LIBRARY
[redacted]
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

NPS-54-86-011

NAVAL POSTGRADUATE SCHOOL
Monterey, California

TESTS OF THE USEFULNESS OF ANALYST EARNINGS FORECAST DATA IN PREDICTING BANKRUPTCY OF PUBLIC CORPORATIONS

by

O. Douglas Moses

October 1986

Approved for public release; distribution unlimited.

Prepared for: Chief of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
Monterey, California

RADM. R. C. Austin
Superintendent

David A. Schrady
Provost

The research summarized herein was supported by the Foundation Research Program of the Naval Postgraduate School with funds provided by the Chief of Naval Research.

This report was prepared by:

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): NPS-54-86-011	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION: Naval Postgraduate School; 6b OFFICE SYMBOL (If applicable): 54	7a NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State, and ZIP Code): Monterey, CA. 93943	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING/SPONSORING ORGANIZATION: Naval Postgraduate School; 8b OFFICE SYMBOL (If applicable): 54	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N0001486WR4E001
8c ADDRESS (City, State, and ZIP Code): Monterey, CA 93943	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61152N	RR000-01	10P3-009	

11 TITLE (Include Security Classification)
Tests of the usefulness of analyst earnings forecast data in predicting bankruptcy of public corporations (unclassified)

12 PERSONAL AUTHOR(S)
O. Douglas Moses

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM Oct 85 TO Oct 86	October 1986	124

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Bankruptcy Prediction Earnings Forecasts

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This study investigates five properties of earnings forecasts made by financial analysts to determine if systematic differences in these properties exists between failing and healthy firms. The five properties are: The level of forecasts, forecast error, forecast bias, forecast dispersion and revisions in forecasts. Measures reflecting the five properties are used in models to distinguish failing and healthy firms and predict future bankruptcy.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL: O. Douglas Moses	22b TELEPHONE (Include Area Code): 408-646-3218; 22c OFFICE SYMBOL: 54

Tests of the Usefulness of Analyst Earnings Forecast Data in Predicting Bankruptcy of Public Corporations.

by

Douglas Moses

Department of Administrative Sciences

Naval Postgraduate School

Monterey, CA. 93943

October 1986

Acknowledgements

Data on earnings forecasts used in this study was taken from the Institutional Brokers Estimate System (IBES) published by Lynch, Jones, & Ryan, New York. Access to historic IBES data provided by Lynch, Jones, & Ryan is greatly appreciated.

Support for this project was provided by the Naval Postgraduate School Foundation Research Program, which is funded by the Office of Naval Research and the Office of Naval Technology.

Contents

EXECUTIVE SUMMARY

The purpose of this study is to determine if various measures developed from financial analysts forecasts of earnings for firms can be exploited in predicting future bankruptcy. The analysis consists of two major parts.

In the first part the properties of analysts forecasts are discussed and measures are developed to reflect these properties. Five properties are investigated: forecast level, forecast dispersion, forecast error, forecast bias and forecast revisions. Various tests are conducted to determine if there are systematic differences in the properties for failing firms as compared to healthy firms in years prior to the bankruptcy of the failing firms. Several statistically significant differences are apparent. Failing firms tend to be associated with lower forecasted earnings, higher dispersion in earnings forecasts across multiple forecasters, greater error in forecasts, over-optimistic forecasts, and greater frequency of downward revisions in forecast estimates. Some differences between failing and healthy firms in how the properties change, both within years and across years, are also apparent.

In the second part of the study, measures reflecting aspects of the five properties are used to discriminate failing from healthy firms. Both univariate and multivariate approaches are attempted. Single measures and linear combinations of measures are able to out-predict a naive model, which classifies all firms as

healthy, in distinguishing between groups. However, overall results are not encouraging. It is possible to develop multivariate models that are highly successful in classifying firms but these models are unstable when applied to data taken from different years prior to bankruptcy and are of questionable validity. Individual measures are more stable across time in identifying failing firms from healthy firms but classification success is not impressive. Perhaps the single best approach identified for using measures taken from analyst forecasts for predicting failure is to look at the mean forecasted value of future earnings. Simply put failing firms are predicted to have lower earnings. An approach in which firms with forecasted earnings below a particular cutoff point are classified as facing bankruptcy is able to correctly classify from 22% to 40% (depending on the year prior to bankruptcy) of firms that are misclassified using the naive rule. While the results indicate some ability of earnings forecast data to assist in the prediction of failure, the ability to exploit analyst forecast data to assist in predicting failure does not appear to be great. The fact that there is an association between low forecasted future earnings and future bankruptcy is not surprising and does not appear to provide any novel insight that may be exploited for predicting bankruptcy.

CHAPTER I

INTRODUCTION AND OVERVIEW

1.1 Introduction

Corporate bankruptcy, failure or distress can result in considerable costs to management, investors, creditors and customers. The prediction of corporate failure ex ante can provide the time to react and minimize those costs. The most common source of information for assessing financial health and developing models to predict failure is corporate accounting reports. Several studies have used statistical techniques to assess the ability of combinations of accounting ratios to predict bankruptcy. Variables used in prediction models typically included liquidity and solvency ratios, and performance and funds flow measures. These studies used different ratios and different analytical techniques, and achieved varying levels of predictability; but each relied on accounting information. (See Zavgren [1983] for a review.)

There are, however, several weakness to the use of accounting data to predict corporate failure. Accounting data is produced only periodically, is historical rather than prospective, and reflects events that are primarily endogenous to the firm. Accounting measures are sensitive to the choice of accounting procedures, subject to "window dressing", and inevitable vary in magnitude across firms and industries as a function of the nature of operations and technology. In addition, because of interrelationships between measures, researchers have found that individual ratios are inconsistent predictors across tests and samples.

This study investigates the usefulness of another approach to the prediction of corporate failure, one involving the use of non-accounting information, specifically financial analysts forecasts (FAF) of a firm's future earnings. The purpose is to see if measures developed from analysts forecasts of earnings can be exploited to predict bankruptcy.

1.2 Background

Earnings are considered by investors and analysts to be the most preferred expectational data (Change and Most [1980]) and have the greatest information content of various accounting variables (Gonedes [1974]); thus there tends to be special importance attached to the information reflected in earnings. Various studies of financial analysts forecasts of earnings have been conducted (See Givoly and Lakonishok [1984] for review). Several qualities of FAF suggest their usefulness as an information source and their potential ability to aid in failure prediction. FAF tend to outperform mechanical models based on past historical earnings in predicting future earnings (Barefield and Comiskey [1975]; Collins and Hopwood [1980]; Fried and Givoly [1982]). This superiority is more pronounced in years where there is a turning point in the earnings trend (Barefield and Comiskey). FAF apparently contain information not captured by historical trends in earnings (Fried and Givoly) and may reflect inside information (Abdel-khalik and Ajinkya [1982]). Analysts revise their forecasts in response to information contained in quarterly earnings announcements (Brown and Rozeff [1979]) but the trend of FAF is smoother than actual trends (Crichfield, Dyckman and Lakonishok [1978]), suggesting that analysts separate a permanent from a temporary component in reported earnings numbers. Studies have indicated an association of FAF and revisions in FAF with stock prices (Neiderhoffer and Regan [1972], Givoly and Lakonishok [1979,1980], Elton, Gruber and Gultekin [1981], Brown, Foster and Noreen [1985]). Securities trading strategies using FAF and revisions in FAF indicate that FAF have information content for the securities market (Givoly and Lakonishok [1980], Abdel-khalik and Ajinkya). Furthermore, FAF appear to be a more adequate surrogate for the securities market earnings expectations than are naive predictions based on historical earnings (Malkiel [1970], Malkiel and Cragg [1970], Fried and Givoly [1982]). Collectively these findings indicate that FAF are a useful, comprehensive piece of information which reflect information exogenous to firms' accounting systems.

Of particular interest in the context of bankruptcy prediction are measures of risk derived from FAF. The error in earnings forecasts has been shown analytically to be an appropriate indicator of uncertainty (Cukierman and Givoly [1982]). The dispersion of forecasts across analysts and the unpredictability of earnings have been shown empirically to be associated with traditional risk measures such as beta and the standard deviation of returns (Givoly and Lakonishok [1983]). In addition the dispersion of FAF has been shown to be superior to measures of beta, economy risk, information risk, and interest rate risk in explaining expected return (Malkiel [1981]). In short, dispersion and unpredictability in FAF may serve as useful proxies for risk. Such measures may be of "unique value to empirical researchers" because unlike most traditional risk measures, these are "ex ante" measures of risk (Givoly and Lakonishok [1984]).

In summary, past attempts to predict corporate failure have in general relied on accounting data, which is historical, reflective of information primarily endogenous to the firm, subject confounding influences such as manipulation and the choice of accounting procedures, and provided only periodically. Financial analysts forecasts are prospective, reflective of a broad information set, and provided and revised in a timely manner. FAF can be expected to reflect macro-economic events, industry expectations and firm-specific non-accounting information (e.g. contracts, order backlogs, capital expenditures). Research has indicated that FAF and risk measures developed from FAF have useful information content.

1.3 Relevance to Department of Defense

The Department of defense continually contracts with corporations to provide material, equipment and services. Corporate failure or distress frequently leads to contract terminations, which can impose direct costs on the Department of Defense in the form of increased legal and administrative expenditures, and indirect costs in the form of delays or failure to perform contracted activities. The ability to predict or anticipate corporate distress has obvious implications for contracting with corporate organizations, the procurement of goods and services from the private sector, and the management of existing contracts.

1.4 Preview and organization of the study

The objective of this study is to empirically investigate the potential usefulness of measures developed from financial analysts forecasts of earnings in predicting corporate bankruptcy. In general the approach used is to identify a sample of failed firms and a matched sample of non-failed firms, to create measures of various properties of analysts earnings forecasts, and to test the ability of the measures to discriminate between the two groups of firms. The remainder of the study consists of three chapters. Chapter 2 provides information on sample firms and sample selection procedures, and information about the data source used to develop measures related to analysts forecasts of earnings.

Chapter 3 draws on the results of previous research to identify properties of analysts forecasts. Five properties of forecasts are addressed: forecast level, forecast error, forecast bias, forecast dispersion, and revisions in forecasts. Measures reflecting each property are developed. Tests are conducted to determine if these five properties of forecasts differ systematically between failing and healthy firms in years prior to bankruptcy. Tests are also conducted to determine if there are systematic differences in the way measures reflecting these properties change both within years and across years. The identification of systematic differences in properties of forecasts provide the basis for assuming that measures reflecting these properties may be useful in distinguishing failing from healthy firms.

Chapter 4 uses a subset of measures in an attempt to distinguish failing from healthy firms. Both univariate and multivariate approaches to discrimination are used. In the univariate approach sample firms are rank ordered on individual measures and a threshold or cutoff value of a measure is selected to classify firms into groups. In the multivariate approach stepwise multiple discriminant analysis is used to construct linear combinations of measures that provide a discriminant score to be used in classifying firms into groups. Various validation tests of the approaches are presented and conclusions concerning the overall usefulness of analyst forecast information in predicting bankruptcy are offered.

CHAPTER 2

DATA AND SAMPLE INFORMATION

2.1 The Data Source - IBES

The data source for analyst earnings forecasts was the Institutional Brokers Estimate System (IBES) published by Lynch, Jones, and Ryan, a New York based brokerage firm. IBES is available in both manual and machine-readable form in monthly issues. An historical summary data tape covering each month from January 1976 through July 1985 was made available by Lynch, Jones, and Ryan. Earnings forecast data for 4305 firms were available on the IBES tape. However, the period covered on the tape for individual firms ranged from one month to the maximum possible nine years, six months.

IBES contains summary statistics related to annual earnings-per-share forecasts up to two years prior to the announcement of the actual earnings number from multiple forecasters who report their predictions to the IBES service. Each month IBES provides information on the mean estimate, median estimate, high estimate, low estimate, standard deviation of estimates, number of upward revisions since the previous month, number of downward revisions, as well as various other data such as monthly stock price and adjustment factors related to stock splits. A more detailed discussion of the IBES data source is provided in Foster [1986] or Brown, Foster and Noreen [1985].

2.2 Sample

The first step was to develop a list of bankrupt firms. The primary source used was the F&S Index of Corporate Changes, volumes from January 1977 through September 1985 inclusive. Lists of bankrupt firms provided in Altman [1983] and Zavgren [1983], as well as the Wall Street Journal Index from 1977 through 1984, were also reviewed. All firms listed in these sources declaring either chapter X or chapter XI bankruptcy were eligible for inclusion.

The list of bankrupt firms was cross-referenced with firms on the IBES tape. The IBES tape contained 98 firms that were identified as having declared bankruptcy. Thirteen firms were dropped because only data for months following bankruptcy were included on the IBES tape. Eighteen additional firms were dropped because the number of months of coverage on the tape was considered too short (less than four months). Thus the sample used in the study consists of 68 firms declaring bankruptcy during the period January 1977 to September 1985.

2.3 Matching Firms

Each bankrupt firm was matched with a non-bankrupt firm from the same industry and of approximately the same size. The procedure for matching was as follows: The annual Wards Directory of Leading U.S. Corporations was examined two years before bankruptcy (the middle of the test period) for each bankrupt firm. Wards lists firms by three digit SIC codes. The firm within the same industry (same SIC code) closest to the bankrupt firm in total asset size was identified as a tentative match. The tentative matching firm was accepted if it was included on IBES and had data coverage overlapping the months covered for the bankrupt firm. If not, the firm in the Wards directory next closest in size to the bankrupt firm was identified as a tentative match. The process of cross referencing Wards and IBES was continued until a match for each bankrupt firm was accepted.

Matching on industry is desirable to control for industry characteristics and conditions. Forecast uncertainty may be related to industry. Furthermore, information events may have industry-wide implications leading to industry-wide revisions in earnings forecasts. Using two digit SIC codes, 29 different industries are represented in the sample.

Matching on size is desirable because size is associated with risk, probability of bankruptcy, analyst attention, and most likely, the number of sources from which consensus forecasts and summary statistics on the IBES tape are developed. Using total assets as a measure of size, 58% (42%) of bankrupt firms were larger (smaller) than their non-bankrupt matched firm. Using total sales as a measure, 50% of bankrupt firms were larger than their non-bankrupt match. Both parametric (t-test) and non-parametric (wilcoxon sign rank) tests revealed no significant difference in mean size between bankrupt and non-bankrupt groups, so the matching process was apparently successful.

The 68 matched pairs represent the maximum sample available for the analysis conducted. However, data for each firm was not available on IBES for each month and year of the test period. In addition, in some months where data was available, IBES included forecasts from only one analyst while certain measures used in the analysis (e.g. standard deviation of multiple forecasts) required forecasts from more than one source. Consequently, many individual tests were conducted on sample sizes less than 68.

Matching on fiscal year-end would perhaps be desirable but was not possible without a great reduction in sample size. Data for each firm in a given matched pair were however taken from the same fiscal year. Within a given year there is substantial evidence that the properties of analysts forecasts change as the year-end approaches. For example forecasts tend to become more accurate as the end of a reporting year approaches. However, data in the study is analyzed in "event" time rather than "calendar" time, which minimizes any problem associated with firms having different fiscal year ends.

2.4 A Word on Notation

Notation used in the study also refers to event time. Two events are of importance: the year in which bankruptcy is declared for the bankrupt firm and the month relative to fiscal year-end within any year. The notation used treats bankruptcy as time "zero" and counts backward in time such that both years and months increase as the time before bankruptcy or year-end increases. Year zero is the year in which bankruptcy is declared for a bankrupt firms (and the corresponding fiscal year for the corresponding healthy firm in a matched pair). Year one is the fiscal year immediately prior to the year in which bankruptcy is declared. Year two is two years prior to the year in which bankruptcy is declared, and so on. Within any given fiscal year, month zero is the last month in the year (e.g. December for a firm with December 31 year-end). Month one is one month prior to year-end (e.g. November), and so on. Since forecast data is available for up to two years prior to the end of a given forecast year, months are numbered 0 through 23.

One other date is also referred to in the study: the month of the public announcement of a firms' annual earnings number. For most firms this is one to three months after the end of a fiscal year. For purposes of this study the month in which a firms actual reported earnings number is available on the IBES tape is considered the month of the announcement. The notation "PA", standing for "prior to announcement" refers to the month immediately preceding the month of announcement. This will vary from firm to firm. It represents the time of the last available earnings forecast information prior to the release of the actual reported earnings.

CHAPTER 3

ANALYSIS OF THE PROPERTIES OF ANALYSTS EARNINGS FORECASTS

3.1 Introduction

This chapter identifies observable properties of analysts forecasts and tests for systematic differences in these properties between failing and healthy firms. The five properties addressed are 1. systematic error or bias in forecasts, 2. forecast accuracy, 3. dispersion in forecasts across different analysts, 4. the frequency of revisions in forecasts by analysts, and 5. the level of the mean or average forecast by the analysts. Both forecast bias and forecast accuracy relate to the issue of whether earnings forecasts are "rational". Givoly and Lakonishok [1984] state that rational forecasts should be unbiased, should incorporate all available relevant information, and should be most accurate. Thus the concept of rationality is discussed and the incorporation of available information in analysts forecasts is also addressed.

