from Chapter 11, in the sense that an agreement is reached among creditors, shareholders, and management, and the company is not liquidated. Chapter 11 is our focus here.

Two Strategies of Bankruptcy Prediction

There are two general strategies of bankruptcy prediction. One involves "mechanical" or "objective" methods in which a model or equation is developed from historical data and then applied to companies of interest. The second strategy involves "subjective" methods employing quantitative and qualitative information (e.g., discussions with analysts, assessments of access to capital markets) and combining it into an assessment of future prospects. We focus on the mechanical approach, partly because the Air Force has focused previous efforts on the more judgmental or subjective methods.

The mechanical approach has several potential advantages: it is seemingly objective and therefore, perhaps, more defensible to outsiders; and it may be more accurate because it incorporates past data systematically. Also, the Air Force might be interested in assessing the prospects of many companies, particularly if subcontractors are of concern. A mechanical approach then would be desirable because a model could be applied to many companies in a short period of time. The criteria for model evaluation are accuracy, reliability, and supportability to outside observers.

Our main questions are: What are the prospects for building a useful model to forecast defense contractor bankruptcy in the current drawdown? What improvements over existing methods seem most likely to enhance these prospects?

This report consists of six sections. Section 2 provides a synopsis of an extensive survey of a literature presenting equations that purport to predict bankruptcy for various types of firms. In Section 3, we assess the applicability of existing equations to the defense context and, in Section 4, the possibility of adapting the existing methods. In Section 5, we consider fundamentally different approaches relying on securities prices. Conclusions are offered in Section 6, including a summary assessment of the current state of technology for the forecasting task and the prospects for improvement.

2. Review of the Literature

Statistical Studies of Financial Distress

Statistical studies comprise almost all the published work on bankruptcy prediction. Since these are directly relevant for the Air Force, we carefully examined this literature. Much of this section is a fairly detailed critique of this work. The critique highlights issues bearing directly on whether it is promising for the USAF to use equations from this literature and, if not, whether this general methodology can be adapted to the defense context.

We reviewed roughly 50 studies. Most have been published in accounting or finance journals in the last 25 years, and almost all focus on bankruptcy of manufacturing or retail firms operating within more or less healthy industries. Prediction time horizons are generally one to seven years; that is, data available during a particular year are used to predict bankruptcy as many as seven years in the future.

Samples of Companies

Bankruptcy is a quite rare event; less than 1 percent of firms fail in any particular year.[1] The rarity of bankruptcy has two important implications for our purposes. First, samples are generally chosen by identifying firms that declared bankruptcy during a particular time period and combining them with either an equal number of somehow-matched firms or a large number of more or less randomly selected firms that (apparently) did not fail.[2] When the matching strategy is used, generally relatively small sample sizes result because business failures are

[1] Zmijewski (1984, fn. 1) reports that the rate of business failure in the United States had not exceeded 0.75 percent in any year since 1934. The rate of bankruptcy among defense contractors is unknown.

[2] Zmijewski (1984, Table 1, p. 61) provides an overview of the sampling strategy used in 17 earlier studies; in 12 of the 17, the matching strategy was used. He also discusses and analyzes statistical biases involved in using this approach to sample construction and another potential source of bias. The first type of bias—a consequence of so-called "choice-based sampling"—results from over sampling failed firms relative to their prevalence in the population of all firms; this example of sampling nonrandomly on the dependent variable generally leads to biases in estimation. The second type of bias pertains to the inclusion of firms only if complete data are available. Zmijewski finds that both these features of the sampling process lead to biases in the estimated coefficients of the prediction equations—compromising their usefulness in assessing the roles of various determinants of bankruptcy—but do not undermine predictive performance.

relatively rare.[3] Second, studies have not been able to focus on firms in a single industry. Most often, the sample includes firms from any of the many industries in the manufacturing sector. Some studies have also considered retail companies.

Review of Key Statistical Techniques

The statistical models are largely atheoretical—they are posited in an ad hoc manner involving searches for correlations of bankruptcy outcomes with variables viewed as plausible predictors. Various statistical techniques have been used. Most studies employ multivariate techniques—i.e., use a prediction model involving several variables—but an early researcher explored the possibility of predicting bankruptcy using a single variable.

Specifically, Beaver (1966, 1968) studied the ability of various financial ratios or market variables used singly to predict bankruptcy. He shows that even a single variable can be a significant help in prediction, and market prices (e.g., rate of return on the company's stock) predict failure slightly earlier than financial ratios (e.g., net income divided by total assets, total debt divided by total assets).

Since Altman (1968), the literature on bankruptcy prediction has focused on better-performing multivariate methods[4]—using multiple predictor variables jointly or in a sequential fashion. By far, the most common techniques are multivariate discriminate analysis (MDA) and logit analysis.[5]

With either method, the basic idea is to use historical data to estimate a relationship between a bankruptcy outcome and a set of predictor variables and use the resulting equation (or model) to forecast bankruptcy for a set of companies in the future. The first step is to specify a population of companies to be studied and a time period during which bankruptcy is to be analyzed. For example, one may choose to examine manufacturing companies and whether they went bankrupt during any of the years 1986, 1987, 1988, and 1989 (the "estimation period"). The second step is to select a sample of companies including some that did and some that did not go bankrupt during the selected time period. Third, one collects data on the candidate predictor variables during

[3]For the 12 studies using the matching strategy reviewed by Zmijewski (1984, Table 1, p. 61), the number of bankrupt firms ranges from 20 to 115.

[4]Cf. Scott (1981, p. 321).

[5]For example, Altman, Haldeman, and Narayanan (1977) use MDA, and Ohlson (1980) and Mensah (1984) use logit analysis. Probit, which is very similar to logit, is used by Zmijewski (1984).

a time period before the bankruptcy prediction period.[6] Next statistical methods are applied to the data (observed predictor variables and subsequent bankruptcy outcomes) to estimate coefficients (or weights) for the predictor variables. Finally, the estimated coefficients are then applied to observed values of the predictor variables for a company to calculate the value of an index that is used to predict bankruptcy.[7]

MDA is a classification scheme: the prediction is in the form of an assignment of a company to a group in which companies are classified as bankrupt or a group in which companies are classified as nonbankrupt. Statistical methods are applied to an estimation sample to choose the weights for the index (called a discriminant score). Companies are then classified as bankrupt if and only if their score—calculated using the values of their predictor variables and the estimated weights—falls on a particular side of some critical value.[8] This critical value is chosen to minimize the expected costs of misclassification, which depend on the costs and probabilities of the two potential types of misclassification: (1) classifying a nonbankrupt firm as bankrupt and (2) classifying a bankrupt firm as nonbankrupt. The appropriate critical value is context specific; most important in this regard may be that the relative costs of the two types of misclassification are likely to be very different for different kinds of decisions that depend on the classification.

For the sake of exposition, assume that two financial ratios are used as predictor variables: net income divided by total assets and debt divided by total assets. Other things equal, we might expect that increases in the former reduce the likelihood of bankruptcy and increases in the latter increase this likelihood.

Geometrically, with two predictor variables the MDA estimation procedure identifies the line (or, if there are more than two predictor variables, the plane or hyperplane) that best separates the observations for which bankruptcy was and was not observed. See Figure 1. The slope of the fitted line is determined by the estimated weights, and its vertical position is determined by the selected critical value for the discriminant score. Every firm whose financial variables place it above the fitted line (i.e., whose discriminant score is above the critical score) is classified as bankrupt, and every one below is classified as nonbankrupt. The

[6]For example, a researcher might observe financial ratios for 1985 to relate to the bankruptcy outcome in 1986, 1987, 1988, or 1989. Separate analyses might be performed for the different outcomes such as whether the company went bankrupt within a year (after 1985), two years, etc.

[7]As discussed below, the predictions might be made for firms not used in estimation but for the years used in estimation, for the companies used in estimation for years after the estimation period, or for both companies and time periods not used in estimation.

[8]See, for example, Altman, Haldeman, and Narayanan (1977) for a discussion of the choice of cut-off or critical values for the discriminant score in the context of bankruptcy prediction.

