Fall 2002

Survival analysis of Internet companies: An application of the hazard model

Khaled Elkhal

SURVIVAL ANALYSIS OF INTERNET COMPANIES: AN APPLICATION OF THE HAZARD MODEL

by

Khaled Elkhal, B.S., M.B.A.

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Business Administration

COLLEGE OF ADMINISTRATION AND BUSINESS
LOUISIANA TECH UNIVERSITY

November 2002

UMI Number: 3066656

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

LOUISIANA TECH UNIVERSITY

THE GRADUATE SCHOOL

September 27, 2002
Date

We hereby recommend that the dissertation prepared under our supervision

by Khaled Elkhal

entitled Survival Analysis of Internet Companies: An Application of the Hazard Model

be accepted in partial fulfillment of the requirements for the Degree of

Doctor of Business Administration

Supervisor of Dissertation Research

Head of Department

Finance
Department

Recommendation concurred in:

Committee

Approved:

Director of Graduate Studies

Dean of the College

Approved:

Director of the Graduate School

GS Form 13
(1/00)

ABSTRACT

The purpose of this study is to develop a model that predicts failure and estimates the time of survival of dotcoms using a number of financial and non-financial factors. This model can be used as a warning tool for stockholders, creditors, and consumers to protect themselves from such failures.

I employ the Cox (1972) Proportional Hazards Model in a cross-sectional and time-varying context using financial data over the 1998 – 2001 period. Results from a cross-sectional analysis reveal that the coefficient estimates for variables CFTL and NSTA are consistently negative and highly significant. This suggests that higher sales and cash flows lower the potential of failure.

The results also show that NITA is negative and significant at the 10% level, suggesting that higher revenues improve the survivability of a firm. Moreover, TLTA and WCTA show no significant effect on failure. On the other hand, the coefficient estimate on TA is positive and highly significant, suggesting that larger firms have higher odds of failure. This could be the result of an unsustainable growth rate among dotcoms. The excessive and rapid need for external sources of funds may raise the concerns of creditors about the financial position of the company and can lead to higher cost of funds and closer monitoring.

The results from event-time data show qualitatively similar findings. However, the coefficient estimate for TA becomes negative. On the other hand, the event-time model does not show much significance in the overall effect of the regressors.

The time-dependent analysis, however, shows a few differences in results, in that; sales have no significant effect on the potential of failure. In contrast, the coefficient estimate on NITA becomes negative, and highly significant.

Results also reveal that stock returns add little to the predictive capability of these models. Moreover, matching companies by size to account for the size effect do not significantly alter the results. Finally, findings from industry-specific models, namely, retail, service and manufacturing, are not conclusive.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my dissertation Committee, Dr. Rohan Christie-David, Chair, Dr. James Cochran and Dr. Anita Pennathur. I am grateful for their support during the stages of the dissertation. Their guidance, comments and suggestions were helpful and greatly appreciated.

I would also like to thank Dr. Ali F. Darrat who introduced me to the field of research. I am grateful for his continuous support and encouragement throughout the years. Special thanks also go to Dr. Maosen Zhong who is always available to help with a cheerful attitude.

Finally, I want to dedicate this work to my parents, Neji and Salha Elkhal, and thank them for their patience, support and prayers over the years.

CHAPTER 1

INTRODUCTION

Over a period of five years after the introduction of the World Wide Web (WWW), the Internet-related industry has bypassed long-existing industries like energy, telecommunications and automobiles in terms of total dollar transactions. A recent study conducted by the Center for Research in Electronic Commerce (CREC) at the University of Texas at Austin provides evidence of this phenomenal growth. This study estimates that total revenues produced by the US Internet economy will reach around $300 billion compared to $350 billion for automobiles, $223 billion for the energy sector, and $270 billion for the telecommunications industry in 1998. Related sources[1] note that, compared to the overall world-wide average economic growth rate of 3.8%, the Internet economy experienced a dramatic compounded average growth rate of more than 174 %. Among the several sources that contribute to the growth in the Internet industry are the shift from the traditionally physical nature of commerce into the new cyber-related commerce and the tremendous role the internet plays in the job market. To the surprise of some, this new industry provided an estimated six million jobs to the US market in 1998.

The Internet economy, as defined by the CREC, constitutes four layers, namely the infrastructure layer, the applications layer, the intermediary layer, and the

[1] Zona Research and the International Monetary Fund.

commerce layer. The infrastructure layer provides background technology to make electronic commerce possible. Examples of this layer are Internet backbone providers, Internet service providers, and PC and server manufacturers. The applications layer is based on the Internet infrastructure and makes electronic transactions feasible. Examples of this layer include Internet consultants, multimedia applications, and web development software. The intermediary layer, such as online brokerages and online advertising, plays the role of catalyst to facilitate interaction between entities involved in e-commerce. Finally, the commerce layer represents the actual transactions between buyers and sellers over the Internet. Total revenues and jobs created by each layer are provided in Table 1 below. The number of Internet customers, as reported by *The Internet Society*, is Table 1. Total number of jobs and amount of revenues per layer of the Internet economy.

TABLE 1. Total Number of Jobs and Amount of Revenues Per Layer of the Internet Economy

Internet Layer	**Revenues (billions of $)**	**Jobs created**
Infrastructure	114,982.8	372,462
Application	56,277.6	230,629
Intermediary	58,240.0	252,473
Commerce	101,893.2	481,990
II. Total Internet Economy	301, 393.0	1,203,799

Source: The Center for Research in Electronic Commerce, The University of Texas at Austin.

estimated to be 20-30 million. These users *surf* approximately 340 million web pages (NEC Corp study) with a growth rate of about 10% per month. The nascent Internet industry is becoming more important, driving away traditional business of physical nature and shaping customers' way of shopping, thus effecting their everyday life, as well as driving companies to constantly change the way they do business to meet the new technology challenges and keep up with the fierce competition. Although the internet economy is getting more important in terms of size of revenues generated and number of jobs it provides, the main focus of this dissertation is on companies belonging to the third and fourth layer of the industry. These companies constitute almost half of the total revenues, and over 60% of the Internet job market.

Statement of the Problem

Contradictory to the rapid growth that the dotcoms have enjoyed for over five years, there is a recent trend of e-business layoffs, closings, and reconfigurations. Internet companies are currently facing enormous financial distress problems and the rate of failure among them is fairly high. Webmergers, a Research Advisory Service, notes that about 800 Internet companies that receive formal funding through private investment or by going public, have shut down while an estimated seven to ten thousand are still in business. Among those still alive, many have filed for bankruptcy and reorganized, or have been acquired by bigger companies or have merged. Bankruptcy is a legal position that a firm undertakes to get protection from creditors. It takes two forms, either filing for Chapter 7, which leads to liquidation, or chapter 11 for reorganization. Bankruptcy is a costly process. In addition to the direct costs such as legal fees, filing fees, lawyer fees, and court fees, associated with bankruptcy, (see

Warner 1977a), indirect costs are also a significant consequence of the process (Altman 1984). Among them is the increased risk of doing business associated with the bankrupt firm (Titman 1984). Bankruptcy not only has reputational effects, but also leads to increased controls over the operations and assets of firms by legal entities.

Several reasons may have contributed to the deterioration in the e-commerce market. Firstly, entry and exit of firms is fairly easy as they don't incur the high costs of building physical stores to conduct their business. Secondly, imitation of business practices is a less costly act than in physical markets, which is a major reason companies in this industry that invest heavily in market research and other R&D activities often suffer great losses. Thirdly, consumers are more willing to pay a price that reflects only the average quality of a product when facing quality uncertainty, resulting in a market for so-called *lemons* (Akerlof 1970). Contrary to physical markets, Internet shopping becomes more difficult for consumers and this asymmetric information about products quality leads to less transaction volume as well as price inefficiency. For instance, some firms are reluctant to provide more information about their products since this product information is often the product itself (Choi et al. (1997), CREC).

Purpose of the Study

Given the important role the e-commerce plays in today's economy, and the recent deterioration in the internet market as dotcom failures have become a familiar scene, it is important to examine this trend carefully in order to better understand why and how internet companies fail. The purpose of this study is to develop a model that predicts failure and estimates the time of survival of dotcoms using a number of financial and non-financial factors. This model can be used as a warning tool for stockholders, creditors as well as consumers to protect themselves from such failures. Failure in this study is defined as either filing for bankruptcy, or a complete shutdown of business. Even though filing for Chapter 11 still keeps the company in business, we consider bankruptcy as a failure since the firm must have default on debt payments to its creditors, and can no longer operate under normal conditions as it faces considerable pressure from external parties and has its reputation at stake.

To the best of my knowledge, none of the studies in the literature have empirically examined firm failure in the Internet industry. Most of the studies related to failure of dotcoms are limited to surveys and intuitive expectations of the future of these firms. Empirical work, however, has been performed mainly on banks and other firms in the manufacturing industry. The contribution of this study to the literature is to provide investors with reliable information about Internet companies' future prospects and their survival potential in order to protect their investment portfolios.

Hypotheses

The main hypothesis in this study is to test if financial data and other company characteristics have significant predictive power on the failure of dotcoms. The first hypothesis tests this proposition using the overall sample of firms. The second hypothesis tests if the industry-specific environment in which Internet companies operate significantly affects their survival probability.

Limitations

The main data source for this study is the Standard & Poor's COMPUSTAT database. As such, data is limited to publicly traded companies, specifically those trading on the New York Stock Exchange (NYSE), American Stock Exchange (ASE), National Association of Securities Dealers Automated Quotations (NASDAQ), Over-the-Counter (OTC) and wholly-owned subsidiaries of companies that are required to file with the Securities and Exchange Commission (SEC). Thus, the data sample is biased since it does not incorporate all categories of Internet companies. However, it is difficult to overcome this bias as data on small, non-publicly traded companies, is not readily available. Moreover, the study is confined only to US dotcoms. Further research may expand this sample to include foreign companies that are active outside the US.

Organizational Plan

Literature review related to failure of banks and other manufacturing firms is presented in Chapter 2. Chapter 3 describes the financial variables and other firm-specific characteristics, data, and methodology to be used in the study.

CHAPTER 2

LITERATURE REVIEW

Beaver (1966) conducted one of the earliest studies on predicting business failure. Using financial ratios on large firms, the paper applies univariate discriminant analysis over the period 1954-1964. The study finds that the ratio of annual cash flow to debt correctly classified 87% of firms one year prior to bankruptcy. The discriminant analysis used in the study is also able to correctly classify 78% of the firms five years prior to bankruptcy.

Altman (1984) investigates the empirical evidence with respect to both the direct and indirect costs of bankruptcy. He uses one sample of 19 industrial firms that went bankrupt over the period 1970-78 and a second sample of seven large bankrupt companies. Based on regression results, he finds that bankruptcy costs are not trivial. In many cases they exceed 20% of the value of the firm measured just prior to bankruptcy, and in some cases measured several years prior. He uses a second method based on the security analysts' expectations of earnings vs. actual earnings, and the results show even more dramatically that bankruptcy costs are significant. He further measures the present value of bankruptcy costs and provides strong evidence that they exceed the present value of tax benefits from leverage. He implies that firms were overleveraged and that a potentially important ingredient in the discussion of optimum capital structure is indeed the bankruptcy cost factor.

Queen and Roll (1987) examine the effect of five readily available market indicators to predict survival of firms. Results reveal that all but beta can be of use in predicting favorable and unfavorable firm mortality. They find that size is the best predictor of both favorable and unfavorable mortality over both the long and short terms. Notably, they also show that the smallest firms have about even odds of disappearing, for favorable or unfavorable reasons, within a decade. The largest firms have a mortality rate of about 20 per cent over two decades. They also provide evidence that as a sole predictor, the market price shows a monotonic negative relation to unfavorable mortality, but that both high and low-priced firms tend to have lower favorable mortality rates than mid-priced firms. When used with other predictors, however, price has a strong positive relation to favorable mortality and no relation to unfavorable mortality. This study further finds that total return and total volatility of return both appear to have strong predictive powers. As return increases the likelihood of unfavorable mortality declines and the likelihood of favorable mortality increases while high total volatility increases the rates of both types of mortality.

Ho and Saunders (1980) show that under certain reasonable behavioral conditions, a catastrophic jump in the probability of bank failure could occur even if the Federal Reserve Bank was willing to act as a continuous source of lender of last resort loans. The important relationship determining catastrophe appears to be the rate of regulatory intervention relative to the rate of deposit withdrawals. In particular, the authors show that when the perceived probability of failure for a bank is very high, neither regulator intervention nor substantial aid would be sufficient to prevent catastrophic jumps. Their model also implies that large banks whose depositors are only partially insured and who have access to the discount window are more susceptible to catastrophe than small banks. They suggest a possible "moral hazard" reason for this aspect. They find that there are circumstances under which micro-catastrophe can become macro-catastrophe, leading to an erosion of the confidence of depositors.

Simpson (1983) examines the proposition that information concerning the financial distress of large commercial banks is embodied in capital market returns. He also evaluates the conditions that would lead to the capital market predictions of distress through an analysis of six cases of major bank failures. He provides evidence that neither the intervention analysis nor the residual analysis give clear indications of financial distress. The results presented in the study may or may not conflict with the accepted theory of capital markets. He further suggests that future research on the transfer of information concerning commercial banks to financial markets should be considered in the development of early warning systems.

Blair and Heggestad (1978) examine several regulations imposed on commercial banks in general and the restriction of risk exposure in bank asset portfolios in particular. They suggest that, although portfolio regulation may reduce the probability of bank failure, its current implementation may produce perverse results. By restricting high-risk, high-return assets, bank portfolio regulation may actually increase the probability of bank failure. They further infer that the present form of portfolio regulation is *inefficient*. They finally conclude that the same goals could be achieved by placing restrictions on total portfolio return and variance without necessitating profit sacrifices by the industry.

Jagtiani and Lemieux (2001) examine pricing behavior for bonds issued by bank holding companies in the period prior to failure of their bank subsidiaries. The results indicate that bond prices are related to the financial condition of the issuing bank holding companies and that bond spreads start rising as early as six quarters prior to failure as the issuing firm's financial condition and credit rating deteriorate. Strong market discipline exists during this critical period in that the size of bond spreads for troubled banking organizations becomes many times that of healthy ones. The results suggest that bond spreads could potentially be useful to bank supervisors as a warning signal from the financial markets. In addition, the finding implies that the proposals to require bank holding companies to issue publicly traded debt in a greater volume and frequency will likely enhance market discipline in the banking system when it is most needed.