3.2 The Concept of Rationality

Recent studies (e.g. Critchfield, Dyckman and Lakonishok [1978] and Givoly [1985]) have explored the nature of the formation of analysts earnings forecasts, with emphasis on their rationality. The importance of rationality of earnings expectations and its link to stock market efficiency is expressed by Givoly and Lakonishok [1984]:

> Evidence of rational earnings forecasts would be consistent with both the finding of stock market efficiency and the important role of earnings in stock valuation. Findings of irrational forecasting by analysts would be inconsistent with stock market efficiency unless either FAF do not represent the true market expectations or earnings expectations do not play the role envisioned for them by the various valuation models. (p. 127)

In principle, rationality of a forecasted measure (such as earnings) means that future expectations of the measure of interest are generated by the same stochastic process that generates the actual measure (Muth [1961]). Actual tests of rationality have generally applied a weaker requirement that future expectations fully incorporate the information available in the past historical series of the measure of interest. Both Critchfield, Dyckman and Lakonishok [1978] and Givoly and Lakonishok [1984] offer two broad testable implications of rationality. First, assuming earnings to be a random variable, rational forecasts should not be expected to be error free, but should in general be close to the mean of the true probability distribution. In short, earnings forecasts should not reflect any systematic error or bias. Second, essentially costless information should be impounded in the forecast. Thus it should not be possible to improve on the forecast by using information on past earnings or forecast errors, and analysts forecasts should not be less accurate than forecasts based solely on past earnings. Investigations of these two implications of rationality have failed to reject the hypothesis that financial analysts forecasts of earnings are rational (Critchfield, Dyckman and Lakonishok [1978], Malkiel and Cragg [1980], Givoly [1985]).

Evidence to date has generally been consistent with the view that analysts produce unbiased forecasts that incorporate available past earnings information.

The following sections (3.3-3.5) present some evidence on the properties of analyst earnings forecasts that relate to the issue of rationality. The findings suggest the degree of rationality may vary across firms depending on financial health. The findings indicate that the nature of analysts forecasts differ between healthy and failing firms, that the assumption of complete rationality for failing firms may be open to question and that variables related to the rationality properties of forecasts may be useful to identify failing firms.

3.3 Systematic Error - Bias

If forecasts are rational they should be unbiased. While forecasters cannot be expected to predict without error, rational forecasters should in general be able to predict without systematic error. Consistent systematic error would be inconsistent with rationality since rational forecasters should use the information in past forecast errors to improve future forecasts. Unbiased forecasts imply that

$$\text{Actual} = \text{Forecast} + e$$

where e is a random error with zero expectation.

Studies by Critchfield, Dyckman and Lakonishok [1978], Givoly [1985], and Malkiel and Cragg [1980] have examined analysts forecasts for bias and have failed to reject the hypothesis that forecasts are unbiased. (However, Givoly and Lakonishok [1984], citing Barefield and Comiskey [1975] and Fried and Givoly [1982] conclude that there is an "accumulation of evidence," though statistically insignificant, that an upward bias may be present in analysts forecasts.)

The finding of no systematic error is consistent with rational forecasts and with the proper processing and utilization of information available in the past realizations of earnings and forecast errors. The immediate concern here is whether there is a difference in the systematic error of forecasts between healthy and failing firms.

3.3.1 Comparison of Errors

Various tests may be conducted to assess the degree of bias in forecasts. One approach is to average forecast errors (across firms or across time). If there is a bias, the average prediction error will be significantly different from zero. Several measures of forecast error were used in this study:

The Relative Forecast Error (RFE) was defined as

$$RFE = (Y - \hat{Y}) / |Y|$$

where $\hat{Y}$ is the mean earnings per share forecast and Y the actual realized EPS. (The mean of analysts forecasts for a given firm is used throughout the study as the average or consensus measure. Tests using the median produced the same findings). The RFE is the "standard" measure used in previous studies. Deflation by $|Y|$ is usually deemed necessary to adjust for magnitude differences in

EPS. However, deflation by $|Y|$ may be questioned in the context of the present study because one might expect the EPS of failing firms to differ systematically (smaller, more negative) from non-failing firms. This is not a real problem if the only objective is to identify statistics that may differ between the two groups. It is perhaps a problem if the objective is to test a hypothesis concerning "true" bias. Consequently, other forecast error measures were also used.

The Absolute Forecast Error (AFE) was defined as

$$AFE = (Y - \hat{Y})$$

where Y and $\hat{Y}$ are defined as before. The AFE removes the problem of inducing a bias through the choice of a deflator, but implicitly assumes that sample firms are of similar size. Since healthy and failing firms were matched on size, use of the AFE may be defensible.

A third alternative, the Price-deflated Forecast Error (PFE) was defined as

$$PFE = (Y - \hat{Y}) \ / \ P$$

where P is the market price of a share of stock in the final month of the fiscal year for which Y is the reported EPS. P is provided on IBES. Neiderhoffer and Regan [1972] argue that using prices to deflate forecast errors is superior to other approaches

The Log Relative Forecast Error (LFE) is defined as

$$LFE = \ln (\hat{Y} \ / \ Y)$$

The LFE has been used frequently (e.g. Givoly [1985], Critchfield, Dyckman and Lakonishok [1978]) because of desirable properties: It provides a symmetric distribution of (positive/negative) errors and reduces the impact of outliers. However, the LFE is undefined when Y is negative. Since many failing firms have negative earnings, use of LFE greatly reduced sample size in tests and its use was considered inappropriate.

The Variability-deflated Forecast Error (VFE) is defined as

$$VFE = (Y - \hat{Y}) / \sigma_Y$$

where σ_Y is the standard deviation of the firms historical EPS series. Imhoff [1982] recommends deflating by σ_Y. Calculation of σ_Y for sample firms was generally not possible to do to insufficient data on IBES. As an alternative, σ_F was used as a deflator where σ_F is the cross-sectional standard deviation of forecasts for an individual firm across analysts at a point in time. The implicit assumption when deflating by σ_F is that earnings variability over time (σ_Y) is reflected in dispersion of forecasts (σ_F). Givoly and Lakonishok [1983] do provide empirical evidence that dispersion in forecasts is related to other measures of risk such as beta and the standard deviation of returns, so the assumption may not be inappropriate. Deflating by σ_F is a problem when the number of forecasts available is only one, which was sometimes the case for sample firms. Thus use of VFE on occasion also created sample size problems. Results from tests using measures deflated by σ_F were conducted but are not reported in the study because they provide no additional insights of interest.

As indicated above, one test of bias is to determine if the average forecast error is significantly different from zero.

TABLE B1

FORECAST ERRORS BY GROUP

YEAR 1	Month	Failing Firms AVE	t	α	Healthy Firms AVE	t	α
	PA	-3.06	-2.78	.009	.06	.54	.594
	0	-3.93	-3.40	.002	-.00	-.04	.971
AFE	3	-5.09	-4.11	.000	-.18	-1.59	.118
	6	-6.63	-4.73	.000	-.27	-1.95	.056
	9	-6.66	-4.99	.000	-.36	-2.21	.031
	12	-6.69	-5.04	.000	-.47	-2.49	.016
	PA	-.17	-.41	.682	-.06	-.30	.766
	0	-.59	-4.44	.000	-.27	-2.27	.027
RFE	3	-.92	-6.74	.000	-.94	-2.12	.038
	6	-1.29	-5.88	.000	-1.36	-2.13	.036
	9	-1.45	-4.51	.000	-2.07	-2.20	.032
	12	-1.52	-4.06	.001	-2.07	-2.21	.032
	PA	-.55	-3.09	.005	-.00	-.09	.925
	0	-.75	-2.84	.009	-.01	-.77	.447
PFE	3	-.93	-3.75	.001	-.03	-2.58	.012
	6	-1.33	-3.13	.005	-.04	-3.03	.004
	9	-1.46	-3.19	.004	-.05	-3.24	.002
	12	-1.47	-3.31	.003	-.06	-3.42	.001

TABLE B2

FORECAST ERRORS BY GROUP

		Failing Firms			Healthy Firms		
YEAR 2	MONTH	AVE	t	α	AVE	t	α
	PA	-.85	-1.65	.108	-.15	-1.23	.225
	0	-1.03	-2.12	.042	-.17	-1.34	.185
AFE	3	-1.86	-3.48	.001	-.33	-2.34	.023
	6	-2.05	-5.02	.000	-.43	-3.15	.003
	9	-2.87	-3.62	.001	-.50	-3.38	.001
	12	-3.24	-3.36	.002	-.57	-3.21	.002
	PA	-.46	-3.11	.004	-.11	-1.48	.144
	0	-.60	-3.38	.002	-.12	-1.72	.091
RFE	3	-1.60	-2.39	.023	-.24	-2.84	.006
	6	-2.15	-2.63	.013	-.33	-3.72	.001
	9	-2.18	-3.33	.002	-.37	-4.22	.000
	12	-2.24	-3.27	.003	-.42	-3.76	.000
	PA	-.13	-2.50	.018	-.01	-1.16	.249
	0	-.21	-2.90	.006	-.01	-1.21	.230
PFE	3	-.38	-3.93	.000	-.02	-2.40	.020
	6	-.43	-3.84	.001	-.03	-3.48	.001
	9	-.55	-3.55	001	-.04	-3.25	.002
	12	-.55	-3.14	.004	-.04	-2.92	.005

TABLE B3

FORECAST ERRORS BY GROUP

		Failing Firms			Healthy Firms		
YEAR 3	Month	AVE	t	α	AVE	t	α
	PA	-.50	-2.71	.010	-.05	-.77	.445
	0	-.68	-3.51	.001	-.12	-1.39	.171
AFE	3	-1.36	-2.62	.013	-.16	-1.45	.152
	6	-1.70	-2.61	.015	-.22	-1.37	.177
	9	-2.35	-2.37	.025	-.35	-1.83	.074
	12	-3.02	-2.41	.025	-.46	-1.98	.054
	PA	-.91	-2.03	.049	-.35	-1.96	.055
	0	-1.23	-2.59	.014	-.38	-1.85	.069
RFE	3	-2.37	-1.61	.118	-.53	-2.05	.046
	6	-2.77	-1.65	.110	-.73	-2.12	.039
	9	-3.45	-1.49	.148	-.96	-2.26	.029
	12	-4.03	-1.43	.168	-1.17	-2.24	.030
	PA	-.06	-2.90	.007	.00	0.58	.561
	0	-.08	-3.61	.001	-.00	-.68	.499
PFE	3	-.12	-3.05	.005	-.01	-1.06	.292
	6	-.15	-3.02	.005	-.01	-.73	.470
	9	-.20	-2.75	.011	-.02	-1.54	.131
	12	-.24	-2.64	.015	-.02	-1.61	.115

TABLE B4

FORECAST ERRORS BY GROUP

YEAR 4	Month	Failing Firms AVE	t	α	Healthy Firms AVE	t	α
	PA	.08	.41	.682	-.02	-.31	.761
	0	.14	.63	.536	-.02	-.21	.834
AFE	3	-.31	-1.08	.292	.00	.02	.987
	6	-1.07	-1.81	.084	.07	.51	.616
	9	-1.32	-1.63	.118	.18	1.00	.325
	12	-1.89	-1.75	.099	.17	.81	.422
	PA	-.58	-1.29	.208	-.01	-.30	.764
	0	-.65	-1.29	.209	-.01	-.29	.771
RFE	3	-.87	-1.79	.085	-.08	-1.00	.323
	6	-1.19	-2.17	.041	-.11	-.92	.364
	9	-1.31	-2.34	.029	-.15	-.95	.347
	12	-1.13	-1.96	.067	-.20	-1.08	.286
	PA	.00	.27	.788	.00	.40	.693
	0	.01	.36	.720	.00	.74	.463
PFE	3	-.02	-1.26	.219	.00	.72	.476
	6	-.04	-1.38	.182	.01	1.33	.190
	9	-.06	-1.35	.192	.02	1.48	.148
	12	-.08	-1.36	.196	.01	1.14	.262

Tables B1-B4 provide data on average forecast errors for both healthy and failing groups at three month intervals for the four years prior to bankruptcy. For each group the tables provide measures of the average (mean) forecast error (AFE, RFE, PFE), the t-values for the test of the null hypothesis that the mean forecast error is zero, and the alpha level for rejecting the null hypothesis.

For years 1-3 prior to bankruptcy, several observations are of note. First, for both healthy and failing firms the average forecast error is generally negative. This is consistent with the upward bias previously noted by Givoly and Lakonishok [1984]. Second, for both groups the absolute magnitude of the forecast errors became smaller, closer to zero, as the earnings announcement data approaches. This is consistent with previous findings and is what one would expect as forecasters update their predictions in response to developing information. Third, t-values for the failing groups are consistently larger (more negative) and significant at lower alpha levels than for the healthy group. Fourth, for forecasts made at year end (month 0) or just prior to announcement of actual earnings (month PA) there is little indication of any significant bias for the healthy firms, but consistent indication of significant bias for failing firms. Assuming that management discloses "news" throughout the year that is used by analysts to update forecasts, and assuming that failing firms on average have more "bad" news to disclose than healthy firms, this result is consistent with implications of some studies (e.g. Penman [1980]) that firms may withhold bad news as long as possible but release good news earlier. It is inconsistent with other studies (e.g. Ajinkya and Gift [1984]) that firms have equal incentives to release good and bad news. Finally, note that there is no indication of significant systematic error for either group in year 4, suggesting that no systematic difference exists between the two groups four years prior to bankruptcy.

Of course Tables B1-B4 only test for the existence of a systematic forecast error within each group. A negative bias is present for both groups and, in spite of indications, the tables do not prove that there is a significantly <u>greater</u> bias for the failing firms. Tables B5-B6 provide results of both parametric tests (t-tests) and non-parametric (wilcoxon) tests for differences in average forecast error between the two groups. The findings support the conclusions drawn from Tables B1-B4. Average forecast errors are significantly more negative for failing firms during years 1-3. Findings are stronger using the wilcoxon test, suggesting that individual extreme negative values are not the

TABLE B5

DIFFERENCE IN FORECAST ERRORS BETWEEN GROUPS

		YEAR 1				YEAR 2			
		T-TEST		WILCOXON		T-TEST		WILCOXON	
	Month	t	α	Z	α	t	α	Z	α
AFE	PA	-2.82	.009	-4.42	.000	-1.32	.196	-2.54	.011
	0	-3.38	.002	-4.69	.000	-1.72	.093	-2.97	.003
	3	-3.95	.001	-5.53	.000	-2.77	.009	-4.06	.000
	6	-4.51	.000	-6.22	.000	-3.75	.001	-3.92	.000
	9	-4.69	.000	-6.29	.000	-2.94	.006	-3.18	.002
	12	-4.64	.000	-5.79	.000	-2.73	.011	-3.47	.001
RFE	PA	-.25	.802	-3.27	.001	-2.13	.038	-2.81	.005
	0	-1.59	.115	-3.31	.001	-2.50	.016	-3.30	.001
	3	.03	.967	-3.54	.000	-2.02	.051	-4.47	.000
	6	.11	.915	-3.96	.000	-2.22	.034	-4.05	.000
	9	.30	.764	-3.83	.000	-2.74	.010	-3.62	.000
	12	.54	.590	-3.25	.001	-2.63	.014	-4.13	.000
PFE	PA	-3.08	.005	-4.44	.000	-2.29	.028	-2.69	.007
	0	-2.81	.009	-4.78	.000	-2.75	.009	-3.23	.001
	3	-3.64	.001	-5.78	.000	-3.69	.001	-4.42	.000
	6	-3.05	.006	-6.19	.000	-3.54	.001	-3.99	.000
	9	-3.08	.006	-6.33	.000	-3.31	.003	-3.60	.000
	12	-3.18	.004	-5.76	.000	-2.90	.007	-3.89	.000

TABLE B6

DIFFERENCE IN FORECAST ERRORS BETWEEN GROUPS

		YEAR 3				YEAR 4			
		T-TEST		WILCOXON		T-TEST		WILCOXON	
	Month	t	α	Z	α	t	α	Z	α
	PA	-2.26	.029	-1.95	.051	.49	.624	-.14	.890
	0	-2.63	.012	-3.21	.001	.66	.513	.34	.734
AFE	3	-2.25	.031	-2.84	.005	-1.02	.314	-.56	.579
	6	-2.21	.035	-2.47	.016	-1.87	.072	-.95	.342
	9	-1.99	.056	-2.11	.035	-1.81	.083	-1.22	.224
	12	-2.01	.056	-2.07	.039	-1.87	.078	-1.05	.292
	PA	-1.17	.247	-1.71	.087	-1.26	.219	-.38	.707
	0	-1.65	.104	-3.00	.003	-1.26	.219	.29	.766
RFE	3	-1.23	.226	-2.60	.009	-1.62	.118	-.88	.380
	6	-1.18	.243	-2.28	.023	-1.93	.066	-1.12	.264
	9	-1.06	.298	-2.11	.034	-1.98	.058	-1.39	.162
	12	-.99	.330	-1.76	.078	-1.54	.139	-.88	.379
	PA	-2.96	.005	-3.04	.002	.13	.897	.06	.956
	0	-3.33	.002	-3.68	.000	.10	.919	.41	.678
PFE	3	-2.82	.008	-3.20	.001	-1.43	.160	-.81	.416
	6	-2.82	.008	-2.94	.003	-1.69	.104	.68	.498
	9	-2.48	.019	-2.69	.007	-1.66	.109	-.97	.332
	12	-2.37	.026	-2.46	.014	-1.54	.143	-.69	.491

cause of the result. No difference between groups is evident in year 4.