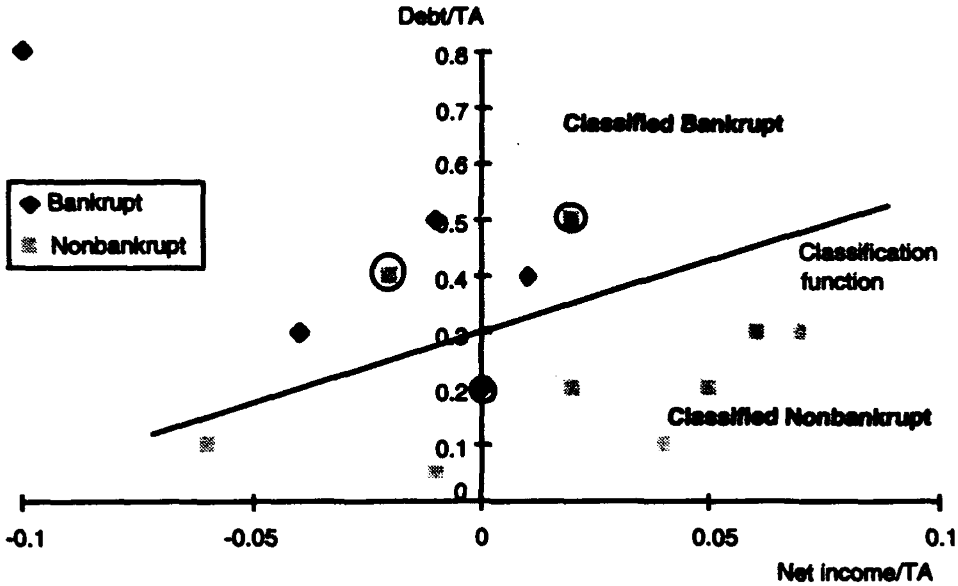

Figure 1—Multivariate Discriminant Analysis with Two Classification Variables

critical discriminant score implicit in the figure reflects a particular value for the relative costs of the two classification errors. If instead the cost of classifying a bankrupt firm as not bankrupt were even larger relative to the cost of the other type of error, then the classification line would be lower to avoid more misclassifications of the former type.

The misclassified observations are circled in Figure 1. One measure of performance of the model (classification scheme) is how well the line separates the two groups of observations used to fit the model, in particular, the percentage of sample observations that are correctly classified. This *in-sample* classification accuracy is reported for all the studies that use MDA.

Logit (or probit) involves statistical concepts similar to linear regression. With logit (see Figure 2), the prediction—the probability of bankruptcy—is a (nonlinear) function of an estimated linear combination of a vector of input variables x, using a functional form constraining these probabilities to lie between 0 and 1. Cumulative distribution functions (cdf's) have these properties. Logit and probit use the logistic and standard normal cdf's, respectively; both of these cdf's have the sigmoid shape illustrated.

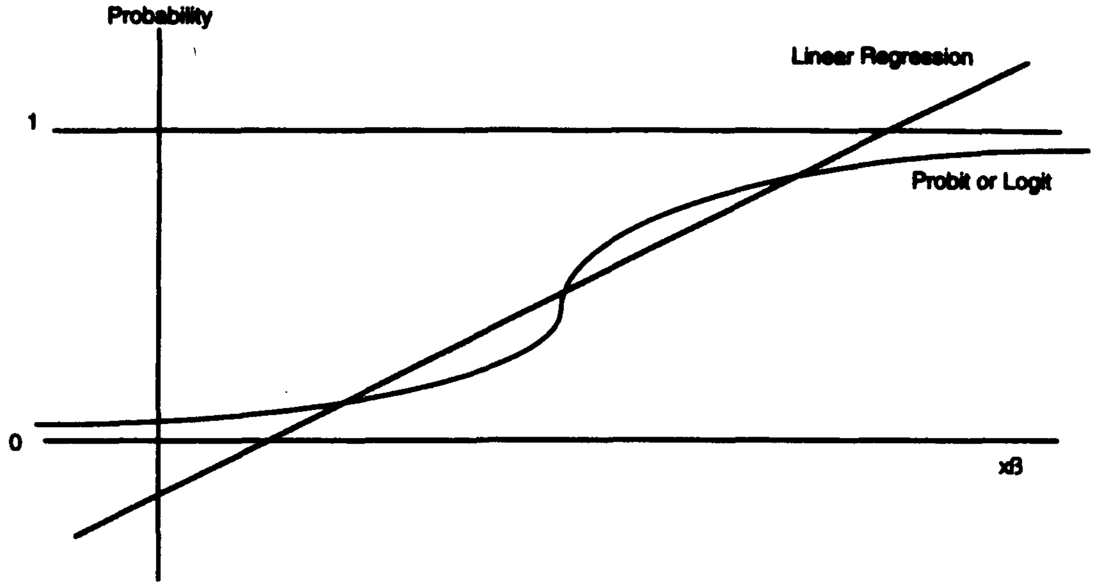

Figure 2—Probit and Logit Are Forms of Regression

Logit (or probit) analysis has an important theoretical advantage over MDA. While the standard output of an MDA is a set of classifications or predictions (i.e., either bankrupt or not), the output of a logit analysis is a probability of bankruptcy. The latter seems more useful in many situations because it varies continuously and is not dependent on a particular ratio of misclassification costs.[9] While we doubt that the choice of statistical technique has major practical consequences compared with a number of the factors discussed below, we favor logit (or probit) analysis over MDA because we find predicted probabilities more useful.[10]

In both MDA and logit, then, a set of predictor variables is used jointly and simultaneously, and variables that can vary continuously have a continuous effect on the index. An alternative statistical framework that shows promise for predicting bankruptcy—recursive partitioning (RP)—uses predictor variables in

[9] In addition, researchers who use logit often emphasize that MDA invokes the assumption that the predictor (x) variables are multinormally distributed, while logit analysis does not involve any assumptions about the distributions of these variables. Such an assumption is clearly violated for some types of variables, but it is not clear that this issue is important in practice. (See for example, Haggstrom [1983].)

[10] We are aware of only one study comparing the relative success of these techniques in predicting bankruptcy: Collins and Green (1982). They conclude that logit analysis does perform better than MDA but not dramatically so. They question whether the advantages of logit compensate for the additional computational complexity relative to linear regression. We tend to view the extra computational burden as rather minor (and much less of an issue than when Collins and Green [1982] did their analysis) because logit and probit routines are now widely available.

a sequential (recursive) fashion and involves dichotomizing of the variables at each step. This approach was proposed and illustrated by Frydman, Altman, and Kao (1985).[11] In their study, RP outperformed MDA in most settings.

Choice of Explanatory Variables

When faced with choosing predictor variables to characterize firms, the basic problem is an embarrassment of riches. Conceptual considerations provide general guidance concerning the types of variables that are likely to be useful for prediction, but there are many potential measures of each concept suggested by a priori considerations.[12] Two basic strategies are employed to select particular variable definitions and sets of variables to be used for prediction purposes. Many, and perhaps most, studies undertake elaborate computerized searches over combinations of predictor variables from a prespecified set of candidates and use statistical criteria to select variables for classification or prediction purposes.[13] Such search techniques are known informally (and pejoratively) as "data mining" or "fishing expeditions." The other strategy is to specify a set of predictor variables on a priori grounds and simply use that specification.[14] This approach often uses the results of previous studies to create a list of 5 to 10 variables.

---

[11]Section I of Frydman, Altman, and Kao (1985) provides a very useful overview of this method.

[12]For example, Altman, Haldeman, and Narayanan (1977) consider (in addition to other types of variables) six different measures of profitability, seven different measures of performance relative to financial obligations, four different measures of other financial means relative to financial obligations, and four different measures of equity value normalized in various ways. (See especially their p. 32 and Appendix B.) A rather extreme case is Rose and Giroux (1984) who (p. 3) consider "... 27 financial ratios most consistently used in previous studies, as well as an additional 130 ratios ..." Mensah (1984) takes a different approach to reducing the number of predictor variables: he starts with 38 ratios and calculates ten factor scores—which can be thought of as estimates of ten unobservable variables that tend to explain the structure of covariation among the original 38 variables—and uses these factor scores as predictor variables in a bankruptcy equation.

[13]For example, Altman, Haldeman, and Narayanan (1977) and Rose and Giroux (1984). Researchers are generally not explicit about precisely what statistical procedure is used to select variables, but it seems that most use some sort of a "forward-selection" procedure. First the single variable is chosen that seems most important or significant—for example, the one with the highest correlation with the bankruptcy outcome or the largest t-statistic in a univariate analysis—from among the set of candidates. This variable is thereby determined to be one of those to be used in the final model, and a search is made from among the remaining variables for the one that seems most significant or important given that the first variable is already in the model. This second variable is thereby selected to be part of the final specification. This strategy is applied repeatedly until no remaining candidate variable satisfies a criterion for inclusion given the previously selected variables.

[14]For example, Ohlson (1980). He explains his choice of predictor variables by (fn. 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.)"