Robertson (1983) examines the changes in the financial situation of firms through ratio analysis. He suggests that the model used provides an interrelationship between the ratios that are carefully chosen for their ability to respond to changes. These variables are sales less total assets/sales, profit before tax/total assets, current assets less total debt/current liabilities, equity less total borrowings/total debt, and liquid assets less short-term borrowings/creditors.

Audretsch (1991) studies survival rates across manufacturing industries. Using 11,000 companies established in 1976, he examines the variation in rates along with the differences in the underlying technological regimes and industry-specific characteristics over a ten-year period. He suggested that the existence of substantial scale economies and a high capital-labor ratio significantly decreases the survival rates. Conducting this same test over different time periods results in different outcomes. Among other findings is the notion that market concentration promotes short-run survival, but in the long run, has no effect. The existence of small-firm innovative activity highly promotes survival of new firms.

Evans (1987) studies the relationship between firm growth, size, and age using a sample of all the firms operating in 100 manufacturing industries. His first finding is that the probability of firm failure, firm growth, and the volatility of firm growth decrease with firm's age. This finding is consistent with Jovanovic's (1982) learning model predictions. The second finding notes that firm growth decreases with firm size at a diminishing rate. Even after controlling for the exit of slow-growing firms from the sample the results still reject Gibrat's law. Nevertheless, this notion is more observable for smaller firms and tends to become less severe for larger firms.

Opler and Titman (1994) study the effect of financial distress on corporate performance. They find that highly leveraged firms are more vulnerable to financial distress than are other more conservative financed competitors. Specifically, market value of equity as well as sales decline by about 26 percent more for firms that lie in the top leverage decile relative to those in the bottom leverage decile. They also suggest that such firms with significant research and development expenditures are more affected in economic downturns than others. These results are consistent with the theory that more product specialized firms are taking higher risks than their competitors with more diversified products. This study, among others, asserts that indirect costs of financial distress are highly significant and therefore should be given more attention in capital structure decisions.

Beaver (1968) examines the degree of association between financial ratios and market prices in predicting firm failure. He applies cross-sectional and time-series analysis to data on 79 failed and 79 ongoing firms over the period 1954-1964. He finds no evidence of perfect association between the two predictors and explains this phenomenon by the fact that investors also respond to non-financial sources of information or overlook the ratios completely. He also suggests that stock prices incorporate all the information revealed by ratio analysis, and therefore, more analysis needs to be done to investigate nonratio information and its effect in shaping investors behavior.

Discriminant Analysis

Altman (1968) attempts to predict firm bankruptcy by applying a multiple discriminant analysis method to financial ratios. He points out that grouping these financial ratios into a single discriminatory model is more powerful than a univariate analysis method. The multiple discriminant analysis (MDA) considers an entire profile of characteristics, as well as the possible interactions imbedded in them simultaneously. Altman selected a sample of 66 manufacturing firms. Of these, thirty-three had filed bankruptcy during the 1946-1965 period. The remaining 33 non-bankrupt firms were matched by size and industry type to the bankrupt group. Assets of the firms in the study ranged between $7 million and $25 million. The financial ratios used in this study as predictors of financial distress include standard measures of liquidity, profitability, leverage, solvency and asset utilization. Other ratios mentioned in the previous studies include working capital to total assets, retained earnings to total assets, operating earnings to total assets, market value of equity to book value of total debt and sales to total assets. Altman shows on an individual basis that the other ratios were not the most significant discriminators. Nevertheless, Altman's model using ratios one period prior to bankruptcy resulted in 95% of the firms being correctly classified as either bankrupt or non-bankrupt. He also tests his model using financial ratios two periods prior to filing bankruptcy. While the "hit ratio" dropped to 72% the evidence still showed significant power to predict bankruptcy two years prior to the event. However, tests on data three years prior to bankruptcy resulted in a 49% hit ratio (which is no better than random guessing given the nature of the sample). Moreover, the results for four and five years out gave even weaker results. Altman concludes that his model is helpful in predicting bankruptcy only for one and two periods out.

Edmister (1972) uses discriminant analysis to predict small business loan defaults. He applies this technique to financial ratios (as is the case in previous studies). All firms he used in his sample are among industries that had received Small Business Administration (SBA) loans. He asserts that by comparing a company's financial ratios against ratios of similar firms in related industries he can single out the information that has predictive power about financial distress. Edmister applies a stepwise discriminant technique to 19 financial ratios taking into account their different variability. After the first variable was included in the model, the next discriminating variable was allowed to enter only if it had a low correlation ($\leq .31$) with the previous variable. The same restriction applied to future variables entering the model. From the original 19 ratios the model ended up with only five ratios, namely equity to sales, working capital to sales, current liabilities to equity, inventory to sales and quick ratio. He then applies this to data from 42 companies, 21 of which had defaulted on SBA loans, whereas, the other 21 companies still maintain their SBA loans in good standing. The final model correctly classified 39 of the 42 companies (a success ratio of 94%). Edmister suggests that his model is a powerful one as it is general enough to be applied to a variety of business sectors, contrary to Altman's model, which is limited to manufacturers.

Laitinen (1994) studies the failure predictive power of traditional and operating cash flow ratios. He finds that traditional cash flows are more stable and reliable than operating cash flows in predicting failure. Using discriminant analysis, logit regression and univariate analysis, he carries out his empirical study with data from 40 failed and 40 similar ongoing firms over a period of five years before failure. As in Gombola and Ketz (1983) and Casey and Bartczak (1984), he defines traditional cash flow as net income plus depreciation, and operating cash flow as an adjustment of net income for accruals and deferrals. He also suggests that operating cash flow is more sensitive to recession unlike traditional cash flow, which stays fairly more stable. The frequent adjustments in operating cash flow is due to decrease in inventories, decrease in accounts receivable, and in the last stage of the bankruptcy process to accounts payable increase.

Casey and Bartczak (1985) examine whether operating cash flow data increase the bankruptcy prediction accuracy of accrual-based cash flow data. They apply canonical correlation techniques, linear multiple discriminant analysis, and conditional stepwise logit regression for each year of data over a period of more than a decade. They reject the hypothesis that operating cash flow data has incremental predictive power over accrual-based ratios. Their results are consistent with the results of Gentry et al. (1985) and Gombola et al. (1983). They also point out four important recommendations for future research. First, while operating cash flows do not have incremental predictive power over other ratios, they may still be useful in predicting other events of interest, such as corporate acquisitions, loan defaults and dividend omissions. Second, the ability of operating cash flows to predict bankruptcy may be better in conjunction with other financial ratios as well as nonfinancial data. Third, that other definitions/forms of cash flow data, such as the variance of total cash flows can be a significant bankruptcy predictor [see Cogger (1982)]. Finally, that sensitivity of the analysis to the time period used may be of great significance [see Gombola et al. (1983)], suggesting that the study should be carried out over different time periods.

Logit Regression

Meyer and Pifer (1970), unlike other studies, use information that is not only limited to financial ratios and current financial position of firms at the bankruptcy period. They also use measures of trends, expected variations, and unexpected changes in values, and other non-financial ratios. In total, they use nine variables, among which only one represents a financial ratio. This study applies a logit analysis to establish a model that can predict bank failures. The success rate on this model is 80%.

Studies note weaknesses associated with multiple discriminant analysis within a bankruptcy context. For example, Ohlson (1980) avoids using the multiple discriminant analysis (MDA) in his study for several reasons. First, MDA assumes that variance-covariance matrices of the predictor variables are equal for both failed and non-failed groups and this may not be the case. Second, the output from an MDA model is a score that has little intuitive interpretation. The score is based on the assumption that predictors are normally distributed and, if this is not the case, future predictive results may be weak. Third, there are also problems related to the matching procedures used in MDA. Firms are usually matched by industry and size. Ohlson argues that such matches are arbitrary and it is not clear what is gained or lost through such matching procedures. Eisenbeis (1977) notes an additional problem in applying MDA. MDA assumes discrete groups when, in fact, some groups are not clearly segregated but rather are somewhat continuous. Instead, he applies the conditional logit model to predict firm bankruptcies. According to Ohlson, the use of logit analysis avoids these problems. Logit requires no assumptions about prior probabilities of bankruptcy or the distribution of predictors. He develops a logit model using nine independent variables and 105 firms that have filed bankruptcy during the period 1970-1976. A non-bankrupt group of 2058 firms is included in the model for a total of 2163 firms. The study correctly predicts bankruptcy one year in advance with 96% accuracy, demonstrating a successful alternative to MDA.

Huyghebaert *et. al* (2000) empirically examine the influence of operating activities and financial and investment decisions in the start-up year on post-entry survival, taking industry effects into account. They find that funds flow measures are superior relative to traditional financial ratios in identifying those start-up characteristics that are related to subsequent failure. They apply multivariate logit estimation technique to accounting data from 823 Belgian start-ups that were founded in 1985. They find evidence that funds flow measures have superior ability in detecting start-up characteristics that influence subsequent failure. The results hold after controlling for start-up size and industry-specific factors, which do not contribute significantly in a multivariate model of post-entry survival. They conclude that their model gives an indication of the relative importance of characteristics specific to the firm and to the industry in explaining heterogeneity in survival probabilities.

Bovenzi, Marino and McFadden (1983) develop three new bank failure prediction models using a probit analysis and other early warning models. They use financial variables in their analysis, including credit risk, liquidity risk, and capital adequacy. When comparing with other models, they show that efficiency of failure detection declines with an increase in potential failure group. Although their models are consistent with bank examiners' CAMEL test rating system, models that use financial ratios are shown to be better able to classify bank failures.

Hazard Models

Lane, Looney and Wansley (1986) develop a model to predict expected survival time for banks. They apply the Cox (1972) proportional hazards model on data from 130 failed banks. The data comes from the Federal Deposit Insurance Corporations. All banks in the sample failed during the period January 1978 through mid-1984. The group of failed banks was then matched with a sample of 334 non-failed banks. Matching each failed bank with at least one non-failed bank is based on five criteria, namely geographic location, charter status (state or national), size (based on total assets), holding company affiliation, and age. Many of the banks are privately held and have no market data available, so the study is based on an analysis of accounting data supplied to regulatory authorities. Among the twenty-one financial ratios used in the study are log of commercial loans to total loans, total loans to total deposits, log of total capital to total assets, log of total operating expenses to total operating income, log of municipal securities to total assets, total loans to total assets, net income to total capital and income taxes to earnings before taxes and security transactions. Of these financial ratios, only six are left in their model after applying a stepwise procedure. The stepwise procedure is applied to two different categories of data. The first data set includes financial ratios measured one year prior to failure and the second data set includes financial ratios measured two years prior to failure. The authors compare the classification results of the Cox model to a traditional multiple discriminant analysis (MDA) model. Using forward and backward elimination techniques, the MDA model returns the same four ratios in the one-year model and five of six ratios in the two-year model. For the first model incorporating data measured only one-year prior to bankruptcy, MDA correctly classifies 89% of the firms, whereas with the data measuring financial ratios two-year prior to bankruptcy, MDA correctly classifies 73% of the firms. On the other hand, the Cox model correctly classifies 80% of the firms in the first case and 74% of the firms in the second case. It is noteworthy that in the Cox model, type I errors (classifying a failed bank as non-failed) are significantly lower than the one in the MDA. Therefore, if the cost of committing a type I error is significantly greater than the cost of committing a type II error (classifying a non-failed bank as failed), the Cox model would be preferred over the MDA model. This would likely be the case for banking regulatory authorities whose goal is to prevent bank failures. Moreover, while the classification ability of the Cox model is comparable to that of the MDA model, the main contribution of the Cox model is the additional information it provides regarding the probability of failure and expected time to survival. In other words, the MDA model simply classifies companies into failure/non-failure categories, whereas the Cox hazard model provides keen additional information relative to the expected time to failure.

Chen and Lee (1993) use a sample of 175 firms active in the oil and gas industry to predict expected corporate survival. The huge drop in oil prices in the beginning of the eighties is considered one of the major causes of financial distress of companies in the oil sector. The authors apply the Cox (1972) hazard model to several financial ratios and other variables including liquidity, profitability, leverage, size of the company (as measured by sales), age of the company and percent of stock owned by management. The ratios were used as independent variables in a Cox proportional hazards model to predict the timing of the onset of financial distress. They define financial distress as either filing for bankruptcy, defaulting on a principal or interest payment, or suspending preferred stock dividends. The time period of the study extends from 1981 to 1988 where 67 firms encountered financial distress, 44 merged, and the remaining 64 are still in business with no major change in their operating and financial status. Chen and Lee developed three different hazards models. Model I contains all ten variables originally selected by the authors. Five of the ten prove to be highly significant. These variables are working capital to total assets, market capitalization to total assets, barrels of oil and gas reserves to book value of oil and gas properties, log of total sales and operating cash flow to total assets. However, two of the ten variables, age of the firm and log of total sales are highly correlated with each other. Therefore, in the second model the log of total sales are dropped. Model II also contains five significant variables though only four of which are common between the two models. In comparing the results of the Cox model to that of a logit model, we notice that the significant variables in the two models are identical (with the exception of one ratio). However, the interdependence of the variables is different. A significant variable in the logit model indicates its ability to discriminate between failed and non-failed firms. On the other hand, a significant variable in the Cox model indicates that the variable contributes toward determining a firm's longevity. Again, the additional valuable information of the proportional hazards model makes it probably more powerful than the logit model.

Bharat and Kini (2000) conduct a survival analysis using the Cox hazard methodology, and find that the involvement of venture capitalists improve the survival profile of IPO firms. They also find that other variables that are potentially influenced by venture capital involvement, namely, research and development allocations, analyst following, and investment banker prestige, are positively related to survival time.

Tveteras and Eide (2000) examine the effect of structural differences between new small firms and new plants of existing firms in the Norwegian manufacturing industry. In their semi-proportional Cox model, they account for plant size, capital intensity and productivity over the period 1977-1992. They suggest that the size of the firm relative to its industry's average does not significantly influence survival of entrepreneurial entrants.

Audretsch and Mahmood (1995) estimate a hazard duration function for more than 12,000 firms that were started in 1976 over a ten-year period. They do not restrict their study only over the effect of technological and market structure environments on firm survival rate, as in Audretsch (1991), but extend this to include establishment-specific characteristics, such as organizational structure and size. They suggest that ownership structure and start-up size can significantly shape the likelihood of survival of new establishments. They also suggest that these results apply only to the manufacturing sector and may not be generalized to include all other firms from different sectors.