3.3.2 Regression Test - Full Sample

Theil [1966], Mincer and Zarnowitz [1969] and Givoly [1985] distinguish between two kinds of bias. Forecasts may contain a "level" bias in that on average forecasts are too low or too high. The previous tests suggest some level bias in forecasts for failing firms. It is possible, however, that forecasts may be unbiased on average but still exhibit a bias that is conditional on the nature of earnings - e.g., forecasts may overstate earnings when earnings are high and understate earnings when earnings are low. This is referred to as a "regression" bias and could occur when no level bias is present. This suggests an alternative approach for assessing bias.

The most commonly used approach is to estimate a regression of the following form:

$$\text{Actual} = \beta_0 + \beta_1(\text{Forecast}) + e$$

where the dependent variable is actual earnings or earnings growth and the independent variable is predicted earnings or earnings growth. The regression can be estimated using either time series data for individual firms or cross-sectional data at a given point in time. The null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ is tested, and rejection of the null hypothesis is treated as evidence of bias. If forecasts are on balance too high (too low) then the intercept term β_0 would be less than (greater than) zero. If forecasts are too high (low) when earnings are high and too low (high) when earnings are low then the coefficient term β_1 would be less (greater) than one.

Two cross-sectional versions of the above regression were estimated using undeflated EPS measures and EPS measures deflated by stock price. Regressions were estimated at selected months for four years prior to failure using the full sample of firms - i.e., including both failing and healthy firms. Clearly this is not a representative sample as it is much more highly weighted toward failing firms than the population of firms in general.

Parameter estimates, F statistics and significance levels for the two hypothesis of interest, $\beta_0 = 0$ and $\beta_1 = 1$, are presented in Table B7. Findings are similar for both deflated and undeflated tests. In years 1-3 the two null hypotheses can be rejected at .05 or better for a majority of the tests. β_0 is universally negative, which is consistent with the level bias noted in previous tests. β_1 is generally greater than one. This is consistent with a regression bias where optimistic forecasts are associated with low earnings realizations and/or pessimistic forecasts are associated with high earnings realizations. (These results are not present in year four, again indicating that the distinguishing properties of the failing firms are not exhibited four years prior to bankruptcy.)

TABLE B7

REGRESSION TEST FOR BIAS - FULL SAMPLE

		UNDEFLATED MEASURES						PRICE DEFLATED MEASURES					
				Test: $\beta_0 = 0$		Test: $\beta_1 = 1$				Test: $\beta_0 = 0$		Test: $\beta_1 = 1$	
	Month	β_0	β_1	F	α	F	α	β_0	β_1	F	α	F	α
YEAR 1	PA	-1.07	1.11	5.32	.024	.48	.493	-.09	1.47	3.01	.087	68.74	.000
	0	-1.51	1.25	9.56	.003	1.82	.181	-.17	2.13	6.87	.011	96.91	.000
	3	-2.53	1.55	16.27	.000	4.02	.049	-.30	2.24	41.32	.000	235.31	.000
	6	-2.97	1.55	14.65	.000	2.66	.107	-.81	5.67	96.25	.000	164.41	.000
	9	-3.25	1.61	14.86	.000	2.78	.100	-.79	4.06	15.96	.000	7.75	.007
YEAR 2	PA	-.22	.95	2.85	.095	2.23	.139	-.05	1.10	7.07	.010	5.98	.017
	0	-.27	.95	3.78	.056	2.13	.149	-.06	1.11	8.41	.005	5.69	.019
	3	-.85	1.09	20.09	.000	2.92	.091	-.15	1.41	20.61	.000	16.31	.000
	6	-1.43	1.28	48.79	.000	19.58	.000	-.20	1.36	20.79	.000	6.78	.011
	9	-2.79	1.79	35.65	.000	17.15	.000	-.19	.94	5.09	.027	.02	.881
YEAR 3	PA	-.39	1.13	14.14	.000	18.32	.000	-.03	1.16	13.54	.001	25.82	.000
	0	-.50	1.12	19.36	.000	13.17	.001	-.04	1.16	22.94	.000	25.28	.000
	3	-1.64	1.59	36.84	.000	33.64	.000	-.13	1.76	48.68	.000	49.94	.000
	6	-2.09	1.66	27.88	.000	18.80	.000	-.18	1.95	37.09	.000	26.82	.000
	9	-.52	.69	.42	.521	1.01	.320	-.02	.52	.10	.749	1.30	.258
YEAR 4	PA	-.01	1.02	.00	.960	.06	.803	.01	.90	1.16	.287	1.15	.288
	0	.02	1.01	.02	.889	.04	.835	.01	.93	.99	.323	.64	.429
	3	-.30	1.09	1.63	.207	.91	.344	-.01	1.07	.58	.449	.33	.567
	6	.65	.55	3.45	.069	12.60	.001	.06	.44	6.25	.015	11.49	.001
	9	1.33	.23	12.95	.001	39.19	.000	.09	.12	19.62	.000	40.03	.000

3.3.3 Regression Tests - By Group

Both Givoly [1985] and Critchfield, Dyckman and Lakonishok [1978] detected no substantial level or regression bias in their tests. Although their sample sizes were perhaps more representative of the population of all firms, they were not fully representative in that firms with poor (i.e., negative) earnings were deleted. The earnings for the failing firms are lower and more frequently negative than earnings for the healthy firms in the current study. (Both t-tests and wilcoxon tests reveal significantly lower earnings, at α = .05, for the failing group as compared to the healthy group for years 1-3. No significant difference exists in year 4.) It is probable that the exhibited bias is driven by the heavy representation of failing firms in the sample. To explore the issue further, a modified version of the above regression was estimated as follows:

$$Y = \beta_{FO}(F) + \beta_{HO}(H) + \beta_{F1}(F \times \hat{Y}) + \beta_{H1}(H \times \hat{Y}) + e$$

where

F = 1 if firm is in failing group
= 0 if not

H = 1 if firm is in healthy group
= 0 if not

Y = actual earnings

$\hat{Y}$ = forecasted earnings

β_{FO} and β_{F1} are equivalent to β_O and β_1 for a regression that includes failing firms only; likewise for β_{HO} and β_{H1} and healthy firms. Estimating the regression jointly including failing and healthy firms allows for testing six hypotheses.

a) $\beta_{F0} = 0$

b) $\beta_{H0} = 0$

c) $\beta_{F0} = \beta_{H0}$

d) $\beta_{F1} = 1$

e) $\beta_{H1} = 1$

f) $\beta_{F1} = \beta_{H1}$

Findings for selected months for the four years prior to bankruptcy are included in Table B8. The Table include intercepts, coefficients, F statistics for each hypothesis, and significance levels.

As in previous tests, results for years 1-3 are most interesting. Columns 3 and 5 of the Table show that significant over-estimation of earnings occurs for failing firms while column 4 indicates no similar level bias for healthy firms. Statistics indicating a regression bias are also occasionally significant (at .05) for failing firms (column 6) but never significant for healthy firms (column 7).

It seems apparent that both the level and regression bias detected in the full sample regressions results in Table B7 are driven by the presence of the failing firms. Earnings are over-estimated for the failing firms, and since these firms tend to have poorer earnings, the regression bias of optimistic forecasts associated with lower earnings results.

TABLE B8

REGRESSION TEST FOR BIAS - BY GROUP (Undeflated)

		1 Intercepts		2 Coefficients		3 $\beta_{F0} = 0$		4 $\beta_{H0}=0$		5 $\beta_{F0} = \beta_{H0}$		6 $\beta_{F1} = 1$		7 $\beta_{H1} = 1$		8 $\beta_{F1} = \beta_{H1}$	
	Month	β_{F0}	β_{H0}	β_{F1}	β_{H1}	F	α	F	α	F	α	F	α	F	α	F	α
YEAR 1	PA	-4.91	-.03	.46	1.04	32.47	.000	.00	.959	6.95	.010	6.96	.010	.02	.880	2.77	.101
	0	-5.21	-.06	.59	1.03	43.26	.000	.01	.924	24.87	.000	3.01	.087	.02	.900	1.45	.233
	3	-5.95	-.43	.60	1.13	58.50	.000	.31	.580	25.59	.000	.99	.322	.17	.679	1.06	.306
	6	-6.61	-.49	.79	1.11	52.54	.000	.36	.549	24.70	.000	.20	.654	.11	.745	.31	.580
	9	-7.05	-.43	1.04	1.03	49.86	.000	.24	.623	24.60	.000	.00	.945	.01	.932	.00	.993
YEAR 2	PA	-.49	.03	.93	.89	4.93	.029	.01	.924	2.27	.136	3.08	.083	.90	.345	.15	.703
	0	-.64	.03	.93	.88	7.38	.008	.01	.905	3.41	.069	3.44	.067	.97	.327	.15	.705
	3	-1.42	.07	1.08	.79	24.40	.000	.03	.853	10.63	.002	2.44	.122	2.26	.136	3.93	.051
	6	-2.19	-.11	1.31	.84	61.49	.000	.09	.767	20.84	.000	21.67	.000	1.31	.254	9.03	.004
	9	-4.78	-.15	2.47	.84	80.06	.000	.06	.797	33.19	.000	38.48	.000	.47	.494	23.91	.000
YEAR 3	PA	-.53	.18	1.15	.90	13.69	.000	.79	.377	8.28	.005	21.93	.000	1.59	.211	8.43	.005
	0	-.67	.20	1.15	.84	18.58	.000	.80	.374	11.12	.001	18.83	.000	3.49	.066	11.31	.001
	3	-2.21	.25	1.92	.81	59.80	.000	.41	.524	25.74	.000	75.86	.000	1.62	.207	36.39	.000
	6	-3.51	.55	2.34	.65	83.77	.000	1.34	.251	43.84	.000	70.41	.000	3.77	.056	49.50	.000
	9	-1.92	.79	.72	.51	2.35	.130	.63	.431	2.87	.095	.26	.612	1.91	.172	.10	.750
YEAR 4	PA	-.04	.08	1.11	.95	.04	.843	.13	.723	.16	.695	1.20	.279	.30	.583	1.38	.246
	0	-.02	.13	1.13	.93	.01	.927	.32	.576	.22	.642	1.81	.184	.57	.453	2.26	.139
	3	-.77	.06	1.22	.98	4.83	.032	.04	.848	3.19	.080	2.29	.136	.02	.886	1.57	.216
	6	.53	.28	.37	.91	1.07	.305	.37	.545	.13	.721	17.14	.000	.24	.627	4.87	.032
	9	1.23	.29	.07	.94	5.58	.022	.29	.594	1.52	.222	52.16	.000	.06	.814	10.04	.003

TABLE B9

REGRESSION TEST FOR BIAS - BY GROUP (Price Deflated)

		1		2		3		4		5		6		7		8	
		Intercepts		Coefficients		$\beta_{F0} = 0$		$\beta_{H0} = 0$		$\beta_{F0} = \beta_{H0}$		$\beta_{F1} = 1$		$\beta_{H1} = 1$		$\beta_{F1} = \beta_{H1}$	
	Month	β_{F0}	β_{H0}	β_{F1}	β_{H1}	F	α	F	α	F	α	F	α	F	α	F	α
YEAR 1	PA	-2.86	-.21	.59	1.14	41.23	.000	.20	.652	17.11	.000	7.77	.007	.11	.745	1.37	.245
	0	-3.38	-.31	.68	1.17	146.63	.000	.43	.516	36.67	.000	5.05	.028	.14	.712	1.07	.304
	3	-3.18	-.50	.74	1.19	55.90	.000	.81	.371	14.69	.000	1.34	.250	.15	.698	.69	.410
	6	-3.82	-.46	1.49	1.04	374.47	.000	.74	.392	35.10	.000	2.53	.116	.01	.929	.71	.403
	9	-4.52	-.46	1.19	.89	970.57	.000	.54	.465	39.11	.000	.26	.610	.05	.819	.25	.621
YEAR 2	PA	-.84	-.12	1.02	1.04	26.61	.000	.20	.653	5.13	.026	.33	.568	.03	.867	.01	.944
	0	-1.05	-.13	1.02	1.04	37.91	.000	.19	.662	7.32	.008	.23	.631	.02	.890	.00	.954
	3	-2.16	-.15	1.30	.94	101.65	.000	.13	.719	19.07	.000	15.25	.000	.03	.867	1.14	.289
	6	-2.88	-.16	1.35	.90	136.33	.000	.17	.733	25.70	.000	12.67	.001	.08	.776	1.43	.235
	9	-3.93	-.03	1.65	.77	58.86	.000	.00	.967	17.33	.000	3.39	.069	.15	.698	1.64	.204
YEAR 3	PA	-.39	-.01	1.15	1.06	17.14	.000	.01	.917	6.86	.011	19.09	.000	.35	.558	.79	.379
	0	-.45	-.02	1.15	1.00	21.32	.000	.02	.877	8.13	.006	18.18	.000	.00	.977	1.96	.167
	3	-.85	-.08	1.71	.98	15.15	.000	.09	.767	4.78	.032	29.11	.000	.01	.924	7.92	.006
	6	-1.94	.18	1.97	.79	42.94	.000	.41	.523	26.62	.000	25.36	.000	.79	.377	14.97	.000
	9	-.62	.23	.21	.67	1.09	.299	.26	.615	1.29	.260	3.87	.054	.87	.355	.76	.387
YEAR 4	PA	.00	.31	.92	.86	.00	.987	3.06	.086	1.58	.214	.52	.473	1.53	.221	.16	.693
	0	-.07	.33	.99	.86	.14	.712	3.31	.075	2.43	.125	.01	.923	1.56	.217	.65	.425
	3	-.32	.29	1.05	.89	2.33	.133	2.02	.161	4.35	.042	.15	.691	.70	.406	.76	.387
	6	.32	.38	.41	.89	1.47	.230	1.87	.177	.03	.866	17.62	.000	.29	.593	3.69	.060
	9	.51	.46	.20	.83	3.36	.072	2.07	.156	.02	.897	41.60	.000	.51	.480	5.49	.023

3.4 Incorporation of Available Information

As previously noted, systematic error captures only one aspect of rationality. Rational forecasts should also reflect all relevant historical information available at the time of forecast, such as earnings history, past forecasts, past forecast errors, and stock price information, as well as other exogenous information. Given unlimited historical information, it is not feasible to test for the incorporation of all relevant historical information in forecasts. Past research (Givoly [1985], Fried and Givoly [1982], Malkiel and Cragg [1980]) has instead tested whether forecasts incorporate two more obvious kinds of available information: past earnings series and past forecast errors. The findings to date are consistent with analysts forecasts reflecting the information in both past earning and past forecast errors.

3.4.1 Past Earnings

It is also not feasible to test if forecasts incorporate all information that may exist in the time series of past earnings. There are too many possible systematic patterns that may exist. The test conducted here instead investigates whether forecasts incorporate the information contained in three naive mechanical prediction models based on past earnings. The models used were as follows: $F_{1T} = Y_{T-1}$

$$F_{2T} = Y_{T-1} + (Y_{T-1} - Y_{T-2})$$

$$F_{3T} = Y_{T-1} + 1/2 \ \ ((Y_{T-1} - Y_{T-2}) + (Y_{T-2} - Y_{T-3}))$$

Where F_{iT} is the mechanical forecast from model 1, 2, or 3. Model

1 is a martingale, models 2 and 3 are submartingales which differ only in the period over which earnings growth is calculated. Past time series studies (See Bao, Lewis, Lin and Manegold [1983] for a review) have found submartingale models to fairly represent the time series behavior of annual earnings and to perform as well as other mechanical models in forecasting next period earnings. Both Givoly [1985] and Fried and Givoly [1982] used submartingale models as a benchmark for assessing forecaster rationality.