The Dangers of Data Mining

As innocent as data mining may appear, it involves some well-known statistical pitfalls. Given the variables we can measure and use in the model, the bankruptcy event is only partially predictable; i.e., it can be thought of as having a predictable and an unpredictable component. The best forecasts can be expected to result from an estimated relationship that is an accurate representation of the predictable component—the part of the bankruptcy determining process that is stable from sample to sample (i.e., stable within the population). But sifting through many candidate predictor variables inevitably leads to an estimated relationship that reflects the random component of the relationship for the estimation sample (i.e., the idiosyncrasies of the sample or the features of the sample that are not stable across the population). This danger is often referred to as "overfitting." By having the model include variables that do not correlate with (or predict) bankruptcy in the population but (by chance) correlate with bankruptcy in the sample, forecasting accuracy is undermined.

As an example of the dangers of including variables with no plausible a priori role, consider the mortality of U.S. presidents over the last 150 years. It turns out that whether the President was elected in a year divisible by 20 is highly correlated with death in office—and a data-mining procedure could pick this up. However, few would take seriously that statistic as a *predictor* of future mortality.

The importance for DoD of overfit models is that forecasts out of sample (the only relevant forecasts for DoD) will probably do worse than need be. We can convey some intuition about why data mining is hazardous for prediction in Figure 3.

As an example, consider a time series of data on a variable y. The straight line is intended to represent the predictable component of the series, and the best forecasts would then result from a straight-line extrapolation (or use of a linear-trend specification). A much better sample fit[15] can be obtained with a complicated form, such as the somewhat complicated function illustrated in the figure. But absent any theory to suggest *why* the predictable component of the time series might look so complicated, it is probably best to fit a simpler form (such as a straight line). Although the complicated function inevitably fits the sample data better (since its form is more general), it is prone to picking up idiosyncrasies associated with that particular sample. In this context it may well be possible to extrapolate a long-term trend successfully outside the sample (i.e.,

[15]In the context of Figure 3, the measure of fit might be an R-squared statistic. In the context of bankruptcy prediction, it would be in-sample classification or forecasting accuracy.

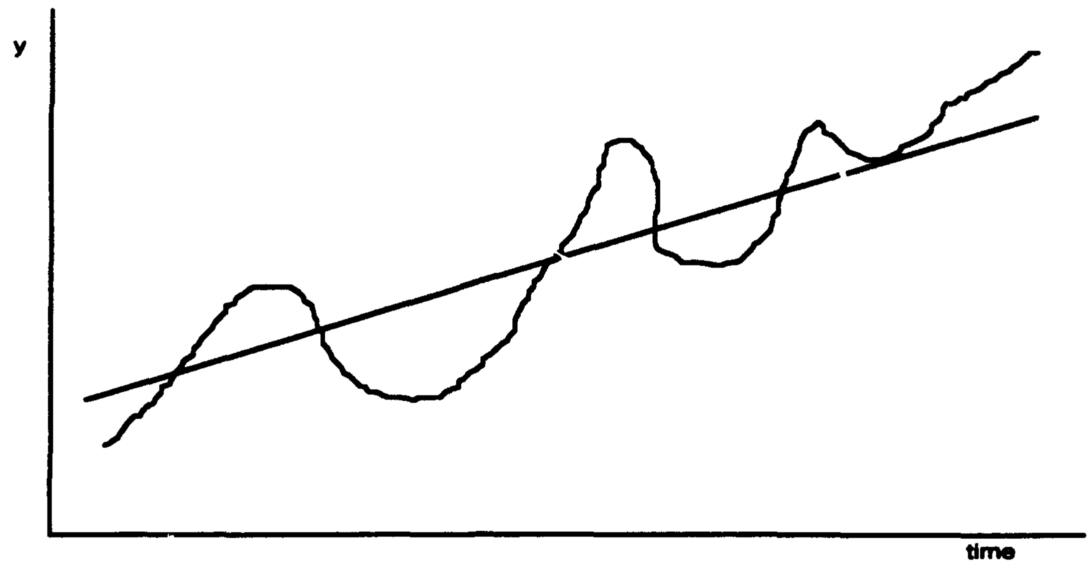

Figure 3—Dangers of Overfitting

into the future) but not the detailed period-to-period fluctuations. Attempting to do the latter can lead to worse predictions: The great is often the enemy of the good.

Variables Selected in the Literature

Studies typically settle on about 5 to 10 predictor variables. These variables capture some, but perhaps not all, of the fundamental determinants of bankruptcy. The studies generally include variables relating to financial vulnerability or recent performance. Broadly, the variables used in the literature can be interpreted as representing financial obligations, or recent financial performance, or a combination (generally a ratio) of the two.[16]

The financial obligations emphasized in the literature are debt and interest payments. Measures of short-term performance prominent in the literature include earnings, cash flow, and sales. Finally, other financial capabilities have been represented empirically by measures of stocks of liquid assets (e.g., working capital, dividend policy, and asset disposability). Performance variables are often normalized by some measure of firm size (total assets, capital, or sales) to construct a measure that is comparable across firms. Often another variable capturing performance relative to financial obligations (cash flow divided by

[16]The general interpretation of the literature provided here is based on our reading of the numerous studies cited in the references; the examples of specific variables are from Altman et al. (1977), which presents the widely-used "ZETA model" and is among the most influential studies of bankruptcy prediction.

debt or earnings divided by interest charges) is used as well. These variables are constructed using annual data from sources such as financial statements, balance sheets, and Forms 10K filed with the Securities and Exchange Commission. Predictions based on such data may be updated only once a year and may await months of delay for audit and release of the company's annual report.[17]

An important issue not addressed in the literature we reviewed is changes in predictive performance from substituting one measure of a particular concept (such as liquidity or leverage) for another. We suspect for two reasons that predictive performance is generally not very sensitive to the choice among various measures of the same concept or combination of concepts. First, alternative measures tend to be very highly correlated with each other. Second, different measures of the same concepts are selected in different studies; thus, the different choices across studies may be more a result of data mining than the superiority of one measure or another.

It seems that any reasonably simple model (e.g., six or so predictors) that incorporates one measure of each key concept—chosen perhaps based on considerations of data availability, cost, and timeliness—would seem to be about as promising as any other. This modeling strategy was used by Ohlson (1980); his model was also used by Burgstahler, Jiambalvo, and Noreen (1989). Differences in equations developed from computer-intensive procedures on one sample of firms are likely to reflect idiosyncratic features of those samples, not stable features of the population that can be exploited for forecasting purposes.[18] Efforts to develop better predictive tools for present purposes would more usefully focus on more substantive issues suggested by the USAF context and the uses to which the tools are to be applied.

Operationally, almost all empirical studies of "financial distress" equate this with bankruptcy. Only one recent study has built and evaluated a statistical model distinguishing more than two forms of financial health. The study, while carefully done, however, could not distinguish more finely than the two states of "distressed" and "not distressed."[19]

---

[17]See, for example, Ohlson (1980, pp. 115–117) and Lau (1987).

[18]Scott (1981, p. 325) suggests that multivariate models do not perform substantially better than univariate ones because the multivariate ones tend to be more complicated than is consistent with the goal of accurate prediction. This is a manifestation of the overfitting problem.

[19]Lau (1987) classifies sample firms into five states of financial health. In decreasing order of health, these are stability, omitting or reducing a dividend, technical default or default on a loan payment, bankruptcy protection, and bankruptcy and liquidation. (Studies of the traditional sort would classify companies in the first three states as "not bankrupt" and companies in the last two as "bankrupt.") She uses a multinomial logit analysis (MLA) to estimate how ten preselected predictor variables affect the probabilities of a company being observed in each of these states. Lau's (1987, p. 137) summary conclusion emphasizes "... the feasibility of constructing MLA models that

All the analyses ignore the perceptions of the financial community regarding crucial intangibles such as management quality and the strategic plan of the company. As a result, it is hard to predict critical things, such as whether the firm has access to capital markets. For example, poor or even negative cash flow may or may not be indicative of financial distress, depending on perceptions of a firm's future business prospects—consider, for example, start-up companies that are still developing their first products. Thus, various relevant factors are not reflected in the financial data on which many of these studies rely exclusively.

Model Performance Experiments and Performance Measures

The objective is to develop a model that will accurately forecast future bankruptcy outcomes. The most obvious way to gauge model performance is to make forecasts about future outcomes and wait to observe the extent to which the forecasts are borne out. This approach, of course, is impractical because we want to act on our forecasts before the event to be forecast does or does not occur. To gauge forecast accuracy, researchers attempt to approximate a forecasting exercise in various ways that allows more immediate evaluation of success. We refer to such exercises as forecasting *experiments*.[20]

As discussed presently, in the literature, models are evaluated (by the analysts who develop them) using various types of experiments. For any type of experiment, model performance is generally measured by a success rate—i.e., percent of observations correctly classified. The crucial issue is the nature of the experiment used to approximate the forecasting objective for which the model is to be used. Some experiments used in the literature are much more likely to be informative about forecasting accuracy than others.