CHAPTER 3

DATA AND METHODOLOGY

Traditionally, multiple discriminant analysis (MDA) and Beaver's univariate analysis (1966) have been widely used in the literature of firm failure. However, some studies note weaknesses associated with multiple discriminant analysis within a bankruptcy context. For example, Ohlson (1980) avoids using the multiple discriminant analysis (MDA) in his study for several reasons. First, MDA assumes that variance-covariance matrices of the predictor variables are equal for both failed and non-failed groups and this may not be the case. Second, the output from an MDA model is a score that has little intuitive interpretation. The score is based on the assumption that predictors are normally distributed and, if this is not the case, future predictive results may be weak. Third, there are also problems related to the matching procedures used in MDA. Firms are usually matched by industry and size. Ohlson argues that such matches are arbitrary and it is not clear what is gained or lost through such matching procedures. Eisenbeis (1977) notes an additional problem in applying MDA. MDA assumes discrete groups when, in fact, some groups are not clearly segregated but rather are somewhat continuous. Instead, he applies the conditional logit model to predict firm bankruptcies. According to Ohlson, the use of logit analysis avoids these problems. Logit requires no assumptions about prior probabilities of bankruptcy or the distribution of predictors. In addition to that, logit regression is superior to MDAs in terms of less restrictive assumptions used.

All these techniques, however, are limited solely to classifying firms as bankrupt or nonbankrupt entities. That is, they are used to predict whether firms are failing or not without giving any indication as to the timing of failure. In contrast, the Cox (1972) proportional hazards model brings an extra dimension, that of providing additional information on the time remaining for a firm to survive. Another feature of the Cox PH model is its ability to examine both time-varying as well as cross-sectional data, unlike traditional discriminant analysis techniques which are limited to cross-sectional data. Perhaps an even more attractive feature of the PH techniques is that they are built on very limited assumptions. In fact, the procedure can be carried out without even explicitly defining the baseline hazard function $h_0(t)$. Moreover, there is no assumption of any kind regarding the distribution of the error terms.

The Cox model is based on the assumption of proportional hazards. That is, the hazard ratio for any variable is assumed to be constant over time. An explanation of methods for checking the validity of this assumption and appropriate remedies is provided later in this chapter. Another attractive feature of the Cox technique is that it makes use of all data available. That is, censored data is used without any restrictions on the number of observations or the time over which the data is available. Each of these characteristics of the PH model makes it more efficient in terms of data usage, and easier to apply because of its limited assumptions.

Cox Hazard Model

The Cox (1972) Hazard Model was originally developed as an application in the health sector and has been employed primarily by medical doctors to serious patients diagnosed with fatal diseases such as cancer, leukemia, etc.

Prior to defining the hazard model we first define the survival function. Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is *time until an event occurs.* In case of more than one event the problem becomes one of *competing risk.* In other words, the survivor function denoted by S(t) gives the probability that a firm survives longer than some specified time t. That is, S(t) gives the probability that the random variable T exceeds the specified time t.

On the other hand, A hazard function h(t) gives the instantaneous potential per unit time for failure given that the individual has survived up to time t. It is essential to

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}$$

note that from the above hazard function formula, the expression is a ratio, with a probability in the numerator and a small time interval, Δt in the denominator. This makes the hazard function a rate rather than a probability and ranges from zero to positive infinity.

Alternatively, the survival function as defined above is a probability of survival up to a given time t and can be derived from the above hazard function as follows:

$$S(t) = \exp\left[- \int_0 h(u) du \right]$$

The survivor function S(t) ranges from 0 to 1. Conversely, if we know the form of S(t), we can derive the corresponding $h(t)$ using the following formula:

$$h(t) = -\left[\frac{dS(t)/dt}{S(t)}\right]$$

Kaplan-Meier Survival Curves and the Log-Rank Test

Kaplan-Meier (KM) Method. The KM method is used to compute the survival probability at a given time. This method uses available information including censored data. The KM formula is often called the ***product limit formula*** and is as follows:

$$S(t_{(j)}) = \prod_{i=1}^{j} P(T > t_{(i)} \mid T \geq t_{(i)})$$

Where $t_{(j)}$ is the j^{th} ordered failure time.

Since the probability of survival past time $t_{(j)}$ implies that the study subject has also survived past time $t_{(j-1)}$, $t_{(j-2)}$, etc., then the probability of survival past time $t_{(j)}$ is a product of all probabilities of survival past times $t_{(j-1)}, t_{(j-2)}, \ldots, t_{(1)}$.

The above KM formula can also be expressed in terms of the product of the probability of surviving past the previous failure time $t_{(j-1)}$ and the conditional probability of surviving past time $t_{(j)}$ given survival to at least time $t_{(j-1)}$:

$$S(t_{(j)}) = S(t_{(j-1)}) \times P(T > t_{(j)} \mid T \geq t_{(j)})$$

Log-Rank Test for Two Groups. This method is used to provide overall comparison of KM curves. It is a large sample χ^2 test that uses observed versus expected counts over categories of outcomes, where categories are defined by ordered failure times for entire set of data. The Log-rank statistic is as follows:

$$LR = \frac{(O_i - E_i)^2}{Var(O_i - E_i)}$$

where $O_i - E_i = \sum_{j=1}^{J}(m_{ij} - e_{ij}),\quad i=1,2$

m_{ij} is the actual number of failures from group i at failure time $t_{(j)}$

e_{ij} is the expected number of failures from group i at failure time $t_{(j)}$ measured as follows: $e_{1j} = \left(\frac{n_{1j}}{n_{1j} + n_{2j}}\right) \times (m_{1j} + m_{2j})$, with n_{ij} being the number of subjects in the risk set i at failure time $t_{(j)}$.

The null hypothesis being tested is H_0: no difference between survival curves. Under H_0, the log-Rank statistic LR is approximately chi-square with one degree of freedom.

<u>LOG-Rank Test for Several Groups</u>. This method is also used to provide overall comparison of KM curves. However, it is able to assess more than two curves at a time. The corresponding test statistic is:

$$\text{Log-rank} = d'V^{-1}d$$

where $d=(O_1-E_1, O_2-E_2, \ldots, O_{G-1}-E_{G-1})'$

$V=(\hat{Cov}(O_i-E_i, O_l-E_l))'$ for $i=1,2,\ldots,G-1$; $l=1,2,\ldots,G-1$; G = # of groups.

The Log-Rank test statistic is also a χ^2 test with G-1 degrees of freedom, where G is the number of curves (groups).

Peto Test. This method, developed by Prentice and Marek (1979), is different from the log-rank test in that it weights observed minus expected score at time $t_{(j)}$ by number at risk, n_j, whereas log-rank test uses equal weights. The test statistic for the Peto test is:

$$LR = \frac{(O_i - E_i)^2}{Var(O_i - E_i)}$$

where O_i-E_i, this time is measured as a weighted average. That is,

$$O_i - E_j = \frac{\sum_{j=1}^{J} n_j (m_{ij} - e_{ij})}{\sum_{j=1}^{J} n_j}, \text{ where } i=1,\ldots,G.$$

This test actually emphasizes beginning of survival curve with early failures receiving larger weights. In contrast, the log-rank test emphasizes the tail of the survival curve.

Cox Proportional Hazards Model and Its Characteristics

The Cox Proportional Hazards (PH) model is defined as follows:

$$h(t,X) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

where $h_0(t)$ is called the baseline hazard function, and X denotes a vector of p explanatory variables X_1, X_2, ..., X_p. The model is nonparametric since $h_0(t)$ is not defined.

There are three main properties of the above model. First, we don't need to specify $h_0(t)$ to be able to estimate the hazard ratio HR as it is computed below. Second, the exponential part of the formula is used to ensure that the fitted hazard is nonnegative. Finally, the model is robust - the model can be used with a wide variety of data types. It makes use of all data including censored data and survival times. For these reasons this approach is more popular and more preferred over other statistical techniques, mainly the logistic model which is limited to a dichotomous outcome. Also, the Cox model is considered a safe choice when we encounter uncertainty about the true parametric model that should be used. Results from the Cox model usually are closely comparable to those from "the correct parametric model" [see Kleinbaum (1996)].

If we want to compare the relative survivability or relative hazard of two firms, we can use the following hazard ratio of two firms with respective explanatory sets X* and X:

$$\frac{h(t, X^*)}{h(t, X)} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

The main assumption of the above HR formula is that it is time independent. That is,

$$\frac{h(t, X^*)}{h(t, X)} = \theta$$

where θ is any real number.

Therefore, for the above assumption to hold the hazard functions of the two subjects of study must be parallel across time. When hazards cross, the above assumption is obviously not satisfied. In this case the extended Cox model is more appropriate. Before proceeding to the Cox extended model, a description of the different procedures for evaluating the PH assumption is provided below.

Evaluating the Proportional Hazards Assumption

The PH model assumes that the hazard ratio is independent of time, that is:

$$\frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \hat{\theta} .$$

There are four different methods of evaluating the HR time-independent assumption. The methods and their pros and cons are discussed below.

Graphical Approach 1

The log-log survival curve is a transformation of an estimated survival curve through taking the natural logarithm of an estimated survival probability twice. This transformation results in the following expression:

$$-\ln[-\ln S(t, X)] = -\sum_{j=1}^{p} \beta_j X_j - \ln[-\ln(S_0(t))]$$

where $S_0(t)$ is the baseline survival function and all other variables are as defined earlier.

The empirical plots of log-log survival curves should be parallel for the PH assumption to be satisfied. Two types of empirical plots can be used to assess this assumption. The first is the usual KM curve defined earlier. The second is an adjusted survival curve where the predictor being assessed is not included in the Cox regression model.

The log-log method, though useful, raises the following three questions: 1) What constitutes a parallel form? 2) How is a continuous variable categorized? and 3) How are several variables evaluated simultaneously?

A possible recommendation to overcome these problems is to use a small number of categories, a meaningful choice, and a reasonable balance. In case we have several variables, however, we can either compare log-log curves from combinations of categories, or adjust for predictors already satisfying the PH assumption.

Graphical Approach 2

The second graphical approach to assessing the validity of the PH assumption uses observed versus expected plots is a graphical analog of the Goodness-Of-Fit (GOF) test. It either uses KM curves to obtain observed plots or adjusts for other variables using a stratified Cox PH model to obtain observed plots. If observed and expected plots are close, then we conclude PH assumption is satisfied.

However, this approach has limitations. One cannot judge how close it should be. Furthermore, we don't know how to categorize a continuous variable. To overcome these problems, we can conclude that PH is not satisfied only if the plots are *strongly* discrepant. When assessing the PH assumption for a continuous variable, we should derive observed plots by forming strata from categories and then obtain KM curves for each category. Also, there are two ways to obtain expected plots for continuous predictors. We can either use a PH model with k-1 dummy variables for k categories, or use a PH model with continuous predictor and specify predictor values that distinguish the different categories.

Goodness-of-Fit (GOF) Approach

This approach provides a test statistic (p-value). Thus, when using it, a clear-cut decision can be made, unlike the graphical methods. The method uses a

χ^2 statistic with 1 degree of freedom. It is based on observed and expected probabilities. If P is small, then we have departure from PH. Note that this method may fail to detect a specific kind of departure from PH.

Time Dependent Covariates

This method uses the extended Cox model assessing either one variable at a time or several variables simultaneously. That is, when assessing the PH assumption for one variable, we incorporate a time-varying factor g(t) in the usual hazard function shown below, and test for the significance of this factor. That is, test for H_0: δ=0 using a Wald test or a Log Likelihood Ratio (LR) test

$$h(t,X) = h_0(t)\exp[\beta X + \sigma X g(t)]$$

When evaluating several predictors simultaneously, however, we incorporate a time-varying factor g_i(t) for each variable i in the model as shown below and then perform a LR test with p degrees of

$$h(t,X) = h_0(t)\exp[\sum_{i=1}^{p} \beta_i X_i + \sigma_i X_i g_i(t)]$$

freedom to test for the null hypothesis H_0: $\delta_1 = \delta_2 = \ldots = \delta_p = 0$. However, the choice of g(t) is not always clear, and different choices may lead to different conclusions about the PH assumption.

Stratified Cox Procedure

The Stratified Cox (SC) model is a modification of the Cox PH model to allow for control by "stratification" of predictors not satisfying the PH assumption. Variables that are assumed to satisfy the assumption are included in the model as predictors; the stratified variables are not included in the model. The form of model is shown below:

$$h_g(t,X) = h_{0g}(t)\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$$

where g=1,2,...k*, strata defined from Z*; Z* has k* categories; $X_1, X_2, \ldots, X_p$ satisfy the PH assumption. In order to define the stratification variable Z*, we proceed as follows: identify $Z_1, Z_2, \ldots, Z_k$ not satisfying PH ; categorize each Z; form combinations of categories (strata); each combination is a stratum of Z*.

The above model is designated as a "no-interaction" model since the β's in the model are the same for each subscript g. The no-interaction assumption means that the variables being stratified are assumed not to interact with the X's in the model. However, the fitted SC model will yield different estimated survival curves for each stratum because the baseline hazard functions are different for each stratum.

The regression coefficients in the SC model are estimated by maximizing a partial likelihood function that is obtained by multiplying likelihood functions for each stratum as follows:

$$L = L_1 \times L_2 \times L_{k^*}$$

To allow for interaction, however, we can use the modified SC model as follows:

$$h_g(t,X) = h_{0g}(t)\exp[\beta_{1g} X_1 + \beta_{2g} X_2 + \ldots + \beta_{gp} X_p]$$

where g=1,2,...k*, strata defined from Z*;

An alternative is to write the interaction model using product terms involving the Z* variable with each predictor. This model uses k*-1 dummy variables to distinguish the k* categories of Z*. Each of these dummy variables is included as a product term with each of the X's.

In order to test for the no-interaction assumption, we perform a likelihood ratio test that compares the no-interaction model to the (full) interaction model. The test statistic is as follows:

$$LR = -2\ln L_R - (2\ln L_F)$$

where the subscripts R and L stand for the reduced and the full models, respectively. Also, LR~ $\chi^2_{p(k^*-1)}$ under the null hypothesis H_0: no interaction.

Extension of the Cox PH Model for Time-Dependent Variables

The extended Cox model for time-dependent variables is as follows:

$$h((t, X(t)) = h_0(t)\exp\left[\sum_{i=1}^{p_1}\beta_i X_i + \sum_{j=1}^{p_2}\delta_j X_j(t)\right]$$

where $X(t)=(X_1,X_2,\ldots,X_{p1},X_1(t),X_2(t),\ldots,X_{p2}(t))$ denotes the entire collection of predictors at time t, X_i denotes the i^{th} time-independent variable, and $X_j(t)$ denotes the j^{th} time-dependent variable.