Following Givoly [1985], the partial correlation between actual earnings and the mechanical model prediction, given the analysts forecast, was calculated. This partial correlation measures any association between actual earnings and the mechanical forecast that is not incorporated in the analysts forecast. A non-zero partial correlation would indicate the availability of unused extrapolative information in the mechanical forecast. Partial correlation results are presented in Table B-10. Correlations between actual earnings for the three years prior to bankruptcy with the mechanical model forecasts for that year (when data availability allowed), while controlling for analysts forecasts available at various times (PA, 0, 3, 6, 9) prior to announcement, as well as cumulative results for data aggregated over the three years, are presented. Both undeflated and price-deflated measures of actual, mechanical model, and analyst forecasts were used. For the healthy group there is no indication of correlation across the tests. R^2s are consistently small and insignificant, and it is safe to conclude that analysts forecasts

TABLE B-10

PARTIAL CORRELATION OF ACTUAL EARNINGS WITH MECHANICAL FORECASTS

		FAILING FIRMS						HEALTHY FIRMS					
		UNDEFLATED			PRICE-DEFLATED			UNDEFLATED			PRICE-DEFLATED		
	Month	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
Year 1	PA	.11	-.10	.13	.59**	.62**	.65**	-.12	-.06	-.20	.00	.02	-.09
	0	.07	-.16	.10	.77***	.80***	.83***	-.10	-.04	-.20	.13	.11	.02
	3	.16	-.16	.16	.41*	.42	.47*	-.28*	-.14	-.31*	-.09	-.05	-.18
	6	.18	-.08	.22	.75***	.60**	.73***	-.15	-.10	-.20	-.05	-.02	-.13
	9	.13	-.15	.15	.94***	.95***	.95***	.01	.01	-.11	.17	.15	.05
Year 2	PA	-.24	-.56**	NA	-.45**	-.39*	NA	.04	-.09	NA	-.07	-.03	NA
	0	-.23	-.46**	NA	-.37*	-.25	NA	.03	-.07	NA	-.03	.01	NA
	3	-.29	-.61***	NA	-.28	-.16	NA	.19	.05	NA	.07	.15	NA
	6	-.24	-.28	NA	.01	.00	NA	.01	-.10	NA	-.05	.04	NA
	9	.50**	.32	NA	.70***	.55**	NA	.02	.00	NA	.12	.22	NA
Year 3	PA	-.00	NA	NA	-.10	NA	NA	-.08	NA	NA	-.09	NA	NA
	0	-.02	NA	NA	-.16	NA	NA	-.07	NA	NA	-.06	NA	NA
	3	-.67***	NA	NA	-.56***	NA	NA	-.12	NA	NA	-.03	NA	NA
	6	-.52**	NA	NA	-.36	NA	NA	-.07	NA	NA	.07	NA	NA
	9	.06	NA	NA	-.01	NA	NA	.13	NA	NA	.23	NA	NA
Years 1-3	PA	.02	-.17	.13	.44***	.53***	.65**	-.05	-.07	-.20	-.03	.00	-.09
	0	-.01	-.20	.10	.62***	.69***	.83***	-.04	-.06	-.20	.05	.07	.02
	3	-.05	-.26	.16	.20	.27	.47*	-.05	-.03	-.31*	-.03	.02	-.18
	6	-.00	-.12	.22	.57***	.42**	.73***	-.08	-.10	-.20	-.03	.00	-.13
	9	.19	.10	.15	.86***	.89***	.95***	.04	.01	-.11	.16*	.17*	.05*

Significance levels: .05*/.01**/.001***

incorporate the information contained in the three mechanical models.

In spite of some indication of significant negative correlations among the results for the undeflated tests for the failing firm group, the evidence is not consistent, and does disappear when aggregated over the three years. It is difficult to make a convincing case for non-incorporation of past earnings information from these tests. The results using price-deflated measures are strongly significant in year 1, with a similar pattern evident in the cumulative tests. The very high positive correlations (up to .95) are perhaps too strong to be true. Undeflated measures that exhibit immaterial correlation surprisingly become highly correlated when deflated by a common deflator. It would appear that the results are spurious and caused by the deflating procedure.

In short, there is no convincing evidence of any consistent lack of incorporation of past earnings series information in forecasts in either the failing or healthy groups; and there is no indication of any pattern in the data that may be exploited to distinguish between failing and healthy firms.

3.4.2 Past Forecast Errors

The second item of information that is readily available to a forecaster and should be considered in prediction is the previous forecast error. If a forecaster uses past forecast error information in setting future forecasts there should be no consistent association between forecast errors over time--i.e., no serial correlation. Consistently over-estimating or under-estimating earnings (positive serial correlation) or consistent mis-adjustment to past forecast errors (e.g. reacting to under-estimation one period with over-estimation the next, followed by under-estimation again--i.e., negative serial correlation) would both indicate the presences of unexploited information in the series of past forecast errors.

Ideally, serial correlation should be tested on a firm by firm basis, since a consistent pattern of mis-adjustment could be present for firms, but the type of pattern could differ from firm to firm. Givoly [1985] used firm specific time series regressions of the $(\text{Actual} - \text{Forecast})_{\tau} = \alpha + \beta(\text{Actual} - \text{Forecast})_{\tau-1} + e$ to test for serial correlation in forecast errors. He was unable to reject the null hypothesis that $\alpha = 0$, $\beta = 0$, and concluded a lack of serial correlation between errors.

There is insufficient times series data to test for firm specific serial correlation in this study. Instead, cross-sectional tests are conducted. Correlations between year τ-1 forecast errors at month PA and year τ forecast errors at selected months are contained in Tables B11 and B12 for undeflated, relative, and price-deflated forecast error measures. Individual correlation statistics are significant, but no consistent pattern is apparent in the tables. When aggregated over the three years available, the cumulative correlations for both the undeflated (AFE) and relative (RFE) measures are small and generally insignificant. Again, there

TABLE B-11

SERIAL CORRELATION BETWEEN FORECAST ERRORS (Pearsons)

		FAILING			HEALTHY		
YEARS	Month	AFE	RFE	PFE	AFE	RFE	PFE
1/2	PA	-.29	-.11	.62**	.08	.04	.12
	0	-.27	-.14	.75***	.11	.07	.17
	3	-.43	-.15	.72***	.09	.01	.15
	6	-.42	-.01	.81***	.11	.01	.16
	9	-.39	-.00	.81***	.11	.01	.14
2/3	PA	-.27	.06	-.15	-.14	-.15	.06
	0	-.26	.03	-.15	-.08	-.15	.06
	3	.18	.10	.06	.01	-.15	.15
	6	.43*	.14	.15	-.01	-.06	.15
	9	.67***	.15	.30	.26	.10	.28*
3/4	PA	.01	.15	.15	.19	.20	.24
	0	.01	.19	.11	.52***	.67***	.32*
	3	-.38*	.07	-.16	.41**	.56***	.33*
	6	-.35	.07	-.15	.33*	.52***	.33*
	9	-.37	.06	-.17	.35*	.51***	.38*
ALL	PA	-.18	.11	.51***	.01	-.02	.11
	0	-.17	.16	.60***	.11	.04	.16
	3	-.23*	.08	.57***	.12	.00	.18*
	6	-.17	.09	.66***	.13	.04	.19*
	9	-.03	.07	.68***	.21**	.05	.29**

Significance Level: .05*/ .01**/ .001***

TABLE B-12

SERIAL CORRELATION BETWEEN FORECAST ERRORS (Spearman)

		FAILING			HEALTHY		
Years	Month	AFE	RFE	PFE	AFE	RFE	PFE
1/2	PA	-.37	-.21	-.03	.16	.24	.18
	0	-.30	-.10	-.02	.27*	.35**	.33*
	3	-.46*	-.16	-.10	.37**	.42**	.41**
	6	-.31	.22	-.03	.40**	.47***	.43**
	9	-.17	.26	.11	.42**	.46***	.44***
2/3	PA	-.10	.13	-.12	.11	.10	.17
	0	-.09	.21	-.10	.08	.06	.16
	3	.02	.37*	.07	.16	.15	.20
	6	.21	.34	.20	.17	.23	.20
	9	.33	.40*	.26	.30*	.36**	.28*
3/4	PA	.33	.53**	.34	-.18	-.11	-.01
	0	.22	.37	.22	-.02	.03	.06
	3	.16	.18	.08	-.06	-.06	.04
	6	.09	.25	.08	-.07	-.04	.01
	9	.13	.33	.12	.02	.03	.12
ALL	PA	-.03	.21	.03	.04	.08	.12
	0	-.03	.23*	.03	.12	.16*	.19*
	3	.03	.25*	.12	.17	.21**	.25**
	6	.09	.30**	.17	.19*	.25**	.26**
	9	.17	.38***	.23*	.28***	.31***	.31***

Significance Level: .05*/ .01**/ .001*** are some large correlations using price-deflated measures, but they may be caused by the deflation process. Overall there is little basis to argue for a consistent significant serial correlation among errors.

To explore the issue further, the signs of forecast errors were investigated. Chi-square tests of association between the sign of forecast errors in successive years were conducted using month PA forecasts from the earlier year and forecasts at various months (PA, 0, 3, 6, 9) from the later year. Frequency percentages (the percent of sample firms falling within each of the four possible sign combination categories), X^2 values and significance levels for rejection of the hypothesis of no association between successive year forecast error signs are in Table B-13.

Several items are of note. First, for the failing firm group, cell sizes are frequently too small for valid X^2 tests. There is always a clear majority of firms that have negative forecast errors in the later year, regardless of which pair of successive years prior to bankruptcy are viewed. In other words, there are more negative forecast errors (over-estimation of income) as the time prior to bankruptcy approaches. There is a strong tendency for negative errors to be followed by negative errors (because of the trend of increasing frequency of negative errors), but there is no corresponding tendency for positive errors to be followed by positive errors (again because of the trend of increasing negative errors), and consequently there is no rejection of the null hypothesis of no association between successive errors. In

TABLE B-13

ASSOCIATION BETWEEN SIGNS OF SUCCESSIVE FORECAST ERRORS

		FAILING						HEALTHY					
		Frequency Percentage (Sign at t/sign at t-1)						Frequency Percentage (Sign at t/sign at t-1)					
Years	Month	-/-	-/+	+/-	+/+	x^2		-/-	-/+	+/-	+/+	x^2	
1/2	PA	50	35	10	5	.07		41	24	15	20	2.15	
	0	50	35	10	5	.07		44	22	11	24	6.18	*
	3	58	42	0	0			44	20	11	26	7.64	**
	6	56	44	0	0			47	20	7	25	11.28	***
	9	60	40	0	0			49	19	5	27	15.48	***
2/3	PA	48	24	21	7	.19		37	19	17	28	4.56	*
	0	48	26	19	7	.26		37	17	17	30	5.87	*
	3	63	25	6	6	.75		40	20	15	25	4.89	*
	6	66	22	3	10	4.07	*	44	20	9	26	8.84	**
	9	61	19	10	10	1.58		43	22	11	24	5.77	*
3/4	PA	43	25	4	29	6.65	**	23	39	23	16	2.00	
	0	36	36	11	18	.36		28	37	19	17	.43	
	3	32	39	14	14	.06		26	40	21	14	1.69	
	6	37	37	11	15	.11		23	40	21	17	1.50	
	9	39	36	7	18	1.20		29	36	16	20	.01	
ALL	PA	47	27	12	14	2.01		34	26	18	22	1.93	
	0	44	32	14	10	.00		37	24	15	24	6.52	*
	3	51	34	8	8	.39		37	25	15	22	5.19	*
	6	53	32	5	9	2.58		39	26	12	23	9.81	**
	9	53	30	6	10	2.86		41	25	10	24	14.42	***

Singnificance Level: .05*/ .01**/ .001*** short, the trend toward negative errors is "interesting" but does not provide evidence that signs of errors are serially associated.

For the healthy group the null hypothesis of no association can be rejected in most of the tests conducted. Negative errors tend to follow negative errors; positive tend to follow positive. This is surprising and conflicts with the findings in Tables B-11 and B-12 using correlations. It suggests that the signs of successive forecast errors may be related (and be reflected in the Chi-square tests) while the magnitude of the forecast errors, given the sign are not related (thus reducing the earlier correlations). This suggests there may be some serial association between successive forecast errors that were not revealed in the correlation tests conducted by Givoly [1985]. Future tests investigating just the signs of forecast errors and using larger samples than this paper may be warranted.

With respect to the various tests of incorporation of past information in forecasts, there is little strong, consistent, and unambiguous evidence to confidently reject the hypothesis that analysts forecasts reflect information available in past earnings and past forecast errors. The most convincing new finding was that the signs of successive forecast errors may be related, but this finding needs additional research. Further, while some differences between failing and healthy firms were evident in the data, there is no obvious difference that appears to be potentially useful in predicting impending bankruptcy.

3.5 Accuracy of Analysts Forecasts

Past investigations of the accuracy of analysts forecasts (see Givoly and Lakonishok [1984] for a review) have generally focused on two questions: Are analysts forecasts more accurate than forecasts from mechanical models? And how do analyst's forecasts compare to management forecasts? Results of studies comparing analysts forecasts with forecasts from numerous types of mechanical models (e.g. random walk model, submartingale models, Box-Jenkins models exploiting serial correlation in the earnings time series, and index models tying earnings predictions to a market-wide index of earnings) have occasionally been contradictory, but in general analyst forecasts appear to out-perform mechanical models. Results of studies comparing analyst forecasts to management forecasts suggest a slight but insignificant advantage to management. These results are not surprising. One would expect analysts to out-perform mechanical models given the wider information set on which analysts may rely. Likewise, one would not be surprised by the essentially similar performance between analysts and management given their similar information sets and the incentives for management to provide information to analysts (Lees [1983], Ajinkya and Gift [1984]).

The objective in this section is to test for systematic differences in analyst accuracy between failing and healthy firms. The two most popular error measures in past studies are the relative error

$$|\text{Forecast} - \text{Actual}| \;/\; \text{Actual}$$

and the relative squared error

$$(\text{Forecast} - \text{Actual})^2 / \text{Actual}$$

The squared error gives greater weight to large errors which is consistent with a quadratic loss function. The two error measures, however, tend to be highly correlated. Because of the problems associated with deflating by actual earnings when earnings are small or negative, as noted in the earlier discussion of bias measures, four different error measures are used in the study:

a) Absolute Error Measure (AEM)

$= |\text{Forecast} - \text{Actual}|$

b) Relative Error Measure (REM)

$= |\text{Forecast} - \text{Actual}| \, / \, |\text{Actual}|$

c) Squared Error Measure (SEM)

$= (\text{Forecast} - \text{Actual})^2 \, / \, |\text{Actual}|$

d) Price-deflated Error Measure (PEM)

$= |\text{Forecast} - \text{Actual}| \, / \, \text{Stock price}$

Each error measure can only take on positive values.

3.5.1 Failing vs. Healthy Firms

The first test of interest is whether there is any difference in analyst forecast accuracy for failing firms and healthy firms. Larger values for the error measures imply less accurate forecasts. Table A-1 provides mean values for each of the four error measures for selected months prior to year-end for year 1. Statistics and probability levels for both t-tests and wilcoxon tests of the null hypothesis of no difference between group means are also presented.

Tables A-2 through A-4 provide similar information for years 2-4.

Findings are robust across the four error measures. For both failing and healthy firms errors steadily decrease as reporting date is approached. This is consistent with the finding of others (e.g. Crichfield, Dyckman, Lakonishok [1978]). For all four years, however, the hypothesis of no difference in accuracy can be rejected. Errors are consistently larger for failing firms. Using wilcoxon tests the difference is significant at .05 or lower for all except 8 of the possible 80 year-month-error measure combinations. Seven of the eight non-significant differences, however, occur in year 4. Observing the magnitude of Z values and significance levels, there is a tendency for the difference in accuracy to be more pronounced as bankruptcy is approached. Evidence from the t-tests, although not quite as consistently strong, support the same conclusions: less accurate forecasts for failing firms, and larger errors and declining comparative accuracy as bankruptcy approaches. The findings suggest that accuracy of forecasts are one possible information item that may be exploited for predicting impending failure.