The most common type of experiment evaluates the extent to which the model *fits* (or reproduces) the bankruptcy outcomes for the observations that were used

compute the probability that a firm will enter each of five different financial states." But inspection of the estimated coefficients in Table 4 suggests a very important conclusion about an issue that is not addressed. In particular, the results suggest very strongly that default without bankruptcy is much more like bankruptcy (in terms of its relationship to the predictor variables) than it is like the other states typically combined with a state of default under the classification "not bankrupt." This issue cannot be examined directly without her data, but it is suggested very strongly from the observation that the coefficients in Table 4 hardly vary across the first two rows (coefficients for the first two states—stability and reducing dividend) or within the last three rows (coefficients for the last three states—default, bankruptcy, liquidation). Put simply, it appears that the predictor variables discriminate between the first two states jointly and the last three jointly but not among states within these two groups.

[20] Analysts generally examine the performance of only the analyst's preferred model. We know of only two studies—Kaplan (1979) and Zmijewski (1984)—in which the predictive performance of the favored model is compared with that of models using different sets of predictor variables.

to estimate the model. Fit, however, is not a reliable guide to forecast accuracy. For example, as discussed above, a model that fits well because it reproduces the idiosyncrasies of a sample may be expected to forecast rather poorly.

Much more revealing about forecast accuracy is the ability of a model (estimated equation) to predict observations *not* used to estimate the model: forecasting is fundamentally about outcomes that have yet to occur (and hence cannot be used in estimation). Thus experiments that involve forecasting for observations not used to estimate the model are likely to be more revealing about forecast accuracy. Some types of experiments involving observations not used in estimation, however, are more useful than others depending on the nature of the forecasting task for which the model is being evaluated.

Regarding observations not used in estimation, is it useful to distinguish between the use of a "hold-out sample" and "out-of-sample prediction." A hold-out sample refers to a set of observations that might naturally have been used as part of the estimation sample—because they are for the same type of companies and same time periods as the estimation sample—but were not used in estimation (i.e., were "held out" from the estimation procedure) specifically so that they could be used to probe model performance. In contrast, "out-of-sample" prediction experiments involve the use of the estimated model to predict observations generated by environments that are different in some fundamental way from those generating the observations used to estimate the model.

For example, out-of-sample observations could be from fundamentally different types of time periods (e.g., from different phases of the business cycle) or companies (e.g., from different industrial sectors) from those reflected in the estimation sample. As developed below, in interpreting the literature on bankruptcy prediction, out-of-sample prediction accuracy is most relevant for present purposes because the environments determining the bankruptcy prospects of defense contractors are different in various ways from those generating observations used to estimate models in the literature.

Model Performance

The models in the literature generally *fit* the estimation sample data very well (partly because variables are often selected by data mining). The models generally also do well in classifying hold-out observations. However, the models perform rather poorly for out-of-sample predictions.

Figure 4 combines results from eight studies published from 1966 to 1984.[21] The figure presents average (over studies) bankruptcy prediction success rates for forecast horizons of one, two, and three years. This information is presented for two types of samples: companies used to estimate the model and hold-out companies.[22] First, note that the two sets of percentages are similar to each other: prediction accuracy for hold-out samples are quite similar to corresponding measures for observations used in estimation. Second, note that performance deteriorates as the forecast horizon is lengthened—classification is correct three years in advance for only 78 percent of the estimation sample companies, as compared to more than 90 percent accuracy for a forecast horizon of one year. As a baseline for comparison, the pure-chance classification accuracy is about 50 percent.

The business cycle is one reason that predictive performance deteriorates as the forecast horizon lengthens. Mensah (1984) seems to provide the most detailed examination of predictive performance over time. He emphasizes that predictive accuracy is much better for models estimated only for recessionary periods than for periods subsuming both recessions and recoveries. Examination of his Table

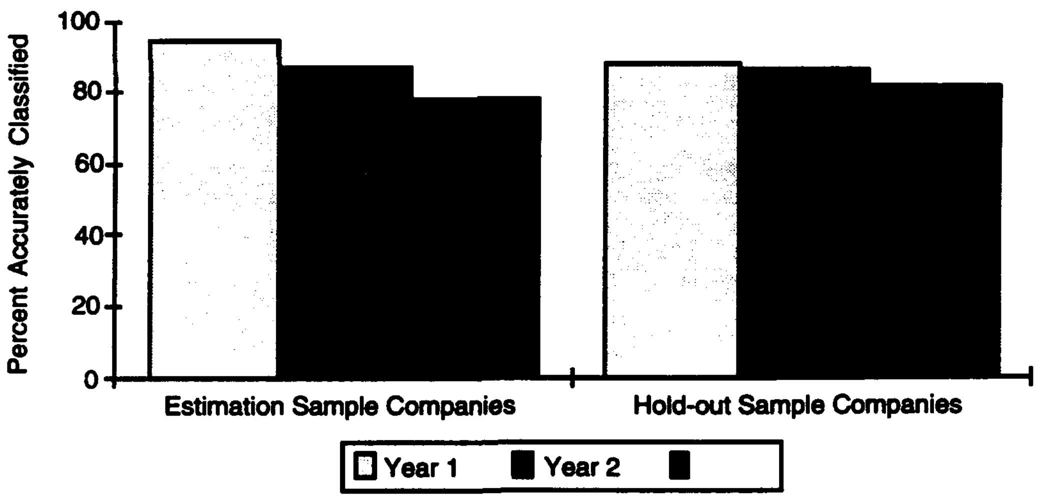

Figure 4—Performance Deteriorates with Lengthening Forecast Horizon

[21]Figure 4 is adapted from Lau (1987, Table 6), which gives citations for the eight studies. Several of the averages reported in Figure 4 are based on less than eight studies because not every study considered all the issues summarized in the figure. For example, some studies did not examine three-year forecast horizons, and three studies did not involve hold-out samples. The results averaged are broadly consistent across the relevant studies.

[22]The nature of the hold-out sample differs across studies with regard to the relationship between the estimation sample years and the predictor variable years. See Lau (1987, p. 136) and the studies cited.

3 establishes a very telling point: across time periods corresponding to different phases of the business cycle, variables can have (statistically significant) coefficients of opposite signs.

Now consider the issue of prediction for years later than those used to estimate the model. For example, a model might be estimated using financial data for 1980–1981 and bankruptcy outcomes for 1981–1982, and predictions about bankruptcy could be made for 1983–1984. In the literature, predictions are typically made from estimates using predictor variables observed up to 1982, rather than (for example) predicting bankruptcy in 1988 using the estimated equation and predictor variables observed up to 1986. Scott (1981, p. 320) cites Altman (1968) as an exception, writing that in this case: "... the model is revalidated over time based on observations after the model was built." Note that this latter exercise mimics more closely the practical use to which any predictive model is to be put: use of a model estimated with historical data to predict events that have yet to occur using information available at the time of prediction. Unfortunately, this exercise seems to have been tried so infrequently that we can say little about its potential success, although the Mensah results suggest that it would not be particularly successful.

Perhaps more important for present purposes, the situation gets much less hopeful when (out-of-sample) prediction involves prediction for companies of a different type than those used to estimate the model. For example, Figure 5 shows the performance of a model estimated from a sample of manufacturing companies and used to predict bankruptcy outcomes (two years out) for a hold-out manufacturing sample and a sample of retail companies.[23] As a baseline for comparison, again the pure-chance classification accuracy is 50 percent.

Performance is good but not outstanding for the hold-out manufacturing sample, but—most importantly—it deteriorates substantially for the retail sample. In fact, the improvement of prediction accuracy for the retail sample relative to the naive prediction is only half that for prediction for hold-out manufacturing companies.

On reflection, this deterioration of predictive performance should not be surprising. Values of financial ratios (prominent predictor variables) that indicate distress in one industry may be typical of financially healthy firms in another. For example, profit margins on sales in retail are much lower than in manufacturing. As a consequence, a model fit to manufacturing firms may

[23]The data are from Mensah (1984).