The Maximum Likelihood (ML) procedure is used to estimate the regression coefficients of the above formula. The model assumes that the hazard at time t depends on the value of $X_j(t)$ at the same time. To calculate the Hazard Ratio Formula for the Extended Cox Model we use the following formula:

$$HR(t) = \exp\left[\sum_{i=1}^{p_1}\beta_i[X_i^* - X_i] + \sum_{j=1}^{p_2}\delta_j[X_j^*(t) - X_j(t)]\right]$$

As we can see here, the PH assumption is not satisfied because the HR(t) is a function of time. The estimated coefficient of $X_j(t)$, however, is time-independent, and represents an "overall" effect of $X_j(t)$.

Predictors of Bankruptcy

The literature relating to bankruptcy does not provide clear results concerning the best predictors of bankruptcy. For example, Lane, Looney and Wansley (1986), Ohlson (1980), and Aziz, Emanuel and Lawson (1988) use pure accounting ratios in their models to predict failure. Such predictors are chosen not necessarily on the basis of rigorous theory. Other studies use accounting ratios for lack of market data. Still others tend to choose their variables based on their frequency of use in the bankruptcy literature. Departing from this *ad hoc* choice of variables, Queen and Roll (1987) confine their failure models to investor perception about prospects of firms. They use stock price return, volatility of stock return, as well as beta, a measure of market risk. Queen and Roll justify their choice of predictor variables by asserting that accounting ratios are historical in nature and suggest little about the future direction of a firm. They also note that all accounting information is already imbedded in market data.

Beaver (1968) employs both accounting ratios and market variables in his model. He suggests that accounting data has lagged information about market prices and should best be put together in the same model to get the best prediction of failure. Given the extant evidence, I employ both market variables and accounting ratios as regressors in my model. Due to lack of other market data, I limit market information to stock returns.

A further dissension of the case of accounting data becomes important since studies in bankruptcy show mixed results on which accounting ratios serve as best predictors of failure. Aziz, Emmanuel and Lawson (1988) show that cash flow data are superior to accrual based data in predicting bankruptcy up to five years prior to the event. However, Altman's Z model (1968), which is an accrual based model, is a better predictor of failure two years prior to bankruptcy. Beaver (1968) provides evidence that non-liquid asset ratios are better than liquid assets ratios in predicting bankruptcy even one to two years prior to the event. He suggests that cash flow, net income, and debt positions of the firm represent permanent aspects of a firm and cannot be easily manipulated. He further suggests that firms with good profit potential but poor liquid asset position are usually able to raise funds to meet their dues. In contrast, he notes that a firm that is adequately liquid but does not have promising future profits will not wait long before declaring bankruptcy. Thus, given these conflicting findings, I employ the most frequently used accounting ratios in the bankruptcy literature, which also cover the more important aspects of a firm's financial position. These variables account for a firm's profitability, solvency, liquidity, leverage, size and asset utilization. A complete a list of the variables used and their definitions is shown in Table 2 below.

TABLE 2. Variables and Their Definitions

Variables	Definitions
RET	Annual stock return
TA	Log(Total Assets / GDP Deflator)[2]
TLTA	Total Liabilities / Total Assets
NITA	Net Income / Total Assets
CFTL	Cash Flow / Total Liabilities
WCTA	Net Working Capital / Total Assets
NSTA	Net Sales / Total Assets

[2] Logarithmic transformation is used to eliminate the growth of the variance of the data over time.

Data

All data on the above listed variables is gathered from Research Insight. The data constitutes 225 fairly large, publicly traded, pure internet-based companies, 26 of which have failed, the balance 199 are still "alive". Most of the companies used in this study feature a ".com" suffix in their names as an indicator of the nature of e-business in which they operate. All observations are gathered over the 1998 - 2001 period. Given that the dates of failure of these companies are different, the gathered information about each company may correspond to a different economic situation. For this reason, I also use an indicator of the strength or weakness of the economy, such as the growth rate of GDP, and study its interaction with other relevant variables, in order to control for economic conditions.

CHAPTER 4

RESULTS

Background on Sample Employed

The initial data set comprise 225 firms, twenty-six of which went bankrupt during the period 1999 – 2001. The data are filtered to exclude mergers and acquisitions. Following merger or acquisition, a firm may no longer exist. Such an event should clearly not be considered a failure, as some of these actions could be favorable to the firm, in that the action can increase the firm's efficiency, productivity, and market share and better meet the needs of management in terms of possible improvement in working conditions. Furthermore, not all acquisitions are in the form of hostile takeovers. Thus, removal of the sixteen acquired and merged companies from the sample is prudent. Without taking this measure, it is not clear whether the variables employed in this study explain bankruptcy or predict mergers and acquisitions.

Due to a few inaccurate measurements, in the COMPUSTAT database, I adjust the data one more time to account for possible outliers. Given the substantial amount of data provided by this database, errors in data entry are likely to occur. Such outliers exist in the form of stock returns in hundreds of thousand percent and total liabilities being larger than total assets for existing firms with continuous operations.

The data frequency is annual and collected over the 1997 – 2001 period. This can result in the number of observations, after adjustment for outliers and accounting for missing data, differing from one year to another. Primarily this is due to two reasons: adjustments to outliers and lack of market data for different years. A complete breakdown of the number of observations in each year is shown in Table 3 below.

TABLE 3. Data Availability

Data	(1) W/O Market Ret	(2) All variables	Failed		% Censored	
			(1)	(2)	(1)	(2)
1997	83	39	12	1	85.54	97.44
1998	161	56	21	4	86.96	92.86
1999	190	82	25	8	86.84	90.24
2000	152	122	14	13	90.79	89.76

Time of Origin and Time of Survival

Studies differ in choosing the best origin from which to measure survival time of an entity. Most state that there is no unique time origin. Indeed, time origin depends on the study at hand. Even though some survival analysis models use age as time of survival, (that is, date of birth as the origin), this may not be suitable and can even be misleading in predicting dotcom failures. This view reflects the fact that many companies in this study did not start as dotcoms, especially those that were founded before the Internet evolution. In my sample, I find about 13% of companies existed before the Internet economy. Thus, it is not appropriate to compare their survival times with the age of pure dotcoms. Such an evolution reflects differing firm strategies, industry structures, economic environments, and types of products. Chen and Lee

(1993) suggest that survival analysis requires an adverse economic situation to figure out what companies were able to survive such a crisis, and assess the characteristics of such survivors. In their study they define the decline in oil prices in early 1980s as an economic adversity in the oil and gas industry and measure financial conditions of companies around the same time. Following their suggestion, I define the period 1997 – 2001 as an adverse economic period for dotcoms due to the increased bankruptcy rates among them. In retrospect, it is apparent that Internet companies were overrated by investors and enjoyed highly inflated market values. When such expected future earnings were not realized, "the bubble burst" and a significant number of dotcoms went bankrupt. To account for the trend of failure during this period, I take measurements on financial positions of Internet firms using the year 1997 as the base year.

Methodology

Bankruptcy is the result of a long process of financial deterioration. It is therefore likely predictable beforehand and most often is not a surprising event. It is also apparent that measurements on the financial situation of companies at different points in time prior to the event should incorporate much more information about the process of failure than a single measurement. Although most of bankruptcy studies acknowledge this concept, few actually take repeated measurements to predict failure. An added problem is that such variables are usually used in their levels format. A more appropriate use of variables could be their first difference format, since this format shows the change in a financial variable from one point in time to another rather than just a simple number.

In this study, I use a cross-sectional model, consistent with the bankruptcy literature, as well as a time-varying one to account for the deterioration process. I also use variables in both formats, i.e., levels and first differences. Most studies of this type also match failed companies with a number of non-bankrupt ones on the basis of size. This is a well-known procedure that attempts to ensure that the analysis is free of any size bias. I follow the same concept and match failed with non-failed dotcoms of approximately the same size, measured in terms of total assets. However, in doing so, the sample size decreases significantly, to approximately 50 observations when I use accounting ratios only, and to approximately 27 observations when I use market data.

Descriptive Statistics

Before proceeding with the survival analysis, I provide descriptive statistics, shown in Table 4 below[3]. Comparing the mean stock returns of bankrupt companies to

TABLE 4. Descriptive Statistics

	1998 data		2000 data		% change (98-00)	
Variables	Failed	Full sample	Failed	Full Sample	Failed	Full Sample
RET	23 %	67 %	-91 %	-29 %	-490.30	-142.72
TA	2.51	1.56	3.70	3.37	47.49	116.80
TLTA	0.74	1.26	0.55	0.60	-25.37	-52.47
NITA	-1.72	-1.94	-3.35	-1.57	95.03	-18.93
CFTL	-3.31	-1.90	-10.22	-3.38	209.17	77.97
WCTA	0.16	-0.34	0.18	0.06	11.55	-118.42
NSTA	0.36	1.98	0.99	0.90	171.82	-54.66

[3] More detailed statistics about standard deviations, minimum and maximum are shown in Appendix A.

those of the entire sample, it is apparent that all relevant information about dotcoms has, to a large measure, already been incorporated in the stock price. Note that the return for failed dotcoms is only 23%, in 1998, below the 67% for the full sample. This also continues through the year 2000 where stock returns fall, on average, from 23% to -91%, more than three times lower than stock returns of the full sample. This indicates that relatively lower stock returns characterize failing dotcoms. Such a finding agrees with prior intuition.

What is remarkable, however, is the evidence that the bankrupt firms continue to have larger total assets compared to those firms in the full sample across these years. Even more surprising is that total assets for bankrupt firms went up by almost 50 percent between 1998 and 2000.

Total liabilities to total assets (TLTA) is lower for failed companies than for those in the full sample. This ratio drops by about 25% over a period of two years. A straightforward explanation can be offered this could simply be due to higher equity financing. For these companies, it becomes more difficult, in a continuous financial deterioration, to raise funds through debt.

Although, net income to total assets ratio (NITA) is already lower for bankrupt firms than that of the entire sample in 1998, it falls by more than 90% in the year 2000, becoming even much lower than that of the full sample. The deterioration process becomes highly visible when this variable is examined.

An even better predictor of bankruptcy is the cash flow to total liabilities ratio (CFTL) which shows the solvency of the firm. This ratio for bankrupt companies continues to be lower than that for the full sample over the years. Note also that it drops by more than 200% from 1998 to 2000, at a much faster rate than the decrease in the same ratio for the rest of the dotcoms.

A higher the net working capital to total assets ratio (WCTA) suggests that more current assets are being financed by long-term funds. This is the case for the bankrupt firms in 1998 as compared to that of the more risk-taking firms in the full sample (whose WCTA ratio is negative). This long-term financing strategy for working capital (current assets) does not change over time for bankrupt companies. It is a less risky financing policy but imposes a higher burden on the cash outflows as it is more costly to borrow long-term. This could have contributed to the higher drop in CFTL that led to bankruptcy.

The 172% increase in net sales to total assets ratio (NSTA) from 1998 to 2000 did not prevent the bankrupt firms from failing. Starting at a much lower level compared to those in the full sample in 1998, bankrupt dotcoms outperformed their counterparts in terms of relative sales in the year 2000. This strongly suggests that costs went up by an even faster rate. This is actually the case as long-term financing lowered the bill-paying capability of the firms and resulted in a drainage of the cash assets of the firm. The NSTA ratio taken by itself may result in misleading results about the prediction of failure. However, the collective evidence must be evaluated, i.e., evaluate revenues and costs together to be able to adequately assess the final outcome.

Partial Likelihood Technique

The Cox Proportional Hazards [PH] Model uses a partial likelihood technique to estimate the parameters while making no assumptions about the shape of the baseline hazard function. Unlike the maximum likelihood [ML] method, the partial likelihood ignores the baseline function. Consequently, there is some information about the vector of coefficient estimates that is lost by discarding this portion of the maximum likelihood. Although this leads to the coefficients not being fully efficient, with larger standard errors as opposed to those from a maximum likelihood estimation, the loss of efficiency is quite small in most cases [Efron 1977]. The benefit of a partial likelihood method is the increased robustness, in that the estimates are consistent and asymptotically normal regardless of the shape of the baseline hazard function.

In order to conduct a survival analysis, one should use data in a specific format. I assume that the structure of the data is composed of three main parts, namely, t_i, δ_i, and X_i, where t_i is the time of the event or the time of censoring and δ_i is an indicator variable with a value of 1 if t_i is uncensored or a value of 0 if t_i is censored. The $X_i=[x_{i1} \dots x_{ik}]$ is a vector of k covariate values. The likelihood function PL is therefore written as a product of the likelihoods for all the events and not the individual observations that are observed. That is, if J is the number of events, we can write the partial likelihood function as:

$$PL = \prod_{j=1}^{J} L_j$$

where L_j is the likelihood for the j^{th} event. To estimate these individual L_j's the data are first sorted in ascending order by survival time. To see how the first partial likelihood is constructed assume that the first event happened to company i at time t_1. Given that an event occurred at time t_1, the partial likelihood L_1 of this event is the probability that this event happened to company i rather than to any of the other n-1 companies. This is equivalent to the ratio of the hazard of company i at time t_1 divided by the sum of the hazards for all the companies that were at risk of bankruptcy at that same point in time t_1. Therefore, L_1 can be written as follows:

$$L_1 = \frac{h_i(t_1)}{h_i(t_1) + h_{i+1}(t_1) + \ldots + h_n(t_1)}$$

Since the hazard $h_i(t_1)$ can be substituted with $\lambda_0(t_1)e^{\beta X_i}$ then a more general expression of the partial likelihood follows:

$$PL = \prod_{i=1}^{n} \left(\frac{e^{\beta X_i}}{\sum_{j=1}^{n} Y_{ij} e^{\beta X_j}} \right)^{\delta_i}$$

where $Y_{ij}=1$ if $t_j \geq t_i$; and $Y_{ij}=0$ if $t_j < t_i$. In order to maximize the function above with respect to β, it is more convenient to use it in its logarithmic form as follows:

$$\log PL = \sum_{i=1}^{n} \delta_i \left[\beta X_i - \log\left(\sum_{j=1}^{n} Y_{ij} e^{\beta X_j} \right) \right]$$

Cross-Sectional Model

A cross-sectional model requires measurement on financial determinants of various companies at a specified point in time. As noted earlier, these measurements should be taken in or around an adverse economic and financial crisis. Some uncertainty arises on whether to use the same specific year for all involved companies (calendar data) or take measurements at a specified period (say one year) prior to the occurrence of the event (event-time data). The use of calendar data can lead to some results that cannot be generalized over other periods of time. The financial condition of all firms during a crisis may not have the same predictive power if these same measurements were taken over another period of expansion for instance. Such results can be very specific to a certain era of the industry and cannot be generalized. Conclusions made using this approach may have little use in predicting failure under different economic and financial conditions. To accommodate this shortcoming, I measure financial conditions at constant time periods prior to the event for each specific company.