3.5.2 Analysts Forecasts vs. Naive Models

The typical approach used in past studies to assess accuracy has been to compare analyst accuracy with the accuracy achieved by mechanical models based on past earnings. Do analysts forecast better than extrapolations of past earnings? Is the performance of analysts relative to mechanical models the same for both groups of

TABLE A 1

DIFFERENCES IN ERROR MEASURES BETWEEN GROUPS

YEAR 1

		MEANS		t-TEST		WILCOXON	
Error	Month	Failing	Healthy	t	α	Z	α
	PA	3.65	.39	3.16	.004	5.89	.000
	0	4.17	.40	3.34	.003	5.97	.000
AEM	3	5.16	.55	3.75	.001	5.46	.000
	6	6.63	.72	4.20	.000	5.95	.000
	9	6.66	.89	4.31	.000	6.01	.000
	PA	1.00	.54	1.13	.266	4.45	.000
	0	.77	.40	2.54	.013	4.55	.000
REM	3	.96	1.06	-.21	.836	3.71	.000
	6	1.29	1.50	-.31	.761	3.96	.000
	9	1.45	1.85	-.48	.635	3.83	.000
	PA	.68	.03	4.10	.000	6.61	.000
	0	.79	.04	2.92	.007	6.52	.000
PEM	3	.94	.05	3.60	.001	6.07	.000
	6	1.34	.06	2.99	.007	6.09	.000
	9	1.47	.07	3.02	.007	6.24	.000
	PA	5.09	.47	2.16	.039	5.33	.000
	0	3.92	.35	2.99	.006	5.50	.000
SEM	3	5.27	1.58	2.43	.019	4.93	.000
	6	7.43	2.87	2.10	.040	5.40	.000
	9	7.99	4.35	1.33	.188	5.38	.000

TABLE A 2

DIFFERENCES IN ERROR MEASURES BETWEEN GROUPS

YEAR 2

		MEANS		t-TEST		WILCOXON	
Error	Month	Failing	Healthy	t	α	Z	α
	PA	1.28	.38	1.82	.077	3.73	.000
	0	1.40	.43	2.06	.046	3.27	.001
AEM	3	2.07	.58	2.83	.008	3.86	.000
	6	2.10	.63	3.52	.001	3.51	.001
	9	2.97	.75	2.82	.008	3.00	.003
	PA	.61	.23	2.51	.015	4.29	.000
	0	.72	.26	2.63	.012	4.06	.000
REM	3	1.64	.36	1.91	.065	4.11	.000
	6	2.18	.42	2.15	.040	3.77	.000
	9	2.23	.49	2.66	.012	3.63	.000
	PA	.17	.03	2.79	.009	4.24	.000
	0	.23	.03	2.90	.006	3.89	.000
PEM	3	.39	.04	3.70	.001	4.21	.000
	6	.43	.04	3.50	.002	3.84	.000
	9	.55	.05	3.25	.003	3.50	.001
	PA	1.00	.48	.91	.363	4.30	.000
	0	1.29	.45	1.48	.144	3.95	.000
SEM	3	3.05	.60	2.24	.031	4.34	.000
	6	4.48	.63	2.26	.031	3.92	.000
	9	6.21	.65	2.68	.012	3.59	.000

TABLE A 3

DIFFERENCES IN ERROR MEASURES BETWEEN GROUPS

YEAR 3

		MEANS		t-TEST		WILCOXON	
Error	Month	Failing	Healthy	t	α	Z	α
	PA	.71	.32	2.31	.026	2.61	.009
	0	.86	.34	2.78	.008	3.76	.000
AEM	3	1.52	.44	2.24	.032	3.39	.001
	6	1.89	.64	1.94	.062	2.54	.011
	9	2.59	.77	1.85	.074	2.18	.029
	PA	1.15	.49	1.41	.166	3.40	.017
	0	1.33	.49	1.66	.104	3.37	.001
REM	3	2.49	.65	1.23	.226	3.38	.001
	6	2.85	.92	1.13	.266	2.14	.032
	9	3.54	1.15	1.02	.319	1.88	.060
	PA	.08	.02	3.06	.004	3.04	.002
	0	.10	.02	3.66	.001	4.07	.000
PEM	3	.14	.03	3.01	.005	3.99	.000
	6	.16	.04	2.54	.017	2.94	.003
	9	.21	.05	2.25	.033	2.38	.017
	PA	1.09	.42	1.59	.117	2.64	.008
	0	1.46	.76	.98	.329	3.68	.002
SEM	3	4.07	1.12	1.28	.207	3.56	.000
	6	4.96	2.08	1.04	.306	2.40	.016
	9	8.35	3.27	1.02	.316	2.09	.036

TABLE A 4

DIFFERENCES IN ERROR MEASURES BETWEEN GROUPS

YEAR 4

		MEANS		t-TEST		WILCOXON	
Error	Month	Failing	Healthy	t	α	Z	α
	PA	.57	.27	1.67	.104	1.84	.065
	0	.62	.26	1.73	.092	2.59	.010
AEM	3	.91	.39	2.15	.039	1.72	.085
	6	1.70	.56	2.13	.044	2.42	.016
	9	2.21	.77	1.98	.060	1.50	.134
	PA	.80	.16	1.47	.152	2.27	.023
	0	.92	.16	1.53	.137	2.46	.014
REM	3	1.06	.28	1.65	.111	1.83	.066
	6	1.45	.36	2.05	.051	2.45	.014
	9	1.59	.48	2.04	.052	1.61	.107
	PA	.05	.02	2.05	.049	2.97	.003
	0	.05	.02	2.04	.050	2.79	.005
PEM	3	.06	.03	2.51	.017	2.34	.019
	6	.09	.04	2.10	.046	2.35	.019
	9	.12	.05	1.95	.064	1.52	.128
	PA	.76	.13	1.75	.090	2.30	.021
	0	.85	.14	1.76	.090	2.53	.011
SEM	3	1.30	.24	2.15	.040	1.98	.048
	6	4.47	.52	1.64	.114	2.84	.005
	9	7.79	.76	1.40	.177	1.87	.062

firms? Givoly and Lakonishok [1984] say that if forecasts are rational they should be the most accurate forecasts. This suggests that no mechanical model should out-perform analysts. Thus a comparison of analyst accuracy with mechanical model accuracy could have implications for the rationality of analysts forecasts.

Tests were conducted in both the failing and healthy groups to compare analysts with a naive no-change model that predicted next period earnings as equal to last period earnings. Thus earnings are forecasted by model F1 as presented in section 3.4.1. (Analogous tests were conducted using models F2 and F3, where data permitted. Findings using models F2 and F3 are not reported here but are consistent with the findings using model F1).

The procedure used was as follows. For each year the four error measures (AEM, REM, PEM, SEM) were determined using the previous years earnings as the forecast. Differences in forecast errors between analysts and the naive prediction were determined by subtracting the naive prediction error measure from the corresponding analyst forecast error measure. The mean error differences would be zero if analysts and the naive model were equally accurate, negative if analysts were more accurate, and positive if the naive model was more accurate. T-tests of the null hypothesis that the mean error difference is zero, using error differences from the four error measures (AEM, REM, PEM, SEM) are summarized in Table A-5.

A fifth variable was also determined as an alternative approach to comparing accuracy. A log relative error measure (LEM) was determined as follows

$$\text{LEM} = \ln\left(\frac{\text{Actual} - \text{Analyst forecast}}{\text{Actual} - \text{Naive forecast}}\right)$$

This measure will also equal zero when analysts are equal in accuracy to the naive prediction and be negative (positive) when analysts are more (less) accurate. T-tests for the null hypothesis that LEM equal zero are also contained in Table A-5.

Findings for healthy firms are quite consistent across the three years and selected months reported in Table A-5. Signs are almost universally negative, suggesting analysts are more accurate than the naive prediction. The mean error differences grow in absolute magnitude as the reporting date approaches indicating that relative accuracy improves. By three months prior to the end of the report year one can reject the null hypothesis of equal analyst vs. naive model forecast accuracy (at .05 or better) across most of the five error measures in the three year period presented. As the report date approaches, the superiority of the analysts over the naive model becomes stronger.

The picture is not as clear for the failing firms. In year 1, signs are as often positive as negative. In no case can the null hypothesis be rejected. It is apparent that when failure is near, analysts have no advantage over a naive no-change model in forecasting actual earnings. Year 2 findings are quite different. Error difference signs are predominately negative, and significantly negative by month 0. Thus by the end of the reporting year analysts out-perform the Naive forecast.

TABLE A 5

MEAN ERROR DIFFERENCES: ANALYST VS. NAIVE MODEL

		FAILING FIRMS										HEALTHY FIRMS									
		AEM		REM		PEM		SEM		LEM		AEM		REM		PEM		SEM		LEM	
YEAR	MO.	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T	MEAN	T
1	PA	-.61	-.86	.05	.10	-.03	-.19	.35	.15	-.02	-.07	-.54	-4.30***	-.98	-2.41*	-.04	-3.68***	-2.81	-2.16*	-1.22	-6.35***
	0	-.35	-.58	-.23	-.75	.09	.34	-1.26	-1.33	.03	.10	-.51	-3.96***	-1.09	-2.08*	-.03	-3.09**	-2.84	-2.05*	-.95	-4.83***
	3	.46	.80	-.05	-.22	.20	.90	-.12	-.13	.19	.71	-.37	-3.45***	-.43	-2.15*	-.02	-2.42*	-1.61	-1.97	-.75	-4.33***
	6	1.27	1.53	.21	1.21	.56	1.29	1.17	.97	.61	2.67+	-.22	-1.97	-.05	-.23	-.01	-.90	-.47	-.53	-.24	-1.50
	9	1.46	2.08	.41	2.94++	.66	1.43	2.02	2.38+	.66	3.05++	-.04	-.38	.33	1.20	.01	.86	1.07	.82	-.09	-.59
2	PA	-1.14	-2.98**	-1.09	-2.79**	-.22	-2.89**	-2.60	-2.22*	-.93	-4.58***	-.34	-4.09***	-.26	-3.15**	-.02	-4.09***	-.28	-1.29	-1.15	-5.71***
	0	-.96	-2.74**	-.90	-2.84**	-.17	-2.72*	-2.15	-2.31*	-.75	-4.06***	-.29	-3.51***	-.23	-3.03**	-.02	-3.49***	-.32	-1.81	-.80	-4.26***
	3	-.44	-2.11*	.10	.25	-.04	-1.44	-.69	-1.18	-.41	-1.80	-.15	-1.62	-.14	-1.96	-.01	-2.25*	-.17	-.89	-.27	-1.65
	6	-.05	-.27	.56	1.14	.03	.91	.97	1.33	-.12	-.55	-.10	-1.29	-.08	-1.20	-.01	-1.44	-.16	-.99	-.14	-.83
	9	.87	2.42+	.60	2.05+	.14	2.74+	2.72	3.09++	.10	.51	.00	.00	-.02	-.29	.00	.18	-.17	-.92	.02	.08
3	PA	-1.24	-1.77	-.54	-1.55	-.11	-1.89	-2.09	-1.82	-.67	-2.86**	-.36	-2.67**	-.32	-1.31	-.02	-3.22**	-1.36	-1.39	-.77	-3.69***
	0	-1.30	-1.75	-.47	-1.27	-.11	-1.98	-2.13	-1.75	-.55	-2.32*	-.36	-2.57*	-.32	-1.47	-.02	-2.92**	-1.11	-1.01	-.80	-4.09***
	3	-.63	-1.42	.69	.99	-.05	-1.52	.42	.23	-.21	-.97	-.29	-2.20*	-.20	-.99	-.01	-2.54*	-.82	-.71	-.72	-3.74***
	6	-.33	-.73	1.08	1.40	-.02	-.60	1.34	.67	.05	.22	-.14	-1.22	.01	.03	-.01	-1.10	-.01	-.01	-.48	-2.17*
	9	.46	1.22	1.83	1.34	.04	1.56	5.08	1.32	.23	1.09	-.00	-.01	.24	1.03	.00	.61	1.21	.76	-.09	-.45

Significance Levels (negative error difference): .05*/.01**/.001***
(postive error difference): .05+/.01++/.01+++

Surprisingly, error differences in year 3, though predominantly negative, are not significant. One might expect the poor analyst performance in year 1 just prior to failure; there would likely be considerable uncertainty concerning earnings. The lack of statistically significant analyst superiority over the naive model in year 3 indicates that the difficulty in out-performing the naive prediction is not limited to the year immediately prior to failure. Greater uncertainty associated with failing firms may prevent analysts from out-performing the naive prediction even when failure is not immediately imminent.

One can compare error differences between the failing and healthy groups using t-tests and Wilcoxon. Comparison of AEM, REM, PEM, SEM error differences are not fully valid. If one assumes greater uncertainty in the failing firm group, then errors for both analysts <u>and</u> the naive model should be larger and the error difference measures may be correspondingly larger, even though the relative superiority (inferiority) of analysts compared to the naive model is the same as in the healthy group. The only fair comparison is using LEM, because the magnitude of the analyst error is deflated by the naive model error. Tests involving LEM (not reported) show LEM to be significantly higher (more positive or less negative) for failing firms in year 1. No significant difference occurs in years 2 or 3.

3.6 Dispersion of Forecasts

Previous research has investigated the variation or dispersion across analysts forecasts as a measure or indication of uncertainty. Cukierman and Givoly [1982] present a model in which the dispersion of forecasts across forecasters is positively associated with the dispersion of the distribution of expected earnings and therefore with the cross-sectional error in forecasts. Empirical evidence supported their model; measures of dispersion were positively associated with measures of forecast error. Results from Elton, Gruber and Gultekin [1984] also document this relationship. Cukierman and Givoly argue that the cross-sectional error in earnings is the empirical counterpart of uncertainty. Dispersion of earnings forecasts have also been found to be associated with traditional risk measures such as beta, the standard deviation of returns and earnings growth variability (Givoly and Lakonishok [1983].)

The purpose of this section is to determine if measures of dispersion in forecasts differ systematically between failing and healthy firms as bankruptcy approaches, with the purpose of identifying dispersion related measures that may be useful in distinguishing failing from healthy firms. The implicit assumption is that forecast dispersion measures may reflect risk that is ultimately manifested in bankruptcy.

3.6.1 Dispersion Measures

For each company for each month, IBES provides information on the high estimate among the individual analysts forecasts available, the low estimate, and the standard deviation of the estimates. Use of the standard deviation as a measure of dispersion is frequent in other studies(e.g. Crichfield, Dyckman and Lakonishok [1978]). Others (e.g., Castanias and Griffin [1984]) suggest that it is not entirely appropriate to combine standard deviations cross-sectionally and suggest the use of alternative measures of dispersion. Tests were conducted using the seven different dispersion measures to follow. Only observations where the number of forecasts per firm was greater than or equal to two were included. Findings are remarkably similar across the measures making the choice of measure somewhat moot. Test results for a representative selection of the seven measures are reported. The seven measures of dispersion are

1. The standard deviation (STND) of earnings forecasts across analysts for individual firms at a particular month. This measure is given by IBES.

2. The variance (VAR), calculated by squaring the STND.

3. The natural log of the variance (LVAR), designed to reduce potentially undesirable effects of not scaling the data for firm size.

4. The coefficient of variation (CVAR), calculated by deflating the STND by the absolute value of the mean forecast, again an attempt to deflate.

5. The price deflated variance (PVAR), determined by dividing the variance by stock price in the corresponding month, an alternative deflating approach.

6. The range (RNG), determined by the high and low forecasts in

any given month on IBES.

7. The price-deflated range (PRNG), dividing the range by stock price to adjust for size effects.

3.6.2 Dispersion Tests

Both parametric (t-tests) and non-parametric (Wilcoxon) tests were conducted to determine if the dispersion of analysts forecasts for failing firms was systematically different than for healthy firms. Given the previous research suggesting that forecast dispersion captures uncertainty, one would hypothesize greater dispersion for failing firms than healthy firms and increasing dispersion as bankruptcy approaches.

Findings from t-tests were consistent with findings from the wilcoxon tests. Results from the wilcoxon tests for three representative measures, the standard deviation, the price-deflated variance and the range, appear in Table D1. (Note that because wilcoxon tests are based on ranks, results using the STND, VAR and LVAR are identical).

The findings reported in Table D1 are not dependent on the specific dispersion measure. In year one, failing firms have significantly ($\alpha \leq .05$) greater dispersion across analysts for a full year prior to fiscal year end. There is consistent indication that uncertainty as reflected in forecast dispersion is associated with impending failure. There is also a steady increase in the failing group average STND, PVAR and RNG as the earnings announcement date approaches. Thus there is also some indication that

TABLE D1
DISPERSION OF FORECASTS - BY GROUPS

		STANDARD DEVIATION				PRICE-DEFLATED VARIANCE				RANGE			
		AVERAGE		WILCOXON		AVERAGE		WILCOXON		AVERAGE		WILCOXON	
YEAR	MONTH	FAILING	HEALTHY	Z	α	FAILING	HEALTHY	Z	α	FAILING	HEALTHY	Z	α
1	PA	1.37	.24	3.07	.002	3.21	.03	3.62	.000	2.49	.70	3.07	.002
	0	.66	.22	3.10	.002	.61	.01	3.82	.000	1.47	.57	2.50	.012
	3	.49	.23	2.94	.003	.10	.01	4.17	.000	1.12	.57	2.51	.012
	6	.44	.25	3.38	.001	.05	.01	4.50	.000	.98	.66	2.28	.023
	9	.44	.24	2.29	.022	.06	.01	3.39	.001	.97	.61	1.79	.074
	12	.45	.22	2.80	.005	.05	.01	3.36	.001	.96	.60	2.03	.042
	15	.34	.27	1.49	.134	.03	.01	2.29	.022	.73	.74	.86	.392
	18	.32	.38	.07	.946	.01	.01	.42	.675	.66	.94	-.69	.490
2	PA	.49	.17	2.37	.018	.21	.01	2.84	.005	.99	.40	2.01	.039
	0	.66	.19	2.01	.044	.64	.01	2.57	.010	1.38	.43	1.53	.126
	3	.43	.25	2.50	.013	.05	.01	3.15	.002	1.01	.60	1.94	.052
	6	.29	.25	1.69	.092	.02	.01	2.23	.026	.63	.65	.83	.410
	9	.31	.29	1.18	.238	.02	.01	1.42	.155	.69	.77	.73	.467
	12	.39	.25	1.76	.079	.03	.01	1.99	.046	.83	.64	1.36	.173
	15	.35	.27	1.20	.229	.02	.01	1.54	.123	.77	.68	1.09	.275
	18	.24	.27	.49	.626	.01	.01	.22	.825	.55	.65	.24	.808
3	PA	.18	.13	1.65	.099	.01	.00	1.86	.064	.49	.32	1.47	.140
	0	.27	.18	2.47	.013	.02	.01	2.49	.013	.65	.43	2.09	.037
	3	.41	.21	1.73	.083	.00	.01	1.89	.058	.82	.51	1.32	.188
	6	.47	.29	1.13	.257	.39	.01	1.03	.303	.90	.68	.99	.319
	9	.20	.24	-.02	.981	.00	.01	.14	.892	.53	.60	.00	.999
	12	.26	.25	1.15	.247	.01	.01	1.21	.227	.67	.63	.86	.389
	15	.58	.21	2.43	.015	.06	.01	2.32	.021	1.13	.52	2.41	.016
	18	.33	.29	1.17	.243	.01	.01	1.36	.173	.83	.68	1.09	.275
4	PA	.24	.10	1.85	.064	.03	.00	1.76	.079	.50	.28	1.60	.109
	0	.29	.13	1.59	.112	.03	.00	1.59	.110	.61	.32	1.74	.081
	3	.32	.14	1.28	.200	.04	.00	1.39	.163	.71	.40	.94	.346
	6	.30	.22	.61	.540	.02	.01	.58	.563	.66	.48	.92	.356
	9	.25	.19	1.14	.251	.01	.00	1.17	.241	.54	.43	1.18	.236
	12	.15	.17	-.28	.777	.00	.00	-.51	.609	.44	.36	.55	.584
	15	.18	.28	-.43	.665	.00	.01	-.63	.525	.48	.55	.07	.939
	18	.28	.29	.51	.609	.01	.01	.77	.443	.61	.61	1.15	.250

uncertainty increases as the year end (and thus bankruptcy) approaches.