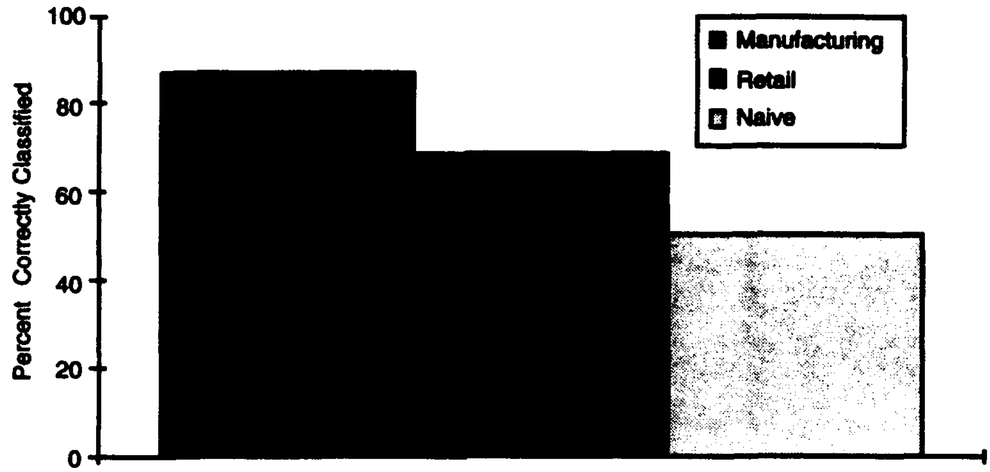

Figure 5—Performance Deteriorates with Nonsample Periods and Different Industries

predict distress for a healthy retailer. Special features of the defense industry, then, should lead us to anticipate difficulties in predicting bankruptcy of defense firms using models developed from data for firms in other types of business.

3. Using an Existing Equation

As reported in Section 2, the statistical models in the literature predict rather badly for out-of-sample time periods and especially for firms in different industries. The question now is: How well can existing equations be applied to the defense industry? Could they be expected to do even as well as the out-of-sample predictions that were shown in Section 2? To answer this question, it is useful to review some of the special features of defense firms.

The Forecasting Problem Facing DoD

The population of interest is defense contractors who supply goods not available on commercial markets. Most fundamentally, such defense firms are likely to become bankrupt for reasons quite different from those that affect commercial companies. Commercial companies tend to go bankrupt because of a combination of a weak financial position (e.g., high debt, low liquidity) combined with poor current-period performance (e.g., poor cash flow, net income). In contrast, defense contractors of the type under consideration here might be threatened with bankruptcy if they perform poorly on fixed-price contracts or if they fail to win contracts after investing large amounts of their own money in a competition. How well, then, might existing equations perform under these circumstances?

Assessment of Use of Existing Equations

We conclude that for defense companies, the forecasting accuracy of existing statistical models is likely to be much worse than would be expected from taking Figures 4 and 5 at face value. As a result we conclude that these models should not be relied on to forecast defense contractor bankruptcy.

The fundamental differences between defense and commercial business (which underlie the differences in the causes of bankruptcy) mean that several variables used in the existing statistical models have different meanings in the defense context. The indications that models perform particularly poorly when applied to companies in businesses that are very different from the one for which it was estimated (e.g., when an equation developed from manufacturing data is used to predict for retail companies) suggest substantial caution in applying models based on manufacturing or retail companies to defense contractors.

Four differences between defense and commercial companies are discussed below:

- Sensitivity to the business cycle
- Use of progress payments
- Existence of government-owned, contractor-operated (GOCO) assets
- Extreme discrepancies between book and market values.

In addition, the methods used in the received literature are not well-suited to the nature of the DoD's need for bankruptcy forecasts. Here we emphasize and discuss the length of forecast horizons.

Unusual Features of the Defense Business

Sensitivity to the Business Cycle

For the manufacturing and retail sectors generally—the sectors for which the models are estimated—bankruptcy is often triggered by poor current-period financial results associated with macroeconomic downturns. Defense activity is typically less sensitive to recessions than manufacturing and retail activities; performance on major contracts or ability to win a design competition is not sensitive to this factor.

Use of Progress Payments

Acquisition of major weapon systems often involves projects of long duration financed in part by progress payments. During such a project, measures of financial performance generally used in the forecasting models—such as cash flow or earnings—often may contain very little information about the company's eventual financial success on the project. For example, cash flow can be negative for years even if the company is performing the contract effectively and the program will eventually be a financial success. The cash flow in a very different situation—e.g., a program with a fixed-price contract that is well behind schedule and over budget—may be very similar to that of a project that is destined for financial success.

Existence of GOCO Assets

Many projects involve contractor use of plant and equipment owned by the government (so-called GOCO assets). In these cases, financial variables often used in the forecasting models—such as total assets—can have a very different meaning for a defense contractor than for the typical manufacturing company. The basic point here is that a variable like net income/total assets may have a very different relationship to the probability of bankruptcy as a result of the existence of GOCO assets. Because profit policy is intended to provide a fair return on the assets owned by the company, it is unclear precisely how this consideration will affect forecasting accuracy. Presumably it depends on such factors as the fraction of a company's plant and equipment that is owned by the government, the fraction of the company's business that is defense related, and the types of defense contracts that it has.

Unusual Discrepancies Between Book and Market Values

Additional difficulties arise because defense industries are declining. Book values of assets are commonly used in constructing predictors for bankruptcy models, but this variable is likely to be very misleading in the context of a declining industry. Consider, for example, net income divided by total (book) assets. This variable may be best interpreted as an attempt to measure a company's rate of return on the *market* value of its assets.[1] In the contemporary defense context, however, the book value of assets may be subject to sharp, somewhat arbitrarily timed (accounting) revisions. Also, some facilities used for defense production are very specialized and are of little use outside defense production; thus, the demand for them (and their market value) may be quite low. Market values will be overstated by book values unless these book values have been written down to reflect the declines in market values of defense production facilities. If they have been written down, historical data series on book values are likely to be much less smooth than the true time path of market values. In sum, relationships between book and market values are likely to be very different for contemporary defense firms than for the companies used to estimate bankruptcy models in the literature. Using variables other than book assets to deflate or normalize current performance measures—such as total sales or employees—seems preferable from this point of view.

---

[1]This is because this factor determines whether a company is currently earning the opportunity cost of its capital, a very useful indictor of future financial health.

DoD's Forecasting Needs

Forecast horizons of varying lengths are relevant to DoD for different contexts. For some purposes—formulation of bankruptcy contingency plans, say—DoD might be very interested in bankruptcy possibilities over the next few months. No study of which we are aware considers the prediction of bankruptcy for time horizons shorter than one year. In addition, the long duration of programs implies that long-term bankruptcy prediction would also be of concern to DoD—for example, if assessments of financial prospects are to be considered in source selection. But the long-horizon, out-of-sample performance of the existing prediction models is very poor. The literature tends to focus on the prediction of bankruptcy over the time horizon of one to no more than seven years. As illustrated above, forecast accuracy seems good for a year or two out[2] but deteriorates substantially for horizons of more than three years or so. No study of which we are aware considers prediction of bankruptcy over time horizons relevant to development or the complete production run of a major weapon system.

Thus, it seems like a very bad idea to apply existing bankruptcy-prediction equations to defense contractors in the contemporary environment. The poor out-of-sample performance of these models is documented, and the discussion here might be reasonably interpreted as establishing that "defense firms in the contemporary environment are *very* out-of-sample."

[2]At least when predictions are made for companies in similar businesses as the sample companies and for the same time periods as used in estimation.

4. Adapting an Existing Method

This section considers the possibility of adapting existing methods for DoD purposes. The question is: To what extent can we overcome the problems with the existing equations by using a sample with the right kinds of companies to develop new equations?

The new equations must apply to defense firms during the drawdown. One might, then, begin by collecting data on companies that are more relevant for defense and for the current circumstances. For example, defense firms during the World War II, Korean, and Vietnam drawdowns might be included, as well as firms from the late 1980s and early 1990s. After the data are collected, new parameters could be estimated using the best of the statistical approaches adapted to the defense context.

However, while using historical data on defense firms in periods of contraction is possible (and is discussed further in Section 6), it is not very promising for several reasons. Three reasons are emphasized and discussed in order:

- Unique nature of current drawdown
- Changes in the structure of defense industries
- Frequency of defense-contractor bankruptcies.

Unique Nature of Current Drawdown

The demise of the Soviet threat has brought a fundamental change in the security environment. The earlier drawdowns—and even the late 1980s—were of a very different character and may provide very little perspective on the financial implications for defense contractors of the current drawdown. Partly as a result of the change in the nature of the security threat, the size of the current drawdown is unusually large and does not appear to be reversible; companies may be less willing and less able to remain in the defense business than during earlier drawdowns.

Changes in the Structure of Defense Industries

The defense business has changed substantially since the earlier drawdowns. Partly because of the historically unprecedented scales of major weapon-system projects and partly because of the drawdown, fewer defense firms exist now than during earlier drawdowns. One implication is that fewer firms are available to acquire failing firms and rescue them from bankruptcy.

Frequency of Defense-Contractor Bankruptcies

Even for the manufacturing sector at large, the relatively small number of bankruptcies of publicly held companies—i.e., the ones for which data are available—over the course of even several years makes it difficult to construct a large enough sample of firms to estimate useful prediction models. This problem seems much more severe when we seek to estimate from a sample including only defense contractors.