In this study I use variables taken one, two and three years prior to bankruptcy of each firm. A major critique of this analysis is that financial determinants are taken at different points in time for each company that could coincide with different economic situations. A possible solution to this problem could be adjusting these regressors using a variable that reflects the relevant economic condition, say, the growth rate of the GDP. However, since this study covers a short period of time of only four years, which also corresponds to more or less the same economic conditions, using such an adjustment process could be of little use and may only contribute to a loss of degrees of freedom.

Results of the year 2000 data are shown below. As expected, TLTA shows a positive sign indicating that higher liabilities of the firm are associated with higher odds of failure. Higher liabilities put more burden on the company to generate cash

TABLE 5. Summary of Cox Regression Results Using the Year 2000

	2000 Data			
	Total	Event	Censored	% Censored
	152	14	138	90.79

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	18.2233	6	0.0057
Score	48.6979	6	<.0001
Wald	24.2012	6	0.0005

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.30394	0.17772	2.9251	0.0872	1.355
tlta	1	0.89618	0.93079	0.9270	0.3356	2.450
nita	1	-0.23775	0.08468	7.8821	0.0050	0.788
cftl	1	-0.05567	0.02289	5.9154	0.0150	0.946
wcta	1	1.14470	0.97220	1.3864	0.2390	3.142
nsta	1	0.11744	0.17026	0.4758	0.4903	1.125

flows to pay back its dues when they mature. In case they default on a payment, they are declared bankrupt. This process gets even worse as creditors start to monitor and restrict the activities of the company. Moreover, suppliers and consumers become more reluctant to do business with a firm in such conditions. The variables NITA and CFTL also exhibit negative signs as expected. The higher the net income, the lower the odds of failure as the potential of paying back the outstanding liabilities is improved. Net income, however, by itself, provides little information about the timing of the cash flows since revenues and expenses are recognized as they occur and not when cash changes hands. It merely gives an indication of the potential of the company to meet required payments. The CFTL variable, however, confirms the result of net income and shows a negative sign. Obviously, the higher the cash flows, the better the chance of solvency and therefore, the lower the probability of bankruptcy. It is worth noting that these two variables are significant at the one percent level of significance.

The third significant variable in the model, TA, exhibits a positive sign, contrary to what is expected. The bankruptcy literature suggests larger firms are better managed and better protected from failure. However, in this case, the model indicates that higher total assets are associated with higher odds of failure. A possible explanation of this phenomenon is that these failed dotcoms could have had an unsustainable growth rate in their total assets. Any increase in assets has to be met by a similar increase in sources of funds as represented by liabilities and stockholders equity. There are three main ways to provide financing for the rise in assets, namely, increase in current liabilities such as accounts payable, increase in retained earnings for profitable firms, and through acquiring external funds. The excessive and rapid need for these external sources of funds may raise the concerns of creditors about the financial position of the company and can lead to higher interest rates charged, closer monitoring, and other restrictions.

The other two insignificant variables in the model are WCTA and NSTA. The signs of their coefficient estimates are opposite to that expected. They both show positive signs indicating that higher net working capital and higher net sales both suggest a higher probability of failure. Higher working capital improves the bill paying ability of the firm and therefore should contribute to a more rigid survival of the firm. Moreover, net sales are the primary revenues of a company and we expect higher odds of survival as this variable increases. Such unexpected signs raise suspicion about the specification of the model. Although the model is highly significant as shown by the three statistics (Likelihood Ratio, Score and Wald test), only 3 out of the 6 variables are significant. These findings are puzzling. It might also be noted that the variables TLTA, NITA and WCTA are highly correlated, suggesting possible multicollinearity problem in the model[4]. A possible remedy to this problem is to reduce the number of variables in the model. In doing so, I test three reduced models. Following common practice, I eliminate 2 of the 3 highly correlated variables. The results of the three reduced models are shown below. The sign on the coefficient estimate for the net working capital changes, but is not significant. However, the sign on NSTA (but not significant) remains positive. The variable CFTL remains highly significant in the reduced models and shows consistency in terms of sign, as does the TA variable.

TABLE 6a. Reduced Model with TLTA Retained

2000 Data				
	Total	Event	Censored	% Censored
	154	14	140	90.91

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	12.4814	4	0.0141
Score	46.0503	4	<.0001
Wald	18.1542	4	0.0012

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.23084	0.17300	1.7804	0.1821	1.260
tlta	1	0.18596	0.25271	0.5415	0.4618	1.204
cftl	1	-0.08132	0.01916	18.0089	<.0001	0.922
nsta	1	0.12776	0.14999	0.7256	0.3943	1.136

[4] A matrix of the correlation of the variables and their significance is provided in appendix B.

TABLE 6b. Reduced Model with NITA Retained

2000 Data			
Total	Event	Censored	% Censored
154	14	140	90.91

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	16.8161	4	0.0021
Score	48.3257	4	<.0001
Wald	22.0990	4	0.0002

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.35646	0.18024	3.9112	0.0480	1.428
nita	1	-0.17817	0.07125	6.2524	0.0124	0.837
cftl	1	-0.06948	0.01954	12.6401	0.0004	0.933
nsta	1	0.15544	0.16214	0.9191	0.3377	1.168

TABLE 6c. Reduced Model with WCTA Retained

2000 Data			
Total	Event	Censored	% Censored
152	14	138	90.79

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	12.2020	4	0.0159
Score	44.7227	4	<.0001
Wald	17.8473	4	0.0013

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.21972	0.17358	1.6023	0.2056	1.246
cftl	1	-0.08099	0.01954	17.1857	<.0001	0.922
wcta	1	-0.14524	0.26857	0.2925	0.5886	0.865
nsta	1	0.12975	0.15073	0.7409	0.3894	1.139

A likely explanation for these results is that the variables are measured in 2000, nearly the same time of the bankruptcy of most failed firms. As stated earlier, all events occurred between the years 2000 and 2001. Taking measurements very close to failure time may not be the best way to proceed. Such ratios are significantly affected by the deterioration process and may not represent the normal financial situation of the firm. To account for this I also take measurements on all observations at the end of the 1998 fiscal year. The results from this procedure, shown below, demonstrate improved significance in all variables employed in the reduced models. Moreover, the coefficient estimate for the variable WCTA, though insignificant, is negative as suggested by the literature. So does the parameter for the variable NSTA, which becomes significant at the 5% level of significance. On the other hand, the variable TA becomes more significant compared to 2000 data and is still consistent in terms of its positive sign.

TABLE 7a. Reduced Model with TLTA Retained

1998 Data			
Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	20.4948	4	0.0004
Score	12.8028	4	0.0123
Wald	13.2632	4	0.0101

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.34733	0.14059	6.1035	0.0135	1.415
tlta	1	0.26025	0.17619	2.1820	0.1396	1.297
cftl	1	-0.15487	0.06695	5.3513	0.0207	0.857
nsta	1	-1.07504	0.49266	4.7617	0.0291	0.341

TABLE 7b. Reduced Model with NITA Retained

1998 Data			
Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	20.9596	4	0.0003
Score	12.9321	4	0.0116
Wald	13.6169	4	0.0086

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.36102	0.14250	6.4183	0.0113	1.435
nita	1	-0.09998	0.05690	3.0879	0.0789	0.905
cftl	1	-0.13412	0.06484	4.2782	0.0386	0.874
nsta	1	-0.98343	0.50335	3.8173	0.0507	0.374

TABLE 7c. Reduced Model with WCTA Retained

1998 Data			
Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	19.8909	4	0.0005
Score	12.7478	4	0.0126
Wald	12.5950	4	0.0134

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.33684	0.14286	5.5592	0.0184	1.401
cftl	1	-0.15005	0.06762	4.9243	0.0265	0.861
wcta	1	-0.21052	0.18414	1.3070	0.2529	0.810
nsta	1	-1.04195	0.49949	4.3515	0.0370	0.353

To confirm the above reasoning I employ event time data instead of calendar data. The latter method employs measurements of financial condition of firms at different time periods prior to bankruptcy. Results given by such models may not be generalizable and could be specific to a certain era in the life of dotcoms. Therefore, to improve the generalizability of the model I measure financial ratios three years prior to bankruptcy of failed companies. The results shown below show consistency in the significance of net sales with the expected negative sign. The variable CFTL also shows the correct negative sign as all other variables do. In addition, the coefficient estimate for the variable TA becomes negative. This result is consistent with the literature that the higher the size of the firm, the more immune from bankruptcy it becomes. On the other hand, the event time model does not show much significance in the overall effect of the regressors.

TABLE 8a. Reduced Model with TLTA Retained: 3 Prior Data

Total	Event	Censored	% Censored
138	11	127	92.03

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	11.4346	4	0.0221
Score	6.5383	4	0.1624
Wald	9.9199	4	0.0418

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	-0.00248	0.20571	0.0001	0.9904	0.998
tlta	1	0.27800	0.15057	3.4087	0.0649	1.320
cftl	1	-0.06286	0.09611	0.4277	0.5131	0.939
nsta	1	-1.26135	0.60844	4.2977	0.0382	0.283

TABLE 8b. Reduced Model with NITA Retained: 3 Prior Data

	Total	Event	Censored	% Censored
	138	11	127	92.03

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	10.6378	4	0.0310
Score	7.7704	4	0.1004
Wald	9.0713	4	0.0593

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	-0.05353	0.20068	0.0712	0.7897	0.948
nita	1	-0.08079	0.05232	2.3841	0.1226	0.922
cftl	1	-0.02921	0.09430	0.0960	0.7567	0.971
nsta	1	-1.13842	0.66732	2.9103	0.0880	0.320

TABLE 8c. Reduced Model with WCTA Retained: 3 Prior Data

	Total	Event	Censored	% Censored
	138	11	127	92.03

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	10.5361	4	0.0323
Score	5.8802	4	0.2083
Wald	8.4302	4	0.0770

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	-0.05200	0.20245	0.0660	0.7973	0.949
cftl	1	-0.04966	0.09485	0.2742	0.6005	0.952
wcta	1	-0.21991	0.14656	2.2515	0.1335	0.803
nsta	1	-1.24450	0.63497	3.8414	0.0500	0.288

Time-Varying Model

As previously mentioned, bankruptcy is the result of a deteriorating process in the financial condition of the firm. To monitor these changes closely I take multiple measurements of financial determinants for each observation. Unlike the cross-sectional case where data is collected at a single point in time (where I implicitly assume that variables are constant across time), a time-varying model could be more appropriate as it accounts for the changes in the financial condition of the firm.

An issue of concern, as stated earlier, is whether to take measurements on all observations at same points in time, say 1998 through 2000, or to collect the financial variables based on event time, i.e. one or two years prior to the event for each of the individual companies. As in the case of cross-sectional analysis, when using event-time data measurements may not be taken at the same period for each company since this again may correspond to different economic situations. I earlier noted that all events occurred between the years 2000 and 2001. Thus using event time should not represent any problem as the time frame of the data collected will range over approximately two years, i.e., just few years prior to 2000. The results from the first method, using calendar data, may not apply to different time periods where the state of the economy and investors perceptions could be different. Therefore, the event time analysis seems more reasonable in this case. To highlight the similarities and differences in results of both methods, I conduct the analysis based on both 1999-2000 data as well as one year prior to bankruptcy.

Since all companies in the sample experience bankruptcy between the years 2000 and 2001, I take measurements twice for each observation. Such measurements are limited to two observations since the data is collected on an annual basis. If bankruptcy occurrence is spread out over several more years then more measurements can be obtained. Results of the time varying model using calendar data are shown below. Three out of four coefficient estimates are highly significant. The coefficient estimate for the variable NSTA exhibits a positive sign but is not significant. Cash flow and net income ratios continue to show results that are consistent with the literature. On the other hand, the total assets variable is significant, but positive.

TABLE 9. Reduced Model with NITA Retained

TV 99 - 00 Data				
	Total	Event	Censored	% Censored
	208	20	188	90.38

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	9.2238	4	0.0557
Score	13.0693	4	0.0109
Wald	11.9641	4	0.0176

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.31054	0.14012	4.9119	0.0267	1.364
nita	1	-0.09008	0.04148	4.7146	0.0299	0.914
cftl	1	-0.07336	0.03117	5.5407	0.0186	0.929
nsta	1	0.09130	0.15043	0.3683	0.5439	1.096

When using event time data, as shown below, three out of six variables in the full model show significance at least at the 5 percent level. However, coefficient estimates for the variables working capital and net sales appear with positive signs. This is partially corrected when reducing the model by taking TLTA and WCTA out.

The variable CFTL is negative, as expected, and highly significant. The parameter estimate for the variable TA still shows a positive sign.

TABLE 10a. Summary of Cox Regression Results
Using Event-Time Data: TV 1 Prior Data

	Total	Event	Censored	% Censored
	208	21	187	89.90

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	15.5501	6	0.0164
Score	16.1470	6	0.0130
Wald	17.4513	6	0.0078

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.34642	0.14425	5.7674	0.0163	1.414
tlta	1	0.17236	0.25374	0.4615	0.4969	1.188
nita	1	-0.25678	0.09862	6.7795	0.0092	0.774
cftl	1	-0.04909	0.04501	1.1894	0.2754	0.952
wcta	1	0.91841	0.46656	3.8749	0.0490	2.505
nsta	1	0.02124	0.12864	0.0273	0.8689	1.021

TABLE 10b. Reduced Model with NITA Retained

TV 1 Prior Data				
	Total	Event	Censored	% Censored
	208	21	187	89.90

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	10.7823	4	0.0291
Score	14.2701	4	0.0065
Wald	12.6929	4	0.0129

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.34293	0.13889	6.0963	0.0135	1.409
nita	1	-0.06526	0.04632	1.9853	0.1588	0.937
cftl	1	-0.08634	0.03007	8.2437	0.0041	0.917
nsta	1	-0.06140	0.13408	0.2097	0.6470	0.940

Market Data

So far, the analysis has been confined to accounting ratios that are historical in nature. Disclosure of a firm's financial statements is not required until approximately 90 days prior to the end of the fiscal year. The results may also be an artifact of accounting practices such as differing accounting methods related to depreciation and cost of goods sold computations or the reporting of plant and equipment and other assets at the lower of cost or market values. Such practices may give different impressions about the performance of two identical firms. To mitigate this problem, market variables are introduced in the model. Since market variables are not available on most of the dotcoms, especially those which are bankrupt, the sample size decreases dramatically.