In contrast, in year four there is no indication of a difference in dispersion for any of the three measures at any of the months. Uncertainty associated with future failure is not reflected in significantly less agreement by analysts in earnings predictions. One year closer to failure, in year three, the first indications of wider dispersion for failing firms shows up; significantly larger dispersion measures are found for failing firms in months zero and 15. Another year closer to failure, year two, provides more convincing evidence;
significantly larger dispersion measures are found for failing firms starting about six months prior to year end and continuing through the earnings announcement.

In short, dispersions measures may reflect uncertainty or risk associated with impending failure and may contain information useful for distinguishing between healthy and failing firms. The results suggest that significant differences in forecast dispersions start appearing and continue to exist from about 18 months prior to the end of the last fiscal year before to the failure year, although hints of greater dispersion for failing firms may be evident even before that period.

3.7 Revisions in Forecasts

Research on revisions in earnings forecasts by analysts is limited. Past investigations have attempted to characterize the forecast formation process by observing the relationship between forecast revisions and information hypothesized to motivate revision. Studies by Brown and Rozeff [1979], Abdel-Khalik and Espejo [1978] and Givoly [1985] provide evidence that forecasters adapt their predictions in a manner consistent with the information contained in successive quarterly or annual earnings releases, lowering (raising) predictions when past earnings were over-predicted (under-predicted). Both Givoly and Lakonishok [1979] and Brown, Foster and Noreen [1985] provide evidence for a small but apparent positive dependence over time in the direction (up or down) of mean forecast revisions. Brown, Foster and Noreen suggest this may happen either because analysts are "spreading" their revision over separate successive months or because separate analysts do not adjust simultaneously to the same information. There does not appear to be any strong clustering of forecast revisions around the time of interim or annual earnings releases, and revisions, up or down, tend to be similar in frequency; there is however some tendency for a greater frequency of revisions within the twelve months immediately prior to an annual earnings release when compared to the period 18 - 24 months prior to release (Brown, Foster and Noreen).

Our concern here is whether forecast revision behavior may differ systematically between failing and healthy firms such that measures related to revisions may indicate future failure. One might suggest various possibilities. First, greater uncertainty associated with failing firms may be reflected in more frequent revisions. On the other hand, a tendency for bad news to be withheld by management may result in less information being released for failing firms and consequently less frequent revisions. Second, poor performance associated with impending failure should lead to more frequent downward earnings revisions and perhaps less frequent upward revisions. Third, if in fact bad news is delayed by failing firms in the hopes of a cure or turn around, there may be a tendency for bad news and therefore downward revisions to cluster toward year end or just prior to release of the actual earnings figure.

3.7.1 Revision Measures

Three different measures of the frequency of analyst forecast revisions were determined. A measure of the frequency of total revisions (up and down) per month was calculated by dividing the total number of revisions in a given month by the number of analysts estimates that were available in that month. These monthly percentage figures were averaged over multi-month periods to smooth out the measures. Frequency of revision measures were averaged over the following three month periods: months 0 - 2, 3 - 5, 6 - 8, 9 - 11; the following six month periods: 0 - 5 and 6 - 11; the full year prior to year end: 0 - 11; and over the 3 months immediately prior to the release of the actual figure: designated PA-PA3. Measures of the frequency of downward revisions and upward revisions were determined in an analogous manner, using the number of downward or upward revisions individually.

3.7.2 Revision Tests

Both t-tests and wilcoxon tests for differences in group means between failing and healthy firms were conducted for the three types of revision measures averaged across the indicated periods in the four years prior to bankruptcy. Results from t-tests and wilcoxon tests were the same; however, wilcoxon tests rest on less restrictive distributional assumptions. Table R1 contains group means, wilcoxon test Z scores and probability levels for rejection of the null hypothesis of no difference between groups. The majority of the test statistics are insignificant at traditional alpha levels (i.e., .05 or lower) but several tests are significant and some regularities exist in the table.

For both groups, down revisions are much more frequent than up revisions. Down revisions tend to be at least four times more frequent regardless of what time period revision measures are averaged over. Since forecast accuracy tends to improve as year end approaches (i.e., as information arrives to forecasters and they update) the dominance of downward revisions indicates over optimistic forecasts occur early and are corrected as year end approaches.

The preponderance of negative Z values for the tests involving upward revisions is consistent with more frequent upward revisions for healthy firms; however, only three of the Z values are significant for the various periods reported in the table.

All except three Z values are positive for the downward revision tests, and about a third of the Z statistics are significant. Thus there is consistent, frequently significant, evidence

TABLE R1
FREQUENCY OF REVISIONS IN FORECASTS - BY GROUP

		MONTHLY AVERAGE % UP REVISIONS				MONTHLY AVERAGE % DOWN REVISIONS				MONTHLY AVERAGE % ALL REVISIONS			
		MEANS		WILCOXON		MEANS		WILCOXON		MEANS		WILCOXON	
YEAR	MONTHS	FAILING	HEALTHY	Z	α	FAILING	HEALTHY	Z	α	FAILING	HEALTHY	Z	α
1	0-2	.9	.9	-1.53	.125	6.4	4.3	2.33	.020	7.3	5.1	2.37	.018
	3-5	.6	.4	-1.28	.201	4.9	4.8	-.38	.701	5.5	5.2	-.09	.922
	0-5	.8	.7	- .71	.477	5.7	4.5	1.76	.070	6.4	5.1	1.96	.050
	6-8	.4	.5	-1.42	.154	5.2	4.5	.88	.380	5.6	5.0	.80	.422
	9-11	.9	1.1	-1.62	.105	4.0	4.2	-.62	.538	4.9	5.4	-.70	.481
	6-11	.7	.8	-1.95	.051	4.9	4.4	.93	.349	5.5	5.3	.73	.467
	0-11	.8	.7	-.74	.457	5.4	4.5	1.81	.070	6.2	5.2	1.73	.084
	PA/PA-3	.7	1.1	-2.28	.023	5.0	4.5	.68	.492	5.8	5.6	-.14	.886
2	0-2	1.2	1.1	-1.22	.221	6.2	4.4	2.25	.025	7.4	5.5	2.56	.010
	3-5	.3	.8	-3.22	.001	5.5	4.0	2.05	.040	5.8	4.8	1.27	.201
	0-5	.7	.8	-1.76	.078	5.8	4.2	2.63	.009	6.5	5.1	2.53	.011
	6-8	.3	.3	-.88	.376	5.4	3.6	2.45	.014	5.7	3.9	2.69	.007
	9-11	.9	.9	-.66	.508	4.0	3.9	-.12	.906	4.8	4.9	-.18	.854
	6-11	.5	.6	.69	.492	4.7	3.8	1.29	.196	5.2	4.4	1.28	.199
	0-11	.7	.7	-.47	.634	5.1	3.9	2.64	.008	5.8	4.6	2.64	.008
	PA/PA-3	1.3	1.9	-2.37	.018	5.3	4.5	.71	.480	6.7	6.4	-.09	.923
3	0-2	.8	1.5	-2.38	.017	4.3	3.9	.76	.447	5.2	5.4	.11	.907
	3-5	.5	1.1	-.68	.499	5.3	2.8	2.67	.008	6.0	3.9	2.31	.021
	0-5	.7	1.2	-1.73	.084	4.9	3.4	2.11	.034	5.7	4.6	1.50	.134
	6-8	.3	.4	-.73	.463	4.5	3.9	.62	.535	4.9	4.3	.64	.520
	9-11	1.5	1.5	.14	.892	3.6	3.0	.06	.949	5.1	4.4	.41	.678
	6-11	1.0	1.0	.17	.864	4.3	3.5	.70	.482	5.3	4.5	.84	.402
	0-11	.9	1.1	-.13	.894	4.7	3.6	1.60	.109	5.6	4.7	1.50	.132
	PA/PA-3	1.4	1.9	-1.27	.203	4.4	4.4	.27	.786	5.8	6.3	-.44	.658
4	0-2	1.5	1.7	-.43	.667	3.9	3.0	1.04	.299	5.4	4.5	.98	.326
	3-5	1.1	1.0	.03	.979	3.5	2.2	1.52	.128	4.5	3.2	1.49	.137
	0-5	1.2	1.4	-.76	.450	3.7	2.6	1.51	.129	5.9	3.8	1.43	.153
	6-8	.1	.2	-.58	.562	4.0	2.3	1.73	.084	4.1	2.4	1.73	.084
	9-11	.5	.9	-.63	.523	4.1	2.5	1.11	.266	4.6	3.5	.67	.502
	6-11	.4	.6	-.94	.348	3.6	2.2	2.37	.017	4.0	2.8	1.84	.065
	0-11	.8	1.1	-1.22	.223	3.6	2.4	2.16	.031	4.4	3.5	1.91	.057
	PA/PA-3	1.9	2.3	-.55	.584	4.2	2.7	1.27	.203	6.1	5.0	.70	.483

that failing firms have a higher frequency of downward revisions. There is also an indication that downward revisions as a percentage of total estimates are more frequent in the 0 - 5 month period rather than in the earlier 6 - 11 month period, and a tendency for larger and more significant Z values to be associated with the 0 - 5 month period (except in year 4). This suggests an increase in downward adjustments, particularly for failing firms, as year end approaches.

The pattern of measures, signs and size of Z statistics, and significance levels for the all (both up and down) revisions column in the table are very similar to the downward revisions column. Analysts make more forecast revisions in general for failing firms, and the frequency of revisions appear greater toward year end. But given that downward revisions are far more frequent than upward ones in the sample, the results for the all revisions tests seem to be driven by the more frequent downward revisions for the failing firm group.

To summarize, there are more frequent downward revisions for failing firms appearing in all four years prior to bankruptcy, and a tendency for the frequency to be greater in the last six months of each year. However, there is no particular individual three month, six month or 12 month period in which the frequency of downward revisions is significantly greater for failing firms in even three out of the four years tested. Thus there is no particular individual revision measure that emerges as one that might consistently help to distinguish failing from healthy firms.

3.8 Mean Forecast Values

One obvious place to look for differences between failing and healthy firms is simply in the level of future earnings predicted for firms in each group. Although low earnings does not imply bankruptcy and high earnings does not insure health, one would expect some relationship between the level of earnings and the probability of future failure. While reported earnings may contain information relevant to distinguishing between groups, forecasted earnings are future looking and consequently have the potential of reflecting aspects of firm health that have not yet been reflected in reported earnings.

Mean forecasted earnings (the average across analysts) was used to determine if there were significant differences between failing and healthy firm groups. Two measures were investigated: the undeflated mean forecasted earnings per share and the mean forecasted earnings per share deflated by stock price to adjust for size differences. Both parametric (t-tests) and non-parametric (Wilcoxon) tests for group differences were conducted. Results for both t-tests and Wilcoxon tests applied to both undeflated and deflated measures of earnings provided similar findings, hence only the results of the Wilcoxon tests on undeflated measures are reported. Table ME 1 provides the average mean forecasted earnings per share for both failing and healthy firm groups at 3 month intervals within each year for the four year period prior to bankruptcy.

Several findings are of note: First, there is a clear

TABLE ME 1
GROUP DIFFERENCES IN MEAN FORECASTS

YEAR	MONTH	AVERAGE FORECAST FAILING	AVERAGE FORECAST HEALTHY	WILCOXON Z	α
1	PA	-1.30	1.55	-5.43	.000
	0	-.86	1.57	-6.17	.000
	3	-.27	1.77	-6.15	.000
	6	.44	1.95	-4.91	.000
	9	.88	2.02	-4.23	.000
	12	1.11	2.13	-3.88	.000
2	PA	-.97	1.67	-5.01	.000
	0	-.81	1.69	-4.94	.000
	3	.07	1.87	-4.25	.000
	6	.58	2.10	-3.36	.001
	9	1.24	2.14	-2.99	.003
	12	1.59	2.23	-1.87	.061
3	PA	.18	1.77	-3.22	.001
	0	.22	1.91	-3.14	.002
	3	.95	1.97	-1.97	.049
	6	1.24	2.07	-1.62	.105
	9	1.78	2.20	-1.06	.291
	12	2.16	2.28	.20	.839
4	PA	1.16	1.92	-1.67	.093
	0	1.11	1.96	-1.78	.075
	3	1.59	1.97	-.89	.373
	6	2.09	1.93	-.09	.924
	9	2.33	1.82	.21	.835
	12	2.85	1.74	1.74	.082

deterioration of the level of earnings forecasted for failing firms across the years as bankruptcy approaches. Second there is a deterioration of earnings within each forecast year as the year end approaches. This holds for both failing and healthy groups (except in year 4 for the healthy firms). Generally the within year deterioration for failing firms appears more severe. (Explicit tests of within year changes follow in a later section.) Third, there are significant group differences in the level of forecast earnings evident in all years except year 4. Looking at Z values and significance levels, the differences become more pronounced as year-end approaches within each year and as the year of bankruptcy approaches. None of these findings is surprising; if information relevant to impending failure is reflected in earnings forecasts, and their change over time, one would expect more significant differences as the time remaining prior to failure decreases. The finding do however suggest some potential ability of measures of mean forecasted earnings to distinguish between the groups.

3.9 Intra-year Changes

The previous sections in this chapter have reported on different properties of forecasts (bias, incorporation of information, accuracy, dispersion, frequency of revision and mean forecast level) as captured by various measures at particular points in time (or, in the case of revision measures, averages over short periods of time). The questions asked were whether particular properties differed between failing and healthy firms. This section expands the analysis by addressing the question of whether <u>changes</u> in particular measures within a given forecast year differ between groups.

Past research has provided evidence on two types of changes that may occur within a given forecast year. One consistent finding is that forecasts become more accurate as the time to the announcement of actual earnings decreases (e.g., Elton, Gruber and Gultekin [1984]). A viewing of the error measures reported in Tables A-1 and A-2 suggest this result holds for firms in the present sample also. A second finding concerns agreement among analysts forecasts. Brown, Foster and Noreen [1985] find that dispersion of forecasts across analysts decreases systematically as announcement date approaches. Crichfield, Dyckman and Lakonishok [1978] and Elton, Gruber, and Gultekin [1984] both find hints of a decrease in dispersion but not convincing evidence.

Four questions related to intra-year changes were investigated in the present study:

a. Do forecast errors decrease as the time remaining to announcement decreases?

Given the strong results from previous studies one would expect increasing accuracy as the announcement date approaches for both failing and healthy firms. This is expected since more and better information becomes available to estimate as the period progresses. Forecast errors, however, are associated with risk and uncertainty. For the failing firm group, the passage of time also implies the approach of bankruptcy. One might hypothesize that approaching bankruptcy will result in a relatively smaller increase in forecast accuracy (less decrease in forecast error) for failing firms. On the other hand, forecasts for failing firms are less accurate. Thus there is more room for "improvement" as new information arrives throughout the year. The larger forecast errors for failing firms may allow for greater improvement and thus greater decrease in forecast errors. These two arguments suggest competing reasons for systematic group differences in changes in forecast accuracy as the announcement date approaches.

b. Does dispersion among forecasters decrease as the time remaining to announcement decreases?