Ultimately, all three of the above reasons boil down to one: constructing a large enough defense sample that is truly relevant to the task of predicting defense bankruptcy in the 1990s and beyond will be very difficult. Thus, we cannot recommend trying to develop a model specifically based on defense contractors using the basic approach popular in the literature.

5. Using Market Data to Forecast Bankruptcy

We have suggested several reasons to avoid the use of existing equations to forecast bankruptcy for defense contractors in the current drawdown and several reasons for pessimism about attempting to adapt the basic statistical forecasting approach. Recognition of these difficulties led us to search for other non-judgmental forecasting strategies. The most promising such strategy appears to be reliance on data from financial markets.

We have reviewed several reasons that historical financial data are unlikely to enable successful bankruptcy prediction. It is useful to categorize these:

- Accounting difficulties such as progress payments, GOCO assets, and discrepancies between book and market values
- Statistical challenges such as the need to forecast for very short and very long time horizons, and the infrequency of defense-contractor bankruptcies
- Practical issues such as the need to make predictions on a very timely basis
- Differences between historical and contemporary conditions such as the unique nature of current drawdown and the changing structure of defense industries
- Substantive factors that may not be measurable with available data such as the heterogeneity and interdependence among defense contractors, types of defense work and contracts, and status of ongoing DoD projects.

Moreover, bankruptcy risk of defense contractors depends on a number of additional, complex factors. First, future source selection can be crucial to a company's future financial viability. Second, emerging patterns of industry consolidation affect bankruptcy prospects of individual companies. Third, financial prospects for some firms depend markedly on the outcomes of future contract disputes (e.g., the Navy's discontinued A-12 fighter). Fourth, some companies are likely to have much more capable management and viable strategic (e.g., diversification) plans than others. These can all be thought of as examples of major factors determining future financial prospects about which historical financial data may be uninformative.

Securities market data—such as stock and bond prices—appear very promising with regard to such considerations. Securities prices reflect the assessments of market participants concerning all factors determining future profitability, including the probability of bankruptcy. Many studies support the accuracy of ex-ante market assessments, so market values should reflect sensible processing of information.

Consider securities-market evaluations of firms in light of the difficulties just categorized. Since market prices reflect assessments of economic prospects, they should be relatively insensitive to accounting issues such as those emphasized above. Since market participants are concerned about both near- and far-term prospects, market prices should contain information relevant to both short- and long-term bankruptcy prospects. Market prices are continually and quickly updated to reflect new information about companies' future financial prospects; thus, they look promising with regard to the practical issue of the timeliness of forecasts. While market participants might be better able to assess future prospects in environments more like historical environments, market assessments would be expected to incorporate (as well as market participants can do so) factors about which history provides little guidance. Market participants would also be expected to condition their valuations on the available information about substantive *factors that* cannot be measured accurately and used in statistical models. Finally, and perhaps most important, unlike backward-looking financial data, market prices are forward looking. In short, these data appear to contain most of the information of interest and much more than analysts could hope to collect and intelligently process.[1]

Stocks or Bonds?

Stocks and bonds are the two major types of financial instruments to consider. For the following reasons, stock prices seem less informative than bond prices or yields for the purposes of bankruptcy prediction.

While extracting bankruptcy probabilities may be done, in principle, for a stock, stock prices involve market-participant expectations over the whole range of possible returns. For example, stock price changes may reflect changes in expectations about the relative likelihood of "good" versus "excellent" outcomes, which have no relevance for the probability of bankruptcy.

[1]An obvious limitation of this approach is that the firm must have its equity and/or debt publicly listed, but this is generally also required for public access to financial data.

The probability of bankruptcy is closely linked to the probability of future losses large enough to trigger financial distress. Bond prices appear more informative about such losses because bonds return their (nominal) face values unless the company defaults, and they can be expected to default only in the event of financial distress. It seems reasonable, then, that bond yields should have a simpler and more stable relation to the probability of bankruptcy than stock prices do. Successfully inferring the probability of default from the yield on a bond appears plausible.[2] If so, historical experience concerning the likelihood of bankruptcy given bond default could be used to link bond yields with the probability of bankruptcy.[3] Moreover, companies often have bonds of very different maturities, which might allow differentiation of short- and long-term bankruptcy or default prospects.

In fact, there is little literature in this area, but what there is focuses on bonds—perhaps for many of the same reasons we are more optimistic about using bonds than stocks.

A Simple Method to Calculate Probability of Default from Bond Yields

Suppose a bond pays face value unless it defaults. Then, bond yields above the (risk-free) treasury rate should reflect default risk. The following general approach for uncovering default risk from the market price of the bond was developed by Fons (1985).

Assume a (risky) bond has a maturity of N years, makes a coupon payment of C per year and pays 1+C at the end of N years. Further, assume that the risky bond carries an interest rate (yield to maturity—YTM) of r, and the current risk-free rate is i. Finally, assume that there is a probability 1–P of defaulting every year and that μ is the capitalized proportion of payments recovered in the event of a default.

Under these assumptions Fons (1985) presents an equation linking C, N, i, r, μ, and P. All of these variables are known, except for μ and P. An assumption about the former[4] allows the equation to be solved for P—the default probability of the bond.

[2] A simple approach to inferring default probabilities is discussed and critiqued below.

[3] For example, Lau (p. 128) reports that 19 of her 30 firms that defaulted declared bankruptcy by the time her paper was written.

[4] In fact, Fons assumes that $\mu = 0$.

Bond Model Performance

The only study located that attempts to infer default probabilities from bond yields—Fons (1985)—was performed by the current head of research at Moody's while he was an economics doctoral student at UC-San Diego. Unfortunately, the effort was largely unsuccessful empirically. In particular, its application produced estimates of default probabilities for risky bonds that were much higher than historical default rates for such bonds (p. 69).

The reasons for these inaccuracies are not known, except that they include the so-called "junk-bond premium." It is widely believed that excess returns are available on junk bonds[5]—or at least were when Fons's data were generated. The surprisingly high yields on such bonds are explicable within the model in only one way: very high default risks on low-grade bonds. This difficulty is particularly disappointing for our purposes since the bonds of a struggling defense contractor would be expected to be low grade.

Extensions and Limitations of the Bond Model

Fons (1985) points out several factors that are not incorporated in his analysis and might underlie the inaccuracies in the calculation of default probabilities. These are discussed below. However, presently there is no obvious solution to the problem.

Fons considers several modifications or qualifications to his basic model. First, he notes that the simple formulation ignores taxes. The tax effects, however, can be ignored if the risk-free and risky debt are accorded the same tax treatment. If the risk-free rate is taken from the yield on Aaa rated bonds, which are given the same tax treatment as inferior bonds, then taxes may be ignored.[6]

Another important limitation of the basic model is that it does not consider callable bonds. Taking account of this requires the application of advanced stochastic calculus techniques that are frequently used in option valuation analysis. For our purposes, simply analyzing non-callable bonds may circumvent this complication.[7]

[5]Michael Milken had great financial success exploiting this observation.

[6] While the populations of Aaa bonds can go long periods without any defaults, the assumption leads to the obvious question of why Aaa bonds bear a risk premium over Treasury bills.

[7]A check of Moody's Industrial Manual reveals that most aerospace contractors have issued some noncallable bonds. Another notable attribute of the industry is the lack of uniformity in the capital structure of firms. Some firms have mostly callable debt, some have mostly noncallable debt, and most firms have a significant amount of what may be termed "idiosyncratic" debt, which is

In addition, the simple model assumes that investors are risk neutral. One way to consider the implications of risk-averse investors is along the lines of the Capital Asset Pricing Model (CAPM). In portfolio theory generally, investors demand an expected-return premium for assuming the risk that they must bear—i.e., that they cannot diversify away. In CAPM, undiversifiable risk is viewed as arising from a correlation between an asset's returns and the returns of an efficiently formed portfolio of other assets—generally referred to as the "market portfolio."[8] If defaults on the bonds of interest here are correlated with the market (which seems plausible), then the basic model does not apply and investors will demand *more* than the risk-free rate in expectation to hold risky debt. Fons (1985) empirically tests the hypothesis that the apparent premium on junk bonds is caused by their higher level of systematic risk; however, he has not found any evidence. The result stems from the fact that the computed measure of default is essentially uncorrelated with the return on the market portfolio.

In light of this result, other explanations for the seemingly excessive risk premium should be explored.[9] In the absence of another explanation, inferring default risk will be a very precarious exercise. As it stands, then, we do not have a satisfactory model of default risk.