The variable stock returns in both time-constant and time-varying models, whether using calendar data or event-time data adds little to the predictive capability of these models. The results below also show that coefficient estimates for the variable RET are not significant, though they exhibit the correct negative sign. This could be due to the lack of synchronization between the market variable and the accounting ratios. All stock returns are taken at the end of the calendar year whereas all other ratios are measured at the end of the fiscal year for each individual company. To account for this discrepancy in the timing of measurements, I employ a variable with values ranging from -6 to 6, representing the number of months between the fiscal year and the calendar year for each observation. Unfortunately, this fails improve the predictive capability of the model.

TABLE 11a. Summary of Cox Regression Results
Using Market Data: 2000 Data

	Total	Event	Censored	% Censored
	128	13	115	89.84

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	21.3026	5	0.0007
Score	50.0426	5	<.0001
Wald	18.9503	5	0.0020

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ret	1	-2.58915	2.47495	1.0944	0.2955	0.075
ta	1	0.36356	0.19686	3.4105	0.0648	1.438
nita	1	-0.14490	0.07701	3.5401	0.0599	0.865
cftl	1	-0.08014	0.02567	9.7475	0.0018	0.923
nsta	1	0.14613	0.16424	0.7916	0.3736	1.157

TABLE 11b. Summary of Cox Regression Results Using
Market Data: TV 99-00 Data

	Total	Event	Censored	% Censored
	208	14	194	93.27

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	12.1706	5	0.0325
Score	11.9552	5	0.0354
Wald	9.8394	5	0.0799

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ret	1	-1.17248	0.96606	1.4730	0.2249	0.310
ta	1	0.39162	0.17441	5.0421	0.0247	1.479
nita	1	-0.08748	0.04932	3.1461	0.0761	0.916
cftl	1	-0.06736	0.03914	2.9624	0.0852	0.935
nsta	1	0.10135	0.16246	0.3892	0.5327	1.107

Other Considerations

Most of the studies on bankruptcies provide models based on matched data. That is, the sample used is composed of bankrupt firms matched with an equal number of non-bankrupt ones. The matches are usually made in terms of company size as proxied by either total assets or total net sales. This method is conducted to eliminate any size bias from the model. Due to the small number of bankrupt firms in my sample (especially after including market data in the model) I opt for the general method that includes all firms in the analysis. Despite the small sample size, however, I follow the methodology in the literature and match the failed companies with a similar number of the survivors in terms of total assets. The results are qualitatively similar. Moreover, most of the individual variables that are significant in the non-matched sample become insignificant in the new model since the sample size is significantly reduced.

Given the practice of using variables in their level format in bankruptcy models, I employ first differences. This is done since bankruptcy is a deteriorating process and taking measurements at a certain point in time may not be enough to predict failure. For this reason, first differences, representing the change in variables, may be more appropriate predictors of bankruptcy. The results again show no marked changes from the earlier evidence.

Industry-Specific Models

So far, the analysis is applied to the full sample of observations regardless of the specific industry of the company. But, in fact, important predictors of survival in one industry may be unimportant in another. This perhaps explains why most of the survival analysis studies are conducted over one specific industry. The study sample is distributed over three main industries, namely, manufacturing, retail, and service industries. A detailed breakdown of the number of firms in each category is shown in Table 12 below.

TABLE 12. Industry Classification

Data	Service	Retail	Manufacturing
1998	126	28	10
2000	116	21	10
3 years to event	114	16	10

Due to the small number of manufacturing companies, I choose this as the omitted category and create two dummy variables *CS* and *CR* with *CS*=1 if the company is in the service industry and 0 otherwise, and *CR*=1 if it is in the retail industry and 0 otherwise. Each of the estimated coefficients of these two dummy variables is a contrast to the omitted category, the manufacturing industry. Therefore, I also perform a second test of the null hypothesis that the two estimated coefficients are equal to compare the service and the retail industries survivability.

The evidence from these tests doesn't provide clear evidence of a meaningful difference[5]. Using calendar and event time data, the estimated coefficients of the two dummies are insignificant at the 5% level. However, results from calendar data show that the parameter estimates of the two dummy variables are jointly significant. On the other hand, the event time model still rejects the joint hypothesis that the coefficient estimates are significantly different from zero. Though it is not clear why in the case of calendar data, the joint hypothesis is rejected, it is important to note that the

[5] Results of reduced models, shown in Appendix C, lead to similar conclusions about the sign and significance of the dummy variables as the full models show.

TABLE 13a. Summary of Industry-Specific Cox Regression Results
1998 Data

	Total	Event	Censored	% Censored
	161	21	140	86.96

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	33.7305	8	<.0001
Score	27.9496	8	0.0005
Wald	27.0128	8	0.0007

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	0.23079	0.14904	2.3976	0.1215	1.260
tlta	1	0.35011	0.42756	0.6705	0.4129	1.419
nita	1	-0.17274	0.23608	0.5354	0.4644	0.841
cftl	1	-0.14290	0.08911	2.5715	0.1088	0.867
wcta	1	0.52805	0.65961	0.6409	0.4234	1.696
nsta	1	-0.90794	0.46758	3.7705	0.0522	0.403
cs	1	-0.07365	1.10827	0.0044	0.9470	0.929
cr	1	1.72693	1.10147	2.4581	0.1169	5.623

Label	Wald Chi-Square	DF	Pr>ChiSq
Test 1	12.6879	2	0.0018
Test 2	12.0212	1	0.0005

coefficient on the service industry is negative almost all the time, whereas the coefficient of the second dummy is positive without exception. This leads one to infer that there is a widening gap between the effects of the two industries. The negative coefficient of *CS* suggests that, if significant, the odds of failure for firms in the service industry are lower than that in the manufacturing industry. Conversely, the positive coefficient on *CR* suggests that, if significant, the probability of failure for retailer is higher than that of manufacturers. Unfortunately, calendar data and event time data give mixed results concerning the significance of these industry effects.

TABLE 13b. Summary of Industry-Specific Cox Regression Results
3 Prior Data

Total	Event	Censored	% Censored
138	11	127	92.03

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > Chi Sq
Likelihood Ratio	13.8821	8	0.0849
Score	11.7656	8	0.1620
Wald	12.2746	8	0.1394

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq	Hazard Ratio
ta	1	-0.07315	0.20679	0.1251	0.7235	0.929
tlta	1	0.52482	0.41874	1.5708	0.2101	1.690
nita	1	-0.05411	0.17080	0.1004	0.7514	0.947
cftl	1	-0.06728	0.10997	0.3743	0.5406	0.935
wcta	1	0.38699	0.58296	0.4407	0.5068	1.473
nsta	1	-1.16052	0.61994	3.5044	0.0612	0.313
cs	1	-0.18105	1.13118	0.0256	0.8728	0.834
cr	1	1.17224	1.29144	0.8239	0.3640	3.229

Label	Wald Chi-Square	DF	Pr>ChiSq
Test 1	2.3890	2	0.3029
Test 2	2.3693	1	0.1237

Similar results are obtained concerning the comparison of the odds of failure of firms in the service industry versus those in the retail business. That is, the models based on calendar data suggest that there is a better chance of survival in the service industry than in the retail one. The event time models however show no significant difference between the two. While the results are not conclusive, some inferences can be offered. The results may be partly due to the lower costs associated with the service industry. Moreover, retail businesses need more financing and more careful management of current assets, mainly inventory, especially with seasonal products. The service sector is probably less sophisticated in this regard. In contrast, the latter probably faces more difficulty in providing the best marketing strategies to target a larger share of the market and advertise their products.

Test of Proportional Hazards

Although the time-varying analysis accounts for the possibility that the hazard ratio is time dependent, it provides little information about the shape of the true hazard function. The log-log survival analysis described earlier is a graphical method that is used to test for the assumption. The plot of the survival time, logged twice, should be a straight line if the PH assumption is satisfied. This plot is the equivalence of a cumulative hazard function. So for the hazard to be constant, the cumulative hazard function should increase as a straight line. If the curve is concave then the hazard is said to be decreasing. If the plot curves upward then the hazard is shown to be rising over time.

Plots of log-log survival time are drawn using four different time periods. For brevity, only one is shown below. The slope of the graph is the proportional hazard function. As seen from the figure below (using the 1998 data), the curve is rising at a faster rate across time. This means that the hazard is not constant and increases as time passes, which is consistent with the concept of the deterioration process of bankruptcy. The more severe the financial distress of a company is, the worse is the investor perception about the prospects of the company. As more information about the distress of a company becomes available to consumers, suppliers, and creditors, the more restrictions and monitoring of the firm's activities will result. Despite the consistency of all the graphs about the shape of the hazard function over time, it might be noted that a graphical analysis may not represent a clear-cut method from which one can

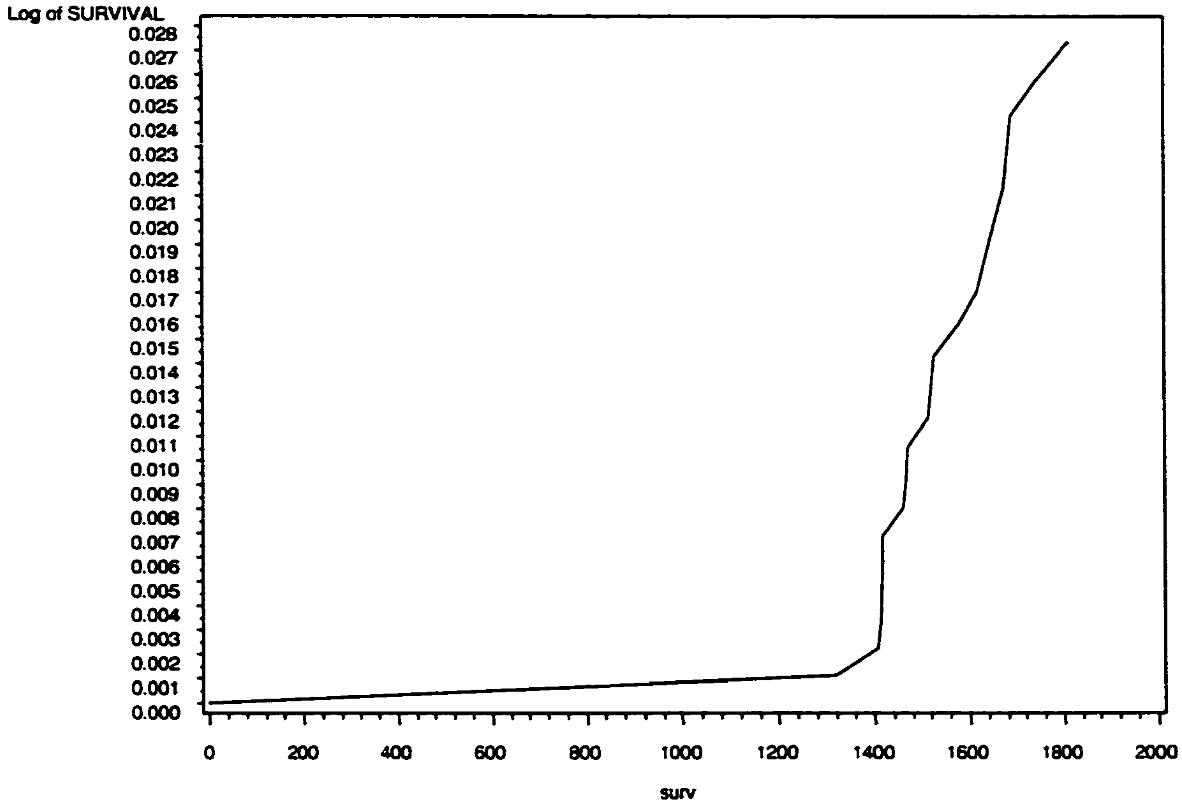

FIGURE 1. Plot of Log-Log Survival Time

can make strong conclusions. For this reason, I also employ a goodness-of-fit test which represents a more robust method. This test is based on the likelihood ratio statistic.

The most commonly suggested models are the exponential and the Weibull distributions. The difference between the two distributions is that the Weibull model is more general, with the exponential being a specific case of the latter through forcing the scale parameter $\sigma=1$. As such, the exponential model is said to be nested in the Weibull model. To test whether the data follows an exponential distribution, I test the null hypothesis that the restriction $\sigma=1$ is true. This is done through taking twice the difference between the likelihoods of both models, resulting in a χ^2 statistic with one degree of freedom. For the 1998 data, the likelihoods of the Weibull and exponential models are -33.57 and -56.39 respectively. This results in a χ^2 statistic of 45.64, thus rejecting the exponential form. Similar conclusions are obtained based on the 2000 data and the 3 years prior to bankruptcy data, with χ^2 statistics equal to 41.74 and 25.44 respectively. The evidence confirms the outcome of the graphical analysis showing that the hazard is not constant and that it actually rises over time.

Logit Regression

The logistic regression technique has been widely used in the bankruptcy literature. Although, the ability of the model is limited to classifying the companies between bankrupt and survivors, I apply this model to the data of dotcoms as a second check on the validity of the survival analysis results. This technique is based on the maximum likelihood estimation of the following model:

$$\log\left(\frac{P_{it}}{1-P_{it}}\right)=\alpha_t+\beta_1 X_{it1}+\ldots+\beta_k X_{itk}$$

where P_{it} is the conditional probability that company i has an event at time t=1,2,3... given that an event has not already occurred to that company. The results shown in appendix D are very similar to that of survival analysis both in terms of significance of the parameters as well as their relative signs. This is consistent with all data sets, time-varying, time-constant, as well as calendar data and event-time data sets.

CHAPTER 5

CONCLUSION

Results from Cross-Sectional Analysis

I use Survival Analysis to identify some important predictors of dotcom failure. The variables CFTL and NSTA are consistently negative and highly significant, at the 5% level or better. This finding is consistent with the literature of bankruptcy. Sales are the main revenues for a firm, and it is expected that higher sales lead to higher revenues and cash flows, which in turn improves the bill paying capability of the firm. Moreover, the results show that these cash flows significantly lower the odds of failure since the coefficient estimate of CFTL is negative.