Again, the expectations are that dispersion across forecasters should decrease as year end approaches, but because dispersion measures are also indicators of risk and uncertainty, approaching bankruptcy may result in a relatively smaller decrease in dispersion for the failing firms group.

c. Does the frequency of revisions in forecasts change as the time remaining to announcement decreases?

Revisions in forecasts for both failing and healthy firms should occur as relevant information becomes available. Firms may have incentives to withhold bad news, and failing firms may have relatively more bad news to withhold. Consequently, there may be a tendency for delayed information release by failing firms and perhaps more revision behavior by analysts as year end approaches. One might hypothesize both a higher relative frequency of forecast revisions, and particularly a higher relative frequency of downward revisions as year-end approaches for failing firms.

d. Do mean forecasts decrease as the time remaining to announcement decreases?

If, as research findings suggest, forecasts are both somewhat optimistic and distinctly more accurate as announcement approaches, then both failing and healthy firms may experience a trend toward lower forecast earnings estimates as the year end approaches. Assuming more negative information becomes available for failing firms, one would expect larger and more negative forecast revisions (i.e., changes in mean forecast estimates) for failing firms.

In short, the above questions ask if there are significant intra-year changes in a) forecast errors, b) forecast dispersion, c) frequency of forecast revisions, and d) mean forecasts.

3.9.1 Measures for Intra-year changes

Four measures of forecast error (absolute, relative, price-deflated, squared), seven measures of forecast dispersion (standard deviation, variance, log variance, coefficient of variation, price-deflated variance, range, price-deflated range), three measures of frequency of forecast revisions (up, down, total), and two measures of the level of mean forecasts (undeflated and price-deflated) were previously explained.

For each of the above variables, measures of the change between values of the variables determined at different months within each forecast year were determined. Changes between monthly values were determined for the following pairs of months.

Later Month	Earlier Month
PA	0
PA	3
PA	6
PA	9
0	3
0	6
0	9
3	6
3	9
6	9

(For the frequency of revision measures, averages over the three months ending in the month listed above were used rather than the point estimates at the months listed above).

For each of the variables of interest, three measures of the change in the variable between the two monthly values were computed:

1. Absolute change = (VL − VE)
2. Relative change = (VL − VE) / |VE|
3. Log relative Change = ln (VL/VE)

Where VL equals the value of the variable at the later month and VE equals the value of the variable at the earlier month.

The log relative measure is preferable because it adjusts for the relative size of the variables and captures only the relative change in magnitude. It also is symmetric for increases and decreases and reduces outlier effects. Unfortunately, the log relative is undefined for zero or negative values of the arguments, which causes substantial sample size reduction.

The relative change measure also has merit because it deflates the difference in values across time for magnitude.

However, it is undefined in some cases, and small or negative values of the arguments may cause a misleading indication of the change that has occurred.

The simple absolute difference measure can be criticized for not adjusting for the absolute magnitudes of the measures, however it is defined for all values. Furthermore many of the variables for which changes are to be determined have already been deflated for size effects.

3.9.2 Tests of Intra-year Changes

Two kinds of questions were addressed. First t-test were conducted to test the null hypothesis that the change measures equaled zero. These tests were conducted separately for failing and healthy firms. Second, Wilcoxon tests were conducted to determine if change measures were different for the failing and healthy groups. All tests were conducted for each of the four years prior to bankruptcy. The t-tests of no significant difference from zero are of secondary interest. The between group tests, however, relate directly to contrasts between failing and healthy firms in intra-year changes and consequently have potential relevance in identifying measures that may be useful for distinguishing between the groups.

Findings using change measures 2 and 3 (the relative change and log relative change) can be summarized as follows: As expected, changes in forecast errors were consistently significantly ($\alpha \leq .05$) negative indicating increasing accuracy as the announcement date approaches. Changes in the frequency of revisions (up, down, and total) were generally positive and frequently significant, indicating more frequent revisions in analysts forecasts as announcement approaches. Changes in dispersion measures were not consistent in sign and were only occasionally significant, but hinted at a decrease in dispersion as announcement approaches. These findings were present within both groups of firms. However, for tests of differences in intra-year changes in forecast errors, forecast dispersion, and frequency of forecast revisions <u>between</u> failing and healthy firms, across all versions of measures, across all four years, results were consistently insignificant.

In short, test results using measures 2 and 3 indicated that accuracy and frequency of forecast revision increase, and dispersion perhaps decreases, as year end approaches, but these observations tend to hold for both failing and healthy firms. There was no significant systematic group difference in intra-year changes in errors, dispersion or revision frequency measures indicated by the tests, and consequently no specific measures were identified that could prove useful in distinguishing between the two groups of firms. Because of this and because the number of tests conducted were extensive, no tabular results using measures 2 and 3 are reported here. (Test results using change measures 2 and 3 to measure intra-year changes in the level of mean forecast were not valuable because of the large number of negative forecasts values in the failing firm group.)

Our remaining interest is then with results using change measure 1, the absolute difference in variables between two months with the year. Some tests provided results worth mention:

1. Findings (relating to question d above) concerning intra-year changes in mean forecasts are interesting. Table C1 shows results for the four year period. Several observations are of note. First, mean values of the change measures are consistently negative in sign for failing firms for the full four-year test period. Most all of the t values are significant except in year four. This provides strong evidence of optimistic forecasts that are subsequently and consistently revised downward as each year end, and bankruptcy, approaches. Second, there are also consistently significant negative changes in years 1 and 2 for the healthy firms. This is not surprising. Forecasts in general could be optimistic and require downward revisions if general economic conditions were deteriorating. The fact that years 1 and 2 reflect times just prior to bankruptcy for the failing firms, coupled with the fact that bankruptcies increase in frequency during periods of overall economic stagnation, is consistent with those years reflecting periods of optimistic forecasts even for the healthy firm group. Third, and most important, despite downward revisions for both groups, the magnitude of the changes in mean forecasted earnings are substantially greater for the failing group. Wilcoxon Z values are universally negative over the four years, universally significant in years 1 and 2, and frequently significant in years 3 and 4. In short, intra-year changes in mean forecasts apparently provide possible measures for distinguishing the groups.

TABLE C 1

INTRA-YEAR CHANGES IN MEAN FORECASTS –
(Absolute difference in undeflated mean forecasted earnings)

		FAILING			HEALTHY			WILCOXON	
	MONTHS	MEAN	t	α	MEAN	t	α	z	
1	P/0	-.71	-2.51	.017	-.06	-2.41	.019	-2.87	.0
	P/3	-1.60	-3.66	.001	-.24	-3.52	.001	-3.95	.0
	P/6	-2.47	-4.55	.000	-.36	-3.67	.001	-4.70	.0
	P/9	-2.80	-4.98	.000	-.45	-3.59	.001	-4.86	.0
	0/3	-.72	-3.27	.002	-.19	-3.31	.002	-2.99	.0
	0/6	-1.34	-4.36	.000	-.29	-3.40	.001	-3.75	.0
	0/9	-1.89	-4.68	.000	-.38	-3.31	.002	-4.54	.0
	3/6	-.52	-4.39	.000	-.11	-2.62	.011	-2.28	.0
	3/9	-1.09	-3.80	.001	-.21	-2.81	.007	-3.72	.0
	6/9	-.48	-1.77	.085	-.09	-2.40	.019	-2.11	.0
2	P/0	-.31	-3.98	.000	-.02	-.61	.544	-3.19	.0
	P/3	-1.03	-5.15	.000	-.19	-3.07	.003	-4.72	.0
	P/6	-1.58	-5.12	.000	-.29	-4.20	.000	-4.32	.0
	P/9	-2.39	-3.52	.001	-.36	-3.21	.002	-4.02	.0
	0/3	-.89	-3.81	.000	-.17	-2.82	.006	-3.71	.0
	0/6	-1.51	-4.12	.000	-.27	-3.95	.000	-3.66	.0
	0/9	-2.28	-3.44	.001	-.35	-3.29	.002	-3.46	.0
	3/6	-.58	-4.05	.000	-.09	-1.24	.221	-3.49	.0
	3/9	-1.26	-2.98	.005	-.16	-1.47	.147	-2.83	.0
	6/9	-.66	-2.10	.041	-.07	-1.00	.323	-2.66	.0
3	P/0	-.12	-3.32	.002	-.07	-1.91	.062	-1.26	.2
	P/3	-.75	-2.26	.011	-.11	-1.79	.079	-2.15	.0
	P/6	-1.16	-2.90	.006	-.14	-1.27	.211	-2.85	.0
	P/9	-1.66	-2.36	.225	-.24	-1.55	.128	-2.27	.0
	0/3	-.64	-2.23	.031	-.03	-.78	.437	-1.20	.2
	0/6	-1.05	-2.65	.012	-.08	-.89	.378	-2.26	.0
	0/9	-1.54	-2.18	.036	-.18	-1.35	.183	-1.81	.0
	3/6	-.38	-3.79	.001	-.05	-.77	.445	-3.01	.0
	3/9	-.89	-2.35	.025	-.16	-1.34	.187	-2.07	.0
	6/9	-.54	-1.78	.085	-.11	-1.75	.086	-.25	.7
4	P/0	.03	.86	.394	.01	.59	.556	-1.16	.2
	P/3	-.37	-2.10	.044	.02	.25	.806	-2.15	.0
	P/6	-.81	-1.86	.075	.09	.96	.342	-1.56	.1
	P/9	-1.07	-1.61	.121	.19	1.34	.189	-1.15	.2
	0/3	-.45	-2.65	.012	.01	.25	.807	-2.48	.0
	0/6	-.93	-2.22	.035	.09	.98	.334	-1.78	.0
	0/9	-1.15	-1.69	.104	.19	1.35	.185	-1.27	.2
	3/6	-.44	-1.81	.081	.07	1.10	.277	-1.52	.1

2. Findings (relating to question a above) concerning intra-year changes in forecast error measures are also interesting. Absolute within-year changes were determined for each of the four forecast error measures (undeflated error, relative error, price deflated error, and squared error). Overall findings were similar for changes based on each of the four error measures. Table C2 shows representative results using the undeflated forecast error. Two items are of note: First, change measure are almost universally negative indicating smaller forecast errors as year-end approaches; this was expected. Second, the reduction in forecast error is consistently greater for failing firms in years 1-3 and significantly so in years 1 and 2. This is probably caused by the overall larger forecast errors for the failing firms; there is more room for improvement as new information arrives during the forecast year. These larger reductions in forecast error provide a potential basis for discriminating failing from healthy firms.

3. Findings (relating to question b above) concerning intra-year changes in dispersion measures are less clear cut. Absolute within year changes were determined for each of the seven dispersion measures. Findings for individual statistics were not completely consistent across the tests involving the seven measures, but the overall pattern of results was. Table C3 shows representative results using the standard deviation across forecasts as the measure of dispersion. The results are not in general significant, but an interesting pattern is evident nevertheless. Changes for failing firms are positive in year 1 and 2 and for most

TABLE C 2

INTRA-YEAR CHANGES IN FORECAST ERRORS
(Absolute difference in undeflated forecast Error)

		FAILING			HEALTHY			WILCOXON	
YEAR	MONTHS	MEAN	t	α	MEAN	t	α	z	α
1	P/0	-.31	-.89	.382	-.01	-.37	.710	-1.47	[illegible]
	P/3	-1.16	-2.03	.054	-.15	-2.57	.012	-2.74	[illegible]
	P/6	-2.15	-3.02	.007	-.34	-4.45	.000	-3.87	[illegible]
	P/9	-2.38	-3.38	.003	-.52	-5.05	.000	-3.69	[illegible]
	0/3	-.87	-2.10	.047	-.15	-2.31	.024	-.79	[illegible]
	0/6	-1.81	-3.28	.004	-.33	-4.06	.000	-2.94	[illegible]
	0/9	-2.01	3.65	.002	-.50	-4.78	.000	-3.37	[illegible]
	3/6	-.77	-3.68	.002	-.18	-4.70	.000	-2.67	[illegible]
	3/9	-1.05	-4.39	.000	-.35	-5.43	.000	-2.86	[illegible]
	6/9	-.18	-1.15	.264	-.17	-4.55	.000	-.32	[illegible]
2	P/0	-.12	-2.78	.009	-.05	-1.87	.067	-.91	[illegible]
	P/3	-.81	-3.14	.003	-.20	-3.52	.001	-3.67	[illegible]
	P/6	-1.21	-3.49	.002	-.25	-4.18	.000	-2.81	[illegible]
	P/9	-2.11	-2.70	.012	-.36	-4.01	.000	-2.63	[illegible]
	0/3	-.64	-2.76	.009	-.15	-2.64	.011	-3.11	[illegible]
	0/6	-1.01	-3.21	.003	-.19	-3.35	.002	-2.48	[illegible]
	0/9	-1.86	-2.56	.016	-.30	-3.64	.001	-2.54	[illegible]
	3/6	-.43	-4.35	.000	-.03	-.63	.528	-3.48	[illegible]
	3/9	-1.30	-2.46	.019	-.14	-1.75	.086	-2.75	[illegible]
	6/9	-.87	-1.90	.067	-.11	-1.64	.106	-1.98	[illegible]
3	P/0	-.10	-2.67	.012	-.04	-.94	.349	-1.61	[illegible]
	P/3	-.77	-2.16	.039	-.12	-2.15	.036	-2.32	[illegible]
	P/6	-1.16	-2.34	.027	-.29	-2.72	.009	-1.23	[illegible]
	P/9	-1.91	-2.28	.031	-.42	-2.97	.005	-1.73	[illegible]
	0/3	-.66	-1.85	.074	-.09	-2.74	.009	-1.21	[illegible]
	0/6	-1.08	-2.15	.040	-.29	-3.33	.002	-.81	[illegible]
	0/9	-1.84	-2.18	.038	-.43	-3.57	.002	-1.29	[illegible]
	3/6	-.34	-2.57	.016	-.18	-2.92	.005	-.29	[illegible]
	3/9	-1.06	-2.38	.024	-.32	-2.93	.005	-1.41	[illegible]
	6/9	-.71	-1.94	.063	-.14	-2.37	.022	-1.42	[illegible]
4	P/0	.01	.53	.599	.00	.00	1.000	.55	[illegible]
	P/3	-.28	-1.40	.173	-.13	-2.95	.005	.79	[illegible]
	P/6	-.84	-1.57	.129	-.31	-4.12	.000	.47	[illegible]
	P/9	-1.35	-1.84	.079	-.51	-4.86	.000	.79	[illegible]
	0/3	-.29	-1.47	.153	-.12	-3.31	.002	1.00	[illegible]
	0/6	-.85	-1.60	.124	-.28	-4.08	.000	.09	[illegible]
	0/9	-1.43	-1.87	.076	-.48	-4.59	.000	.61	[illegible]
	3/6	-.50	-1.61	.121	-.16	-2.70	.010	-.39	[illegible]
	3/9	-1.05	-1.98	.062	-.37	-3.97	.000	-.14	[illegible]
	6/9	-.49	-2.07	.051	-.19	-2.93	.006	.34	[illegible]

TABLE C 3

INTRA-YEAR CHANGES IN DISPERSION
(Absolute difference in Standard deviation
of forecasts across forecasters)

		FAILING			HEALTHY			WILCOXON	
YEAR	MONTHS	MEAN	t	α	MEAN	t	α	z	α
1	P/0	.73	1.14	.274	.02	.53	.600	.92	.354
	P/3	.84	1.06	.307	.01	.23	.819	-.28	.783
	P/6	.95	1.14	.273	-.00	-.05	.962	-1.05	.292
	P/9	1.05	1.25	.230	.02	.32	.753	-.16	.874
	0/3	.09	.99	.331	.00	.12	.902	-.44	.662
	0/6	.22	1.50	.147	-.03	-.90	.375	.31	.757
	0/9	.28	1.78	.087	-.00	-.06	.951	.82	.414
	3/6	.05	.62	.539	-.03	-1.04	.303	.15	.878
	3/9	.10	.95	.351	-.00	-.05	.962	1.12	.260
	6/9	.06	1.00	.325	.01	.47	.643	1.03	.299
2	P/0	.08	.55	.585	-.05	-2.12	.038	-.30	.761
	P/3	.13	.85	.401	-.08	-3.13	.002	.90	.368
	P/6	.25	1.62	.116	-.09	-2.66	.011	1.51	.129
	P/9	.29	1.42	.168	-.15	-3.17	.003	1.12	.261
	0/3	.25	1.00	.324	-.05	-1.88	.066	.34	.731
	0/6	.42	1.39	.175	-.04	-1.13	.265	1.78	.074
	0/9	.49	1.28	.211	-.10	-2.70	.009	1.81	.069
	3/6	.12	1.71	.097	.01	.25	.805	2.11	.034
	3/9	.13	1.25	.224	-.04	-1.17	.248	1.35	.175
	6/9	.00	.08	.933	-.06	-2.26	.028	1.21	.228
3	P/0	-.07	-2.32	.028	-.06	-1.40	.168	-1.96	.050
	P/3	-.07	-1.60	.122	-.07	-1.83	.074	-.49	.620
	P/6	-.02	-.27	.791	-.14	-1.92	.062	.26	.789
	P/9	.02	.31	.757	-.08	-2.04	.048	1.32	.186
	0/3	.01	.31	.758	-.01	-.61	.542	.95	.337
	0/6	.05	1.04	.311	-.08	-1.34	.189	1.16	.244
	0/9	.09	1.69	.107	-.05	-.91	.371	2.54	.011
	3/6	.02	.50	.621	-.06	-1.31	.197	.01	.993
	3/9	.06	1.05	.306	-.03	-.78	.442	1.48	.136
	6/9	.03	.91	.376	.04	.88	.385	.57	.567
4	P/0	-.04	-1.00	.329	-.01	-.81	.420	-.57	.566
	P/3	-.06	-.32	.749	-.03	-2.11	.043	.39	.693
	P/6	-.03	-.19	.855	-.11	-2.58	.014	.44	.656
	P/9	.01	.07	.948	-.07	-1.88	.070	.55	.579
	0/3	-.02	-.12	.902	-.01	-.36	.717	.76	.447
	0/6	-.03	-.16	.871	-.08	-2.13	.040	.23	.815
	0/9	.06	.44	.665	-.07	-1.66	.107	.79	.427
	3/6	.04	.48	.639	-.07	-1.80	.081	.03	.972
	3/9	-.11	-2.25	.042	-.01	-.28	.780	-1.94	.051
	6/9	-.03	-1.18	.259	.06	1.88	.070	-1.17	.241

measures in year 3, while changes for healthy firms tend to be negative. Thus in general dispersion across forecasters tends to increase during the year for failing firms and decrease for healthy firms. This is what one would expect if there is greater unpredictability for failing firms due to uncertainties associated with future bankruptcy.