Elements of a Promising Model of Default Risk

At least six elements might be essential to a successful approach to inferring default risks from bond yields:

- Fons assumes a total loss to investors if there is a default, i.e., $\mu = 0$.[10] This proportion might be substantial, however, for firms with major defense contracts.
- Fons assumes a flat term structure of interest rates. But the term structure should affect bondholder behavior, and inferences about default probabilities might be sensitive to variation in term structure.

---

unique to the firm and may be difficult to use as the basis for a bankruptcy prediction in a generic model.

[8]Risk that is "unique" to the financial instrument can be diversified away in a large portfolio.

[9]It may be that the measured premium is not "real" because of unmeasured factors that make junk bonds less attractive to investors. In any case, the model as specified cannot be relied on to predict bankruptcy because it seems to dramatically overestimate default probabilities for companies with low-grade bonds.

[10]However, this should bias default probability estimates *downward*; i.e., the high estimates of the model may be even higher after this adjustment.

- Fons assumes that historical default rates are equal to investors' expectations of future default rates. Some method of directly incorporating information about current investor expectations might be very useful.
- Bond ratings contain information about default risk, and the incorporation of bond rating data could be very helpful.
- Some form of bond price or risk spread might be included. Fons demonstrates that the monthly holding period return is superior for prediction to YTM, so the former should be used.
- Business-cycle factors are likely to contain default-risk information beyond that incorporated in the default risk premium because the expected return bondholders require to hold risky bonds depends on the evolving prospects of their entire portfolios. Including variables to reflect these business-cycle factors might be very worthwhile.

Consideration of the econometric issues in such a model would be premature. However, at least one paper appears to be particularly relevant here. Kao and Wu (1990) have used an ordered probit model to predict bond ratings, using indenture provisions and the characteristics of bonds and the issuing firms. The ordered probit respects the ordinal nature of bond ratings and avoids the assumption of multivariate normality for the independent variables required by multiple discriminant analysis.

Conditional and Unconditional Forecasts and the Needs of DoD

The previous discussion indicates that undertaking the estimation of an accurate bond price model would be very difficult. But suppose we were to press on and improve on this basic framework for the defense context and uncover from bond-market information probabilities of default for defense firms. A remaining complication is that the probability of default implicit in the bond yield often may not be the right probability for the DoD to use to support its decisions. The problem lies in the difference between a "conditional" and an "unconditional" forecast.

Bond yields reflect default probabilities given investor beliefs about the company's future prospects, including investor guesses of future DoD actions (e.g., selecting a winner of a major contract); this is referred to as the *unconditional* forecast. In many situations, however, DoD wants a *conditional* forecast—e.g., the probability of default *given* a contract award. We can be reasonably confident that the (conditional) probability of default given a contract award (alternatively, not given the award) is lower (higher) than the unconditional probability. To estimate the difference between the conditional and unconditional probabilities, DoD must guess at what the market assumes about the contract award and adjust its interpretation of the bond market data accordingly.

Such issues arise even if—and perhaps especially if—DoD thinks that the market is guessing incorrectly about future contract awards or, say, black programs. The critical point is that the adjustments should be based on differences between what the USAF expects to happen and what the market participants appear to expect to happen. Projections of future military business prospects in the business press or in research reports prepared by investment banking firms might be informative in this regard.

No matter what type of data are used in forming statistical predictions, if future activity on black programs or future DoD contracts is expected to involve neither major losses nor major profits (because of profit policy), then there is little reason to adjust for these. If, however, these programs do have major cash flow implications (for example, because progress payments do not reimburse completely for current-period financial expenditures), then these might be relevant to adjusting predictions for companies that are viewed as vulnerable financially. Foreign military sales may be more profitable or have more favorable cash-flow profiles than domestic military sales. If so, projections of major differences in foreign sales predicted by the market and the DoD might call for adjustments of the sort under discussion.

The following example illustrates that the computed conditional probabilities of default can be very sensitive to DoD's guess at the market assessment. Suppose three identical firms, each with a stock price of $40, are all competing for a lucrative contract worth $60 in stock price.[11] Furthermore, suppose that through analysis it is determined that any stock price below $65 means the firm is a bankruptcy risk.

First, note that if DoD naively applies these facts, it appears that all three contractors are bankruptcy risks, given their current stock prices below $65. In particular, if DoD confuses the probability the market is using with the probability of bankruptcy given the contract award, all three firms look like major risks to DoD. But (by assumption) the firm that receives the contract will not be a bankruptcy risk. DoD is interested in the stock price that is conditional on the award contract. That is, the *less* likely investors thought the award to be

[11]The logic of the example applies as well to bond prices, but it is easier to convey in terms of stock prices.

(i.e., the more surprising it is), the higher a firm's stock price will go afterward. As Figure 6 details, DoD's determination of whether the winning contractor is an acceptable bankruptcy risk depends on its guess of the market's assessment of the probability of winning.

In summary, the market-based approach, while seemingly promising, has two serious difficulties of its own:

- Implausibly high predicted probabilities of default using the best existing model
- Difficulties in adapting the method to the defense context.

These difficulties prevent us from recommending further pursuit of this bankruptcy-prediction strategy, even though it seems the most promising mechanical approach.

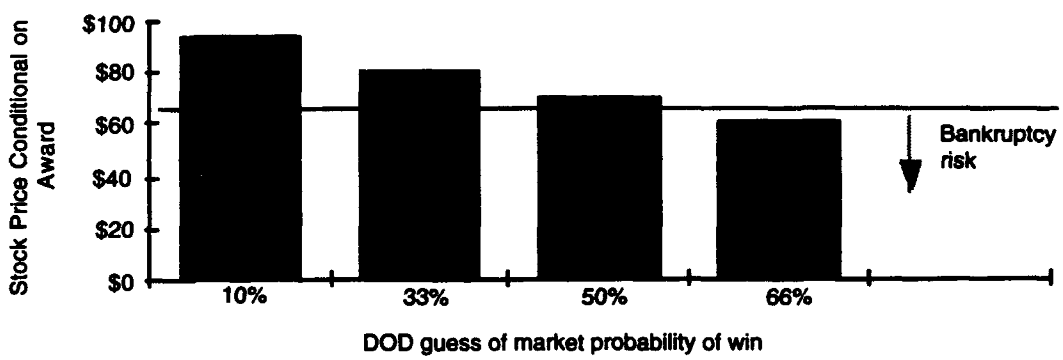

Figure 6—Small Mistakes in Guesses Can Create Large Errors in Bankruptcy Estimates

6. Where Should One Go from Here?

Three objective strategies for bankruptcy prediction have been scrutinized for DoD purposes:

- using existing statistical equations from the literature
- developing a DoD-specific statistical model using approaches similar to those most popular in the literature
- inferring default probabilities from bond yields.

All three strategies look somewhat unpromising. We are not optimistic that improved methods that could be developed over the next few years, say—using any of the strategies—would perform better than judgmental methods.

Direct support for this disappointing conclusion is provided by the results of another study. In particular, Doukas (1986) found that bankers, given three minutes to evaluate each firm, beat a leading statistical prediction model—the Springate model, which is the Zeta model adapted for Canadian use (Altman et al., 1977). The in-sample prediction accuracy of the Springate model is shown on the left of Figure 7. Doukas applies the model to his data set to obtain the out-of-sample predictions and then asks bankers to predict distress based solely on financial information given them. The bankers clearly outperform the model for one and two years out, and while the model does better three years out, neither bankers nor the model predict well for this forecast horizon.

Caution is warranted in interpreting these results; the sample size comprised only 20 bankrupt firms and 20 not-bankrupt firms. Still, the results are suggestive of the limited power of objective, mechanical methods relative to the application of expert judgment.

Some threads of hope exist. Our review does provide guidance for those attempting to develop more useful objective forecasting tools. In closing, we emphasize:

- Methods that may be quite unreliable for estimating bankruptcy probabilities might be much more reliable for *ranking* companies according to these probabilities. For some DoD decisions, being able to rank companies this way may be very useful.