The results also show that NITA is negative and significant at the 10% level of significance. This is consistent with the above reasoning that higher revenues improve the survivability of a firm. It is important to note however, that revenues and costs are recognized as they occur and not when cash changes hands. Thus, it is possible to have an unexpected, unfavorable change in the timing of cash flows, and this latter variable becomes more important since firms need cash to pay their debt and not simply revenues.

As expected, the coefficient estimate of TLTA is positive, suggesting that higher liabilities increase the odds of failure. This finding is intuitive and reflects the increased risk and higher payments on firm's debt that drain cash flows. However, the coefficient estimate of this variable is insignificant at the 10% level of significance. Thus, it is likely that the effect of this variable could have been already captured by the CFTL. The value of this variable is a direct result of the firm's level of debt outstanding. The net working capital has a negative coefficient estimate that is also insignificant at the 10% level of significance. Higher net working capital reflects the use of long-term debt to finance current assets. There is less risk embedded in such a financing plan. Therefore, lower risk improves the likelihood of survival.

On the other hand, the coefficient estimate on log total assets (TA) is positive and significant at the 5% level. This is in contrast to findings in other bankruptcy studies. My results show that larger firms have higher odds of failure. However, a plausible explanation for such a finding is that these failed dotcoms could have had an unsustainable growth rate in their total assets. Any increase in assets has to be met by a similar increase in sources of funds as represented by liabilities and stockholders equity. The excessive and rapid need for these external sources of funds may raise the concerns of creditors about the financial position of the company and can lead to higher interest rates charged, closer monitoring, and other restrictions.

The above analysis uses calendar data. Findings from such a study may not be generalized and could be specific to a certain era in the life of dotcoms. Therefore, to improve the generalization capability of the model, I measure financial ratios three years prior to bankruptcy of failed companies. The results show consistency in the significance of net sales with the expected negative sign. The variable CFTL also shows negative sign. In addition to that, the coefficient estimate for the variable TA becomes negative. This result is consistent with the literature that the higher the size of the firm, the more immune from bankruptcy it becomes. On the other hand, the event-time model does not show much significance in the overall effect of the regressors.

Results from Time-Dependent Analysis

Results using the years 1999 – 2000 data show that the coefficient estimate on the sales variable is insignificant. This contradicts the findings in the bankruptcy literature-- that higher sales provide higher revenues for the firm and therefore a better chance of survival. It is important to note that the sales volume, by itself, provides little information about the profits of a firm. To adjust for the effect of costs, the net income variable, NITA, can provide better information about the financial condition of a firm. The coefficient estimate on this variable is negative, as expected, and significant at the 5% percent level.

The parameter estimate on CFTL is also significant with the expected negative sign. This is consistent with the findings from the cross-sectional model. The coefficient estimate on the variable TA is significant and exhibits a positive sign. This is also similar to results obtained from the cross-sectional study. Results from an event-time model using measurements one year prior to bankruptcy show similar findings.

Other Considerations

Accounting ratios consist of historical information and may be less useful compared with market variables in predicting the future. Also, differing accounting methods can provide different impressions about financial performance. On the other hand, the financial situation of a firm highly depends on investors' perceptions about the future earnings. To capture this effect of investors' perceptions, I introduce market data into the model. However, the results show that stock returns, RET, add little to the predictive capability of these models. Its coefficient estimate is insignificant, though it exhibits the correct negative sign.

Matching firms by total assets is a common practice in the bankruptcy literature. This method is conducted to eliminate any bias induced by size. Despite the small sample size, however, I follow the methodology in the literature and match the failed companies with a similar number of the survivors in terms of total assets. The results are qualitatively similar. However, most of the individual variables that are significant in the non-matched sample become less significant in the new model.

In addition, the most commonly suggested predictors of failure may not apply to different industries. Some important predictors of survival in one industry may be unimportant in another. For this reason, I follow the common practice of confining the firms under study to specific sectors, namely, retail, service and manufacturing. Further, I create two dummy variables representing the service and the retail industries. The evidence from these tests does not provide clear-cut signals. Using calendar and event time data, the estimated coefficients of the two dummies are insignificant at the 5% level. However, these models provide mixed results about the joint significance of the dummy variables. While the results are not conclusive, it is important to note that these industries have different characteristics, such as costs, financing needs, current asset management, and advertising strategies. These indicators that shape each of the stated industries may player a more important role in the survivability of firms, perhaps under different economic circumstances.

Policy Implications

It is important to note that the coefficient estimate on the cash flow variable is consistently negative and significant at the 5% level or better. This variable represents the solvency of the firm and therefore enhances the likelihood of survival. Therefore financial managers need to examine cash flows more carefully and manage their timing to provide more synchronization between bill payments and actual receipt of cash. This could be achieved through better examination of credit standards, terms of trade and perhaps, improved collection policies. Moreover, cash flows are directly related to all other indicators such as total liabilities, net working capital, sales, and net income. Thus, a careful management of these aspects of a firm should ultimately maximize cash flows and expedite the receipt of cash.

On the other hand, larger firms appear to be more exposed to risk of bankruptcy. This finding suggests that financial managers should slow down the process of acquiring external funds. As a firm's growth in terms of total assets accelerates, its need for funds to finance this growth also accelerates. Such funds are more than likely to come from external sources. This rapid growth may raise concerns of creditors and investors about the financial risk of the firm. Such perceptions can lead to higher cost of capital and therefore a decline in shareholders wealth.

Recommendation for Future Research

In this study, the information used is limited to publicly traded companies. This results in a bias as the sample does not incorporate all categories of Internet firms. Although it is more difficult to gather data on non-publicly traded dotcoms, findings of such extended study can be generalized to cover the entire Internet sector.

Moreover, further research may expand this sample to include foreign companies that are active outside the US.

While the frequency of the data is annual, an issue of concern is the short time span that this study covers. A follow-up study may expand this time frame to cover more years. Findings from such a study could, perhaps, be more appealing in that it can be more generalized.

APPENDIX A

DESCRIPTIVE STATISTICS

APPENDIX A

DESCRIPTIVE STATISTICS

TABLE 14a. Descriptive Statistics for the Year 1998 Data

1998 Data						
Descriptive Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ret	59	0.67026	2.48011	39.54507	-0.96861	12.00000
ta	170	1.55601	2.01629	264.52131	-6.93925	6.48281
tlta	169	1.25673	2.63459	212.38748	0	21.52569
nita	164	-1.94090	5.79079	-318.30723	-63.26877	0.42342
cftl	161	-1.89953	2.74490	-305.82395	-14.21596	9.66667
wcta	170	-0.33939	2.49934	-57.69682	-20.11858	1.00000
nsta	164	1.97616	7.46516	324.09106	0	89.89362

TABLE 14b. Descriptive Statistics for the Year 2000 Data

2000 Data						
Descriptive Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ret	128	-0.28636	4.17608	-36.65381	-0.99758	46.04133
ta	157	3.37341	1.91159	529.62561	-1.60944	7.59864
tlta	157	0.59729	1.02139	93.77421	0.01064	10.17857
nita	156	-1.57347	2.88514	-245.46062	-18.49000	0.42923
cftl	154	-3.38064	7.02981	-520.61821	-74.84561	2.10905
wcta	155	0.06253	1.08407	9.69289	-9.86286	0.98632
nsta	156	0.89588	1.52841	139.75792	0	13.73137

APPENDIX B

PEARSON CORRELATION COEFFICIENTS

APPENDIX B

PEARSON CORRELATION COEFFICIENTS

TABLE 15a. Pearson Correlation Matrix for the Year 1998 Data

1998 Data							
Pearson Correlation Coefficients							
	ret	ta	tlta	nita	cftl	wcta	nsta
ret	1.00000						
ta	-0.02063	1.00000					
tlta	-0.09394	-0.34197	1.00000				
nita	0.07262	0.40674	-0.76411	1.00000			
cftl	-0.12188	0.11494	0.10488	0.09155	1.00000		
wcta	0.10295	0.35029	-0.94110	0.78499	-0.08461	1.00000	
nsta	-0.10909	-0.27596	0.31565	-0.25285	0.09505	-0.34755	1.00000

TABLE 15b. Pearson Correlation Matrix for the Year 2000 Data

2000 Data							
Pearson Correlation Coefficients							
	ret	ta	tlta	nita	cftl	wcta	nsta
ret	1.00000						
ta	-0.00630	1.00000					
tlta	-0.04489	-0.44591	1.00000				
nita	0.04594	0.46518	-0.66943	1.00000			
cftl	0.01004	0.13921	0.08840	0.32801	1.00000		
wcta	0.07672	0.42750	-0.96121	0.69127	-0.10271	1.00000	
nsta	-0.05661	-0.17904	0.10078	0.00654	0.07554	-0.04151	1.00000

APPENDIX C

INDUSTRY SPECIFIC COX

REGRESSION RESULTS

APPENDIX C

INDUSTRY SPECIFIC COX

REGRESSION RESULTS

TABLE 16a. Industry-Specific Cox Regression Results for the Year 1998. Reduced Model with TLTA Retained

1998 Data

Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	32.9021	6	<.0001
Score	27.1845	6	0.0001
Wald	27.2384	6	0.0001

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	0.23163	0.14881	2.4228	0.1196	1.261
tlta	1	0.29141	0.16730	3.0342	0.0815	1.338
cftl	1	-0.17882	0.07288	6.0204	0.0141	0.836
nsta	1	-1.00186	0.43149	5.3909	0.0202	0.367
cS	1	0.06238	1.09165	0.0033	0.9544	1.064
cR	1	1.88470	1.08101	3.0397	0.0813	6.584

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	13.6151	2	0.0011
Test 2	12.6184	1	0.0004

TABLE 16b. Industry-Specific Cox Regression Results for the Year 1998. Reduced Model with NITA Retained

1998 Data

Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	32.9493	6	<.0001
Score	27.4592	6	0.0001
Wald	27.1291	6	0.0001

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	0.23442	0.14941	2.4617	0.1167	1.264
nita	1	-0.10288	0.05707	3.2492	0.0715	0.902
cftl	1	-0.15637	0.07174	4.7507	0.0293	0.855
nsta	1	-0.89378	0.44430	4.0467	0.0443	0.409
cS	1	-0.10511	1.09142	0.0093	0.9233	0.900
cR	1	1.70142	1.06688	2.5433	0.1108	5.482

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	13.2061	2	0.0014
Test 2	12.3316	1	0.0004

TABLE 16c. Industry-Specific Cox Regression Results for the Year 1998. Reduced Model with WCTA Retained

1998 Data

Total	Event	Censored	% Censored
161	21	140	86.96

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	32.1636	6	<.0001
Score	27.0976	6	0.0001
Wald	26.8364	6	0.0002

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	0.22009	0.15046	2.1397	0.1435	1.246
cftl	1	-0.17347	0.07337	5.5898	0.0181	0.841
wcta	1	-0.23814	0.17195	1.9180	0.1661	0.788
nsta	1	-0.96494	0.43247	4.9784	0.0257	0.381
cS	1	-0.02649	1.09061	0.0006	0.9806	0.974
cR	1	1.78863	1.07338	2.7767	0.0956	5.981

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	13.5665	2	0.0011
Test 2	12.6412	1	0.0004

TABLE 17a. Industry-Specific Cox Regression Results For the Data 3 Years Prior To Failure. Reduced Model with TLTA Retained

3 Prior Data

Total	Event	Censored	% Censored
138	11	127	92.03

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	13.4312	6	0.0367
Score	9.3748	6	0.1536
Wald	11.8310	6	0.0658

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	-0.08144	0.21088	0.1491	0.6994	0.922
tlta	1	0.27259	0.14919	3.3384	0.0677	1.313
cftl	1	-0.07059	0.10052	0.4932	0.4825	0.932
nsta	1	-1.25009	0.56987	4.8122	0.0283	0.286
cs	1	-0.28221	1.11357	0.0642	0.7999	0.754
cR	1	1.07654	1.27329	0.7148	0.3978	2.935

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	2.4668	2	0.2813
Test 2	2.4646	1	0.1164

TABLE 17b. Industry-Specific Cox Regression Results for the Data 3 Years Prior to Failure. Reduced Model with NITA Retained

3 Prior Data

Total	Event	Censored	% Censored
138	11	127	92.03

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12.7391	6	0.0474
Score	11.1180	6	0.0848
Wald	11.3807	6	0.0773

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	-0.13296	0.20526	0.4196	0.5171	0.876
nita	1	-0.08151	0.05198	2.4595	0.1168	0.922
cftl	1	-0.03872	0.10009	0.1496	0.6989	0.962
nsta	1	-1.12185	0.62804	3.1908	0.0741	0.326
cs	1	-0.50045	1.10376	0.2056	0.6503	0.606
cR	1	0.87595	1.25084	0.4904	0.4837	2.401

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	2.5554	2	0.2787
Test 2	2.5405	1	0.1110

TABLE 17c. Industry-Specific Cox Regression Results for the Data 3 Years Prior to Failure. Reduced Model with WCTA Retained

3 Prior Data

Total	Event	Censored	% Censored
138	11	127	92.03

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12.7391	6	0.0474
Score	11.1180	6	0.0848
Wald	11.3807	6	0.0773

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
ta	1	-0.12828	0.20721	0.3833	0.5359	0.880
cftl	1	-0.05775	0.09965	0.3359	0.5622	0.944
wcta	1	-0.21976	0.14370	2.3387	0.1262	0.803
nsta	1	-1.25211	0.59364	4.4488	0.0349	0.286
cS	1	-0.44717	1.11045	0.1622	0.6872	0.639
cR	1	0.93127	1.26301	0.5437	0.4609	2.538

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	2.5960	2	0.2731
Test 2	2.5879	1	0.1077

APPENDIX D

DATA

APPENDIX D

DATA

TABLE 18a. Summary of Logistic Regression Results for the Year 2000. Full Model

2000 Data

Log Likelihood	-63.90832552

Value	bankrpt	Frequency
1	1	14
2	2	744

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-6.1240	1.0772	-8.2352	-4.0127	32.32	<.0001
ta	1	0.2629	0.1782	-0.0864	0.6123	2.18	0.1402
tlta	1	0.7586	0.9556	-1.1144	2.6315	0.63	0.4273
nita	1	-0.1939	0.0857	-0.3618	-0.0259	5.12	0.0237
cftl	1	-0.0287	0.0211	-0.0701	0.0128	1.84	0.1753
wcta	1	0.9462	0.9730	-0.9608	2.8532	0.95	0.3308
nsta	1	0.0929	0.1779	-0.2558	0.4416	0.27	0.6014

TABLE 18b. Summary of Logistic Regression Results for the Year 2000.
Reduced Model with TLTA Retained