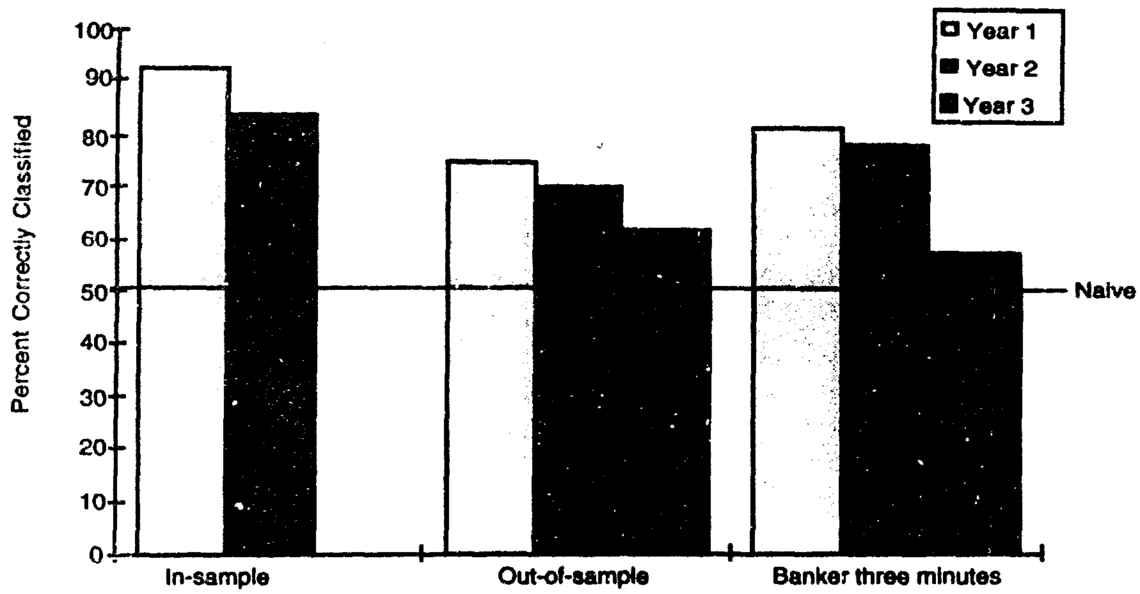

Figure 7—Time-Constrained Bankers Beat the Model

- If sufficient bankruptcies are available historically, future empirical effort should concentrate on estimating models using data from defense contractors.
- Bankruptcy is fundamentally a sequential or interactive process. Companies first fall into a vulnerable position from which poor performance can push them over the edge. Use of statistical methods that reflect such characteristics—e.g., the recursive partitioning approach used by Frydman, Altman, and Kao (1985) or classification and regression trees (e.g., see Breiman et. al [1984])—warrant serious consideration in future efforts.
- Regarding choice of predictor variables, it seems very worthwhile to consider using bond yields (or ratings) in conjunction with accounting data. Use of book values of assets should be avoided. In addition, it could be very useful—albeit very challenging—to develop variables to reflect interdependencies among defense contractors. At least two sources of interdependency warrant attention: those between competitors and those between prime contractors and their subcontractors.

Bibliography

Altman, Edward I., Robert G. Haldeman, and P. Narayanan, "Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations," *Journal of Banking and Finance*, 1977, pp. 29–54.

Altman, Edward I., "A Further Empirical Investigation of the Bankruptcy Cost Question," *Journal of Finance*, 39(4), July 1984, pp. 1067–1089.

Altman, Edward I., "Measuring Corporate Bond Mortality and Performance," *Journal of Finance*, 44(4), September 1989, pp. 909–922.

Altman, Edward I., "Multidimensional Graphics and Bankruptcy Prediction: A Comment," *Journal of Accounting Research*, 21(1), Spring 1983, pp. 297–299.

Anderson, Dan. R., "Financing Asbestos Claims: Coverage Issues, Manville's Bankruptcy and the Claims Facility," *Journal of Risk and Insurance*, 54(3), September, 1987, pp. 429–451.

Beaver, William H., "Market Prices, Financial Ratios, and the Prediction of Failure," *Journal of Accounting Research*, 6, 1968, pp. 179–192.

Boyes, William J. and Roger L. Faith, "Some Effect of the Bankruptcy Reform Act of 1978," *Journal of Law and Economics*, 29(1), April 1986, pp. 139–149.

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, *Classification and Regression Trees*, Wadsworth International Group: Belmont, CA, 1984.

Burgstahler, David, James Jiambalvo, and Eric Noreen, "Changes in the Probability of Bankruptcy and Equity Value," *Journal of Accounting Research*, 11(2–3), July 1989, pp. 207–224.

Chiang, Raymond and Richard H. Pettway, "A Note on the Segmentation of Bond Default Risk," *Southern Economic Journal*, 49(4), April 1983, pp. 1153–1161.

Chiang, Raymond, "Some Results on Bond Yield and Default Probability," *Southern Economic Journal*, 53(4), April 1987, pp. 1037–1051.

Collins, Robert A. and Richard D. Green, "Statistical Methods for Bankruptcy Forecasting," *Journal of Economics and Business*, 34(4), 1982, pp. 349–354.

Dietrich, J. Richard, "Discussion of: Methodological Issues Related to the Estimation of Financial Distress Prediction Models (Zmijewski 1984)," *Journal of Accounting Research*, 22 Supplement, 1984, pp. 83–86.

Doukas, John, "Bankers versus Bankruptcy Prediction Models: An Empirical Investigation," *Applied Economics*, 18(5), May 1986, pp. 479–493.

Eberhardt, Allan C., William T. Moore, and Rodney L. Roenfeldt, "Security Pricing and Deviations from the Absolute Priority Rule in Bankruptcy," *Journal of Finance*, 45(5), December 1990, pp. 1457–1469.

Fons, Jerome Switzer, "On the Pricing of Risky Corporate Debt," unpublished Ph.D. dissertation, University of California at San Diego, Department of Economics, 1985.

Frydman, Halina, Edward I. Altman, and Duen-Li Kao, "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress," *The Journal of Finance*, 40(1), March 1985, pp. 269–291.

Gentry, James A., David T. Whitford, and Paul Newbold, "Predicting Industrial Bond Ratings with a Probit Model and Funds Flow Components," *The Financial Review*, 23(3), August 1988, pp. 269–286.

Gilson, Stuart C., John Kose, and Larry H.P. Lang, "Troubled Debt Restructurings: An Empirical Study of Private Reorganization of Firms in Default," *Journal of Financial Economics*, 27(2), October 1990, pp. 315–353.

Giroux, Gary A. and Casper E. Wiggins, "An Events Approach to Corporate Bankruptcy," *Journal of Bank Research*, 15(3), August 1984, pp. 179–187.

Guffey, Daryl M. and William T. Moore, "Direct Bankruptcy Costs: Evidence from the Trucking Industry," *Financial Review*, 26(2), May 1991, pp. 223–235.

Haggstrom, Gus W., "Logistic Regression and Discriminant Analysis by Ordinary Least Squares," *Journal of Business and Economic Statistics*, 1(3), July 1983, pp. 229–238.

Kao, Chihwa and Chunchi Wu, "Two-Step Estimation of Linear Models With Ordinal Unobserved Variables: The Case of Corporate Bonds," *Journal of Business & Economic Statistics*, 8(3), July 1990, pp. 317–325.

Kaplan, Robert S. and Gabriel Urwitz, "Statistical Models of Bond Ratings: A Methodological Inquiry," *Journal of Business*, 52(2), 1979, pp. 231–261.

Lau, Amy Hing-Ling, "A Five-State Financial Distress Prediction Model," *Journal of Accounting Research*, 25(1), Spring 1987, pp. 127–138.

Lawrence, Edward C., "Reporting Delays for Failed Firms," *Journal of Accounting Research*, 21(2), Autumn 1983, pp. 606–610.

Mensah, Yaw M., "An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A Methodological Study," *Journal of Accounting Research*, 22(1), Spring 1984, pp. 380–395.

Moody's Industrial Manual, Moody's Investors Services, Inc., New York, 1993.

Morse, Dale and Wayne Shaw, "Investing in Bankrupt Firms," *Journal of Finance*, 43(5), December 1988, pp. 1193–1206.

Ohlson, James A., "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, 18(1), Spring 1980, pp. 109–131.

Rose, Peter S. and Gary A. Giroux, "Predicting Corporate Bankruptcy: An Analytical and Empirical Evaluation," *Review of Business and Economic Research*, 19(2), Spring 1984, pp. 1–12.

Scott, James, "The Probability of Bankruptcy: A Comparison of Empirical Predictions and Theoretical Models," *Journal of Banking and Finance*, 5, 1981, pp. 317–344.

Schmidt, Rachel, "Z-Score Analysis and Combat Industry Health," internal RAND Memorandum, February 1991.

Warner, Jerold B., "Bankruptcy, Absolute Priority, and the Pricing of Risky Debt Claims," *Journal of Financial Economics*, 4, 1977, pp. 239–276.

Weiss, Lawrence A., "Bankruptcy Resolution: Direct Costs and Violation of Priority of Claims," *Journal of Financial Economics*, 27(2), October 1990, pp. 285–314.

Zmijewski, Mark E., "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," *Journal of Accounting Research*, 22 Supplement, 1984, pp. 59–82. (Also see: Dietrich (1984) for comments.)