2000 Data

Log Likelihood	-66.12820491

Value	bankrpt	Frequency
1	1	14
2	2	754

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-5.2393	0.8796	-6.9633	-3.5152	35.48	<.0001
ta	1	0.2145	0.1733	-0.1252	0.5542	1.53	0.2159
tlta	1	0.1644	0.2685	-0.3619	0.6907	0.37	0.5403
cftl	1	-0.0534	0.0163	-0.0853	-0.0216	10.80	0.0010
nsta	1	0.1152	0.1552	-0.1891	0.4194	0.55	0.4582

TABLE 18c. Summary of Logistic Regression Results for the Year 2000.
Reduced Model with NITA Retained

2000 Data

Log Likelihood	-64.55506003

Value	bankrpt	Frequency
1	1	14
2	2	754

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-5.7498	0.8890	-7.4922	-4.0074	41.83	<.0001
ta	1	0.3045	0.1757	-0.0399	0.6489	3.00	0.0831
nita	1	-0.1514	0.0726	-0.2938	-0.0090	4.34	0.0371
cftl	1	-0.0398	0.0171	-0.0734	-0.0063	5.41	0.0201
nsta	1	0.1287	0.1695	-0.2035	0.4609	0.58	0.4476

TABLE 18d. Summary of Logistic Regression Results for the Year 2000. Reduced Model with WCTA Retained

2000 Data

Log Likelihood	-66.06278678

Value	bankrpt	Frequency
1	1	14
2	2	754

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-5.0744	0.8005	-6.6433	-3.5054	40.18	<.0001
ta	1	0.2022	0.1737	-0.1383	0.5426	1.35	0.2445
cftl	1	-0.0530	0.0167	-0.0856	-0.0203	10.11	0.0015
wcta	1	-0.1178	0.2889	-0.6839	0.4484	0.17	0.6836
nsta	1	0.1162	0.1559	-0.1894	0.4218	0.56	0.4560

TABLE 19a. Summary of Logistic Regression Results for the Year 1998. Reduced Model with TLTA Retained

1998 Data

Log Likelihood	-87.32111903

Value	bankrpt	Frequency
1	1	21
2	2	774

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-4.3894	0.6615	-5.6860	-3.0929	44.03	<.0001
ta	1	0.3600	0.1497	0.0665	0.6535	5.78	0.0162
tlta	1	0.2883	0.1814	-0.0672	0.6437	2.53	0.1120
cftl	1	-0.1561	0.0722	-0.2976	-0.0146	4.68	0.0306
nsta	1	-1.0449	0.4877	-2.0007	-0.0891	4.59	0.0321

TABLE 19b. Summary of Logistic Regression Results for the Year 1998. Reduced Model with NITA Retained

1998 Data

Log Likelihood	-87.0079749

Value	bankrpt	Frequency
1	1	21
2	2	774

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-4.3709	0.6506	-5.6462	-3.0957	45.13	<.0001
ta	1	0.3778	0.1518	0.0804	0.6753	6.20	0.0128
nita	1	-0.1157	0.0610	-0.2351	0.0038	3.60	0.0578
cftl	1	-0.1326	0.0702	-0.2702	0.0049	3.57	0.0588
nsta	1	-0.9211	0.5003	-1.9017	0.0595	3.39	0.0656

TABLE 19c. Summary of Logistic Regression Results for the Year 1998. Reduced Model with WCTA Retained

1998 Data

Log Likelihood	-87.6029485

Value	bankrpt	Frequency
1	1	21
2	2	774

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-4.1373	0.6055	-5.3240	-2.9506	46.69	<.0001
ta	1	0.3522	0.1520	0.0543	0.6502	5.37	0.0205
cftl	1	-0.1521	0.0728	-0.2948	-0.0093	4.36	0.0368
wcta	1	-0.2465	0.1864	-0.6118	0.1189	1.75	0.1861
nsta	1	-0.9957	0.4943	-1.9645	-0.0268	4.06	0.0440

TABLE 20a. Summary of Logistic Regression Results for the Data 3 years Prior to Failure. Reduced Model with TLTA Retained

3 Prior Data

Log Likelihood -51.07964304

Value	bankrpt	Frequency
1	1	11
2	2	677

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-3.8273	0.7273	-5.2527	-2.4019	27.70	<.0001
ta	1	0.0085	0.2136	-0.4101	0.4271	0.00	0.9684
tlta	1	0.2669	0.1616	-0.0498	0.5836	2.73	0.0985
cftl	1	-0.0681	0.0988	-0.2618	0.1256	0.47	0.4909
nsta	1	-1.2082	0.5996	-2.3834	-0.0331	4.06	0.0439

TABLE 20b. Summary of Logistic Regression Results for the Data 3 Years Prior to Failure. Reduced Model with NITA Retained

3 Prior Data

Log Likelihood -51.18865893

Value	bankrpt	Frequency
1	1	11
2	2	677

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-3.6149	0.6773	-4.9424	-2.2874	28.48	<.0001
ta	1	-0.0291	0.2061	-0.4331	0.3749	0.02	0.8877
nita	1	-0.0930	0.0584	-0.2075	0.0215	2.53	0.1114
cftl	1	-0.0317	0.0974	-0.2225	0.1591	0.11	0.7449
nsta	1	-1.0994	0.6613	-2.3955	0.1967	2.76	0.0964

TABLE 20c. Summary of Logistic Regression Results for the Data 3 Years Prior to Failure. Reduced Model with WCTA Retained

3 Prior Data

Log Likelihood	-51.34906838

Value	bankrpt	Frequency
1	1	11
2	2	677

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-3.4794	0.6280	-4.7102	-2.2485	30.70	<.0001
ta	1	-0.0308	0.2095	-0.4415	0.3799	0.02	0.8833
cftl	1	-0.0552	0.0978	-0.2468	0.1365	0.32	0.5726
wcta	1	-0.2352	0.1610	-0.5508	0.0805	2.13	0.1442
nsta	1	-1.2144	0.6236	-2.4366	0.0079	3.79	0.0515

TABLE 21. Summary of Logistic Regression Results for the 1999 – 2000 Data. Reduced Model with NITA Retained

99 - 00 Data

Log Likelihood	-71.40137935

Value	bankrpt	Frequency
1	1	20
2	2	325

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-4.4646	0.7251	-5.8859	-3.0434	37.91	<.0001
ta	1	0.3312	0.1480	0.0411	0.6213	5.01	0.0252
nita	1	-0.0903	0.0481	-0.1845	0.0039	3.53	0.0603
cftl	1	-0.0855	0.0412	-0.1662	-0.0048	4.31	0.0379
nsta	1	0.1038	0.1557	-0.2013	0.4089	0.44	0.5050

TABLE 22a. Summary of Logistic Regression Results for the Data One-Year Prior to Failure. Full Model

TV 1 Prior Data

Log Likelihood -91.50839949

Value	bankrpt	Frequency
1	1	21
2	2	885

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-5.6222	0.7197	-7.0327	-4.2116	61.03	<.0001
ta	1	0.2856	0.1408	0.0097	0.5615	4.12	0.0425
tlta	1	0.2213	0.2893	-0.3457	0.7884	0.59	0.4443
nita	1	-0.3831	0.1398	-0.6571	-0.1091	7.51	0.0061
cftl	1	-0.0318	0.0500	-0.1299	0.0662	0.40	0.5245
wcta	1	0.8977	0.4542	0.0076	1.7878	3.91	0.0481
nsta	1	0.0217	0.0730	-0.1213	0.1646	0.09	0.7666

TABLE 22b. Summary of Logistic Regression Results for the Data One-Year Prior to Failure. Reduced Model with NITA Retained

TV 1 Prior Data

Log Likelihood -94.8040717

Value	bankrpt	Frequency
1	1	21
2	2	899

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-5.0165	0.5908	-6.1744	-3.8586	72.11	<.0001
ta	1	0.2556	0.1339	-0.0069	0.5180	3.64	0.0563
nita	1	-0.0978	0.0455	-0.1869	-0.0087	4.63	0.0314
cftl	1	-0.0843	0.0351	-0.1530	-0.0156	5.79	0.0162
nsta	1	0.0090	0.0548	-0.0984	0.1163	0.03	0.8702

REFERENCES

Altman, E. I. (1984), "A Further Empirical Investigation of the Bankruptcy Cost Question." *Journal of Finance*, pp. 1067-1089.

Altman, E. I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance*, pp. 589-609.

Altman, R. Halderan and P. Narayman (1977), "Zeta Analysis, A New Model to Identify Bankruptcy Risk of Corporations." *Journal of Banking and Finance*, pp. 29-54.

Andertsch D.B. (1991), "New-Firm Survival and the Technological Regime." The Review of Economics and Statistics, pp. 441-450.

Andertsch, D. B. and T. Mahmood (1995), "New Firm Survival: New Results Using a Hazard Function." The Review of Economics and Statistics, pp. 97-103.

Aziz, A., Emanuel D. C. and G. H. Lawson (1988), "Bankruptcy Prediction-An Investigation of Cash Flow Based Models." *Journal of Management Studies*, pp. 419-437.

Aziz A., and G. H. Lawson (1989), "Cash Flow Reporting and Financial Distress." *Financial Management*, pp. 55-63.

Beaver, W. H. (1966), "Financial Ratios as Predictors of Failure." *Empirical Research in Accounting*, selected studies, pp. 71-111.

Beaver (1968a), "Alternative Accounting Measures As Predictors of Failure." *The Accounting Review*, pp. 113-122.

Beaver (1986b), "Market Prices, Financial Ratios, and the Prediction of Failure." *Journal of Accounting Research*, pp. 179-192.

Bharat A. J. and O. Kini (2000), "Does the Presence of Venture Capitalists Improve the Survival Profile of IPO Firms?." *Journal of Business Finance & Accounting*, pp. 1139-1176.

Blair, R. D., and A. A. Heggestad (1978), "Bank Portfolio Regulation and the Probability of Bank Failure." *Journal of Money, Credit, and Banking*, pp. 88-95.

Blum (1974), "Failing Company Discriminant Analysis." *Journal of Accounting Research*, pp. 1-25.

Bovenzi J. F, Marino J. A., and F. E. McFadden (1983), "Commercial Bank Failure Prediction Models." *Economic Review*, pp. 14-26.

Casey C. and N. Bartczak (1985), "Using Operational Cash Flow Data to Predict Financial Distress: Some Extensions." Journal of Accounting Research, pp. 384-401.

Chen, K. W. and C. J. Lee (1993), "Financial Ratios and Corporate Endurance: A Case of the Oil and Gas Industry." *Contemporary Accounting Research*, pp. 667-694.

Cox, D. R. (1972), "Regression Models and Life Tables." *Journal of the Royal Statistical Society*, pp. 187-201.

Deakin (1972), "A Discriminant Analysis of Predictors of Business Failure." *Journal of Accounting Research*, pp.167-179.

Dimitras A. I. (1996), "A Survey of Business Failures with an Emphasis on Prediction Methods and Industrial Applications." European Journal of Operational Research, pp. 487-513.

Edmister, R. O. (1972) "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction." *Journal of Financial and Quantitative Analysis*, pp. 1477- 1493.

Eisenbeis, R. A. (1977), "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics." *Journal of Finance*, pp. 875-900.

Evans (1987), "The Relationship between Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries." *Journal of Industrial Economics*, pp. 567-581.

Gentry et al. (1985), "Classifying Bankrupt Firms with Funds Flow Components." Journal of Accounting Research, pp. 146-60.

Ho, T., and A. Saunders (1980), "A Catastrophe Model of Bank Failure." *The Journal of Finance*, pp. 1189-1196.

Huyghebaert, N., A. Gaeremynck, F. Roodhooft, and L. M. Van De Gucht (2000), "New Firm Survival: The Effects of Start-up Characteristics." *Journal of Business Finance and Accounting*, pp. 627-651.

Jagtiani J. and C. Lemieux (2001), "Market Discipline Prior to Bank Failure." *Journal of Economics and Business*, pp. 313-324.

Johnson C. G. (1970), "Ratio Analysis and the Prediction of Firm Failure." *The Journal of Finance*, pp. 1166-1168.

Kleinbaum, D. G. (1996), *Survival Analysis*. Springer-Verlag New York, Inc.

Laitinen (1994), "Traditional Versus Operating Cash Flow In Failure Prediction." *Journal of Business Finance and Accounting*, pp. 195-217.

Laitinen (1992), "Prediction of Failure of a Newly Founded Firm." *Journal of Business Venturing*, pp. 323-340.

Lane, W. R., S. W. Looney and J. W. Wansley (1986), "An Application of the Cox Proportional Hazards Model to Bank Failure." *Journal of Banking and Finance*, pp. 511- 531.

Libby (1975), "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence." Journal of Accounting Research, pp. 150-161.

Mata J. and P. Portugal (1994), "Life Duration of New Firms." *Journal of Industrial Economics*, pp. 227-245.

Meyer, P. and H. Pifer (1970), "Prediction of Bank Failures." *Journal of Finance*, pp. 853-868.

Moyer (1977), "Forecasting Financial Failure: A Re-examination." *Financial Management*, pp. 11-17.

Ohlson, J. A. (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research*, 109-131.

Opler and Titman (1994), "Financial Distress and Corporate Performance." *Journal of Finance*, pp. 1015-1040.

Platt H. D. and M. B. Platt (1994), "Business Cycle Effects on State Corporate Failure Rates." Journal of Economics and Business, pp. 113-127.

Queen, M., and R. Roll (1987), "Firm Mortality: Using Market Indicators to Predict Survival." *Financial Analysts Journal*, pp. 9-26.

Robertson J. (1983), "Company Failure: Measuring Changes in Financial Health Through Ratio Analysis." *Management Accounting*, pp. 24-28.

Simpson, W. G. (1983), "Capital Market Prediction of Large Commercial Bank Failures: An Alternative Analysis." *The Financial Review*, pp. 33-55.

Tveteras R. and G. E. Eide (2000), "Survival of New Plants in Different Industry Environments in Norwegian Manufacturing: A Semi-Proportional Cox Model Approach." *Small Business Economics*, pp. 65-82.

Wilcox (1973), "A Prediction of Business Failure Using Accounting Data." *Empirical Research in Accounting*, pp. 163-190.

Zavgren (1985), "Assessing the Vulnerability of Failure of American Industrial Firms: A Logistic Approach." *Journal of Business Finance and Accounting*, pp. 19-45.

Zmijemski (1984), "Methodological Issues Related to the Estimation of Financial Distress Models." *Journal of Accounting Research*, pp. 59-82.