4. Findings (relating to question c above) concerning intra-year changes in the frequency of forecast revisions were inconclusive. For intra-year changes in the three measures (up revisions, down revisions, total revisions) there were no systematic or consistent patterns of findings or significance in the tests. No tables are presented.

In summary, the tests conducted in this section, when using the absolute change measure, provide strong evidence of systematic groups differences in intra-year changes in the level of mean earnings forecasts, weaker evidence of group differences in intra-year changes in forecast errors, suggestive evidence of group differences in intra-year changes in forecast dispersion, and no evidence of group differences in intra-year changes in revision measures.

3.10 Inter-year Trends

The previous section reported on tests of whether forecast errors, forecast dispersion, forecast revisions, or the level of forecasts changed systematically within a given forecast year. This section addresses the question of whether properties of analysts forecasts (forecast error, bias, dispersion, level and revisions) differ from year to year, and more importantly whether year to year trends differ between failing and healthy firms.

Five year-to-year trend questions were investigated for the two groups of sample firms.

a. Does the magnitude of forecast errors change from year to year? One would hypothesize increasing forecast errors as bankruptcy approaches for failing firms due to increasing uncertainty.

b. Does the degree of bias in forecasts change from year to year? One would hypothesize a more negative bias (over-optimistic forecasts) as bankruptcy approaches for failing firms. This might be expected because of the negative bias for failing firms previously noted in section 3.3, and might be explained by the withholding of bad news information by failing firms.

c. Does the magnitude of forecast dispersion change from year to year? One would hypothesize increasing dispersion for failing firms as bankruptcy approaches due to greater risk and uncertainty.

d. Does the frequency of forecast revisions change from year to year? If information is disclosed one would expect increasing frequency of downward revisions for failing firms as bankruptcy approaches. There is no reason to expect any particular pattern of year to year changes for healthy firms.

e. Does the mean level of forecasted earnings change from year to year? One would expect forecasted earnings for both failing and healthy firms to reflect year to year trends in both economy wide and industry wide conditions, and hence expect some year to year trends that are similar between the two groups. If in addition, forecasted earnings for failing firms reflect the conditions that ultimately lead to bankruptcy, failing firms may show more of a year to year decline than healthy firms.

Ultimately for all five properties the question of interest is not whether there are systematic year to year trends for failing firms, but whether those trends are significantly different from year to year trends experienced by healthy firms.

Year to year trends in errors, bias, dispersion, level, and revision measures were determined by computing the year to year difference for particular measures

$$\text{Difference} = (\text{Measure}_{t} - \text{Measure}_{t-1})$$

And the log relative

$$\text{Log Relative} = (\text{Measure}_{t}/\text{Measure}_{t-1})$$

Error, bias, dispersion and level measures were previously determined at three-month intervals prior to year end, within each of four years prior to bankruptcy (i.e., at the last month of each quarter, months 0,3,6, and 9, as well as at the month just prior to announcement of actual earnings). The year to year trend variables were computed for each measure within each forecast year using the corresponding measure from the same month in the previous year. (e.g., differences between year 2, month 6 measures and year 3, month 6 measures). Changes in revision frequency were based on yearly average frequencies only.

The log relative measure is conceptually appealing because it treats increases and decreases symmetrically and reduces outlier effects. But the log relative measure is undefined when either argument is zero or negative, and a substantial number of dispersion or bias measures in the sample are zero or negative. The absolute difference measure fails to deflated for size effects but can accommodate more of the observations available in the sample. The findings resulting from tests performed on the two measures are consistent, but slightly stronger for the absolute difference measure. Since the objective is to identify measures that may distinguish failing and healthy firms, results using the difference measures are reported.

3.10.1 Tests of Inter-year Trends

Two types of tests were conducted. T-tests were conducted to answer the question of whether mean year to year trends, within each group, were significantly different from zero. Wilcoxon tests were conducted to answer the question of whether mean year to year trends were significantly different between groups. Test results for changes from year 2 to year 1 prior to bankruptcy are presented in Table T1 (forecast errors), Table T2 (forecast bias), Table T3 (dispersion), and Table T4 (revision frequency) and Table T5 (mean forecast level).

Table T1 is consistent with expectations. For three out of

TABLE T 1

YEAR 2 TO YEAR 1 CHANGE IN FORECAST ERROR

Error	Month	FAILING MEAN	FAILING t	FAILING α	HEALTHY MEAN	HEALTHY t	HEALTHY α	WILCOXON Z	WILCOXON α
	PA	3.58	2.46	.024	.07	.52	.609	2.23	.025
	0	3.65	2.54	.019	.05	.34	.731	2.92	.004
AEM	3	4.11	2.60	.017	.08	.53	.598	1.68	.093
	6	5.00	2.75	.014	.21	1.38	.174	1.74	.082
	9	4.22	2.57	.019	.30	1.96	.056	1.85	.064
	PA	.53	1.05	.309	.32	1.41	.163	.47	.639
	0	.02	.08	.940	.15	1.06	.292	1.45	.147
REM	3	-.24	-.48	.633	.74	1.52	.135	-.63	.528
	6	-.55	-.71	.485	1.08	1.56	.124	-.61	.543
	9	-1.37	-1.30	.211	1.37	1.55	.127	-.80	.424
	PA	.57	3.42	.003	.01	.62	.541	4.49	.000
	0	4.50	3.25	.004	.37	3.86	.000	4.97	.000
PEM	3	.63	2.43	.025	.01	1.25	.217	2.99	.003
	6	.89	2.07	.055	.02	1.89	.064	3.93	.000
	9	1.00	2.24	.037	.03	1.92	.061	4.32	.000
	PA	5.91	1.99	.062	.05	.12	.905	1.86	.063
	0	3.50	2.20	.040	-.04	-.13	.897	2.22	.027
SEM	3	3.07	1.30	.209	1.15	1.26	.213	1.34	.181
	6	2.59	.68	.504	2.48	1.45	.153	1.05	.296

the four error measures, the relative error measure excepted, there is evidence of significant increase in forecast error for the failing firm group. No such significant trend is apparent for the healthy firms. Furthermore, the Wilcoxon tests indicate that the increase in forecast errors are significantly greater for the failing firms as compared to the healthy firms. These findings are clearly strongest for the absolute error measures (AEM) and the price-deflated error measures (PEM).

Table T2 shows similar findings for the bias measures. Forecast errors are significantly more negative in year 1 than in the prior year 2 for failing firms. No similar increase in bias is apparent for the healthy group. Wilcoxon tests support the conclusion that the increase in negative forecast errors (over-estimation) is significantly different between the groups, particu-larly when differences are determined using the price-deflated forecast error (PFE) measures. As in section 3.6, tests were conducted using seven different dispersion measures, but as findings were roughly consistent across the seven measures, results for three representative dispersion measures are reported. Table T3 provides evidence of increasing dispersion of forecasts for failing firms. Means and t values are typically greater for the failing groups than the healthy group and several of the t values are significant. Thus there is some support for concluding that dispersion has increased for the failing firms. Wilcoxon Z scores are consistently positive and predominantly significant, thus the increased dispersion for the failing group is a trend not apparent

TABLE T 2

YEAR 2 TO YEAR 1 CHANGE IN BIAS

BIAS	Month	FAILING MEAN	FAILING t	FAILING α	HEALTHY MEAN	HEALTHY t	HEALTHY α	WILCOXON Z	WILCOXON α
	PA	-3.48	-2.12	.047	.13	.87	.389	-1.63	.103
	0	-3.99	-2.57	.018	.07	.53	.597	-1.89	.058
AFE	3	-4.37	-2.61	.017	.07	.45	.653	-1.46	.145
	6	-5.08	-2.78	.013	.07	.43	.667	-2.14	.032
	9	-4.36	-2.66	.016	.04	.19	.848	-2.37	.018
	PA	.37	.62	.543	1.72	.09	.926	-.39	.692
	0	-.01	-.02	.986	1.03	-1.28	.205	-.45	.651
RFE	3	.22	.43	.672	3.57	-1.51	.137	.89	.379
	6	.53	.67	.509	4.96	-1.50	.139	.40	.687
	9	1.33	1.25	.227	6.23	-1.50	.140	.71	.479
	PA	-.44	-2.23	.038	.07	.11	.914	-2.90	.004
	0	-4.25	-2.96	.008	.81	-.01	.992	-3.38	.001
PFE	3	-.64	-2.45	.024	.08	-.65	.518	-2.54	.011
	6	-.89	-2.08	.053	.09	-.62	.537	-4.09	.000
	9	-1.00	-2.26	.036	.13	-.85	.399	-4.44	.000

TABLE T 3

YEAR 2 TO YEAR 1 CHANGE IN DISPERSION

		FAILING			HEALTHY			WILCOXON	
Measure	Month	MEAN	t	α	MEAN	t	α	Z	α
STND DEV	PA	1.15	1.38	.189	.09	1.27	.211	1.42	.155
	0	.29	2.08	.047	.05	1.56	.125	1.92	.054
	3	.17	2.30	.031	-.02	-.43	.667	2.67	.008
	6	.21	4.56	.000	.02	.59	.558	4.08	.000
	9	.16	1.52	.144	-.04	-1.03	.308	2.45	.015
VAR/ PRICE	PA	3.57	1.33	.203	.02	1.41	.164	2.13	.033
	0	.56	1.22	.234	.01	1.74	.087	2.61	.009
	3	.08	1.94	.065	-.00	-.30	.769	3.37	.001
	6	.04	2.77	.011	.00	.20	.841	4.83	.000
	9	.05	1.71	.102	-.01	-1.85	.071	3.05	.002
RANGE	PA	1.80	1.45	.169	.35	1.59	.118	1.64	.099
	0	.59	2.02	.054	.18	2.37	.022	1.41	.159
	3	.31	2.44	.023	-.01	-.12	.905	2.30	.021
	6	.49	4.15	.000	.05	.61	.546	2.89	.004
	9	.35	1.94	.065	-.11	-1.06	.296	2.42	.015

for the healthy firms.

Table T4 provides no evidences for any year to year trends in the average frequency of forecast revisions for either group nor any difference between the groups.

Analogous tests were conducted for trends from year 3 to year 2 and year 4 to year 3 prior to bankruptcy. Throughout those tests there was no evidence of significant year to year change for either group and no evidence of any differences in year to year changes between the two groups. This was true for tests involving forecast errors, bias, dispersion and revisions. Tables for those tests are not reported here.

Table T5 shows year 2 to year 1 trends in the mean forecasted earnings level for both undeflated earnings per share and price deflated earnings per share. For failing firms measures are consistently negative indicating the declining level in forecasts which was previously suggested by the results reported in Table M1. There is evidence that the year to year declines are significantly greater for failing firms as compared to healthy firms. Results for year 3 to year 2 and year 4 to year 3 changes (not reported) also provided statistically significant evidence of greater year to year declines in forecast level for failing firms when compared to healthy firms for a least some of the monthly comparisons.

In short, the set of tests of year to year trends supports the following general conclusions: in the year just prior to bankruptcy there is a statistically significant increase in

TABLE T 4

YEAR 2 TO YEAR 1 CHANGE IN YEARLY AVERAGE REVISION FREQUENCY

	UP REVISION FREQUENCY CHANGE	DOWN REVISION FREQUENCY CHANGE	TOTAL REVISIONS FREQUENCY CHANGE
FAILING FIRMS			
Mean Change	-.001	.003	.002
t	-.50	.44	.27
α	.624	.664	.789
HEALTHY FIRMS			
Mean Change	-.000	.006	.006
t	-.03	1.40	1.52
α	.975	.168	.134
FAILING VS. HEALTHY			
Wilcoxon Z	-.879	-.627	-.724
α	.379	.531	.469

TABLE T 5

Year 2 to Year 1 change in Mean Forecast Level

		Failing			Healthy			Wilcoxon	
Measure	Month	Mean	t	α	Mean	t	α	Z	α
Undeflated earnings	PA	-1.20	-1.47	.153	-.07	-.49	.623	-3.48	.001
	0	-.08	-.15	.878	-.03	-.26	.796	-2.53	.011
	3	-.07	-.16	.876	-.01	-.05	.960	-2.73	.006
	6	-.11	-.23	.818	.01	.10	.917	-2.65	.008
	9	-.42	-1.66	.106	.03	.30	.768	-2.95	.003
	PA	-.82	-2.86	.008	-.02	-1.63	.108	-3.85	.000
Price deflated earnings	0	-.24	-2.31	.026	-.01	-1.15	.254	-2.51	.012
	3	-.12	-2.13	.039	.00	.50	.618	-1.42	.155
	6	-.12	-1.52	.136	.00	.02	.984	-1.42	.157
	9	-.06	-2.34	.025	.00	-.48	.630	-1.78	.075

forecast errors, increase in negative bias (over-estimation), increase in dispersion across forecasts and decrease in the level of forecasted earnings for firms facing impending bankruptcy.

3.11 Summary

The preceding sections in this chapter have provided findings on a series of tests designed determine if failing firms and healthy firms differ systematically with respect to various properties of earnings forecasts. Five specific properties were investigated:

a) Mean earnings forecast estimates

b) Forecast bias

c) Forecast errors

d) Dispersion in forecasts across forecasters

e) Frequency of revisions in forecast estimates

In addition, tests were conducted to determine if measures reflecting the five properties just listed changed significantly either a) within a given forecast year (intra year changes) or b) between successive forecast years (inter year trends); and whether such changes differed systematically between failing and healthy firm groups.

Results from the tests can be summarized as follows:

Mean forecasts: Failing firms had lower mean earnings forecasts than healthy firms in years 1-3. Failing firms had more negative intra-year changes in forecasts and more negative interyear trends in forecasts than healthy firms.

Bias: Forecasts were more over-optimistic (negative forecast error) for failing firms than for healthy firms in years 1-3. The magnitude of the negative bias for failing firms increased significantly from year 2 to year 1. The greater bias for failing firms however could not be attributed to failure to use the information available in past earnings and past errors; past earnings and error information was incorporated into forecasts for both healthy and failing firms.

Accuracy: Failing firms had consistently larger forecast errors for all four years of the study. Forecasts for failing firms were not significantly better than a naive forecast of no change in earnings. Within forecast years, accuracy increased more for failing firms than healthy firms as the earnings announcement date approached (years 1-3). Across years, there was a significant decrease in accuracy for failing firms from year 2 to year 1.

Dispersion: There was greater dispersion in forecasts for failing firms during year 1 and year 2 prior to bankruptcy. Within forecast years there was mild evidence that dispersion decreased less for failing firms as year end approached. Across years, dispersion increased for failing firms from year 2 to year 1.

Revisions: Results were least interesting for frequency of revision measures. There was some indication of greater frequency of downward earnings revisions for failing firms in years 1-4. There was no indication of any group differences in the pattern of either intra-year or inter-year changes in revision frequency.

While systematic groups differences exist to some degree for each of the five properties of analyst forecasts, the important issue is whether those differences are strong enough and stable enough to be exploited for predictive purposes. The following chapter addresses that question.

CHAPTER 4

TESTS OF DISCRIMINATION AND PREDICTION

4.1 Introduction

The previous chapter documented numerous systematic differences in the properties of analysts earnings forecasts between failing and healthy firms. This chapter reports on both univariate and multivariate approaches to using those properties to discriminate between the groups and predict bankruptcy.

4.2 Selecting measures to be used in prediction

Each documented property may be captured by various measures. Which specific measures may be most appropriate for prediction? For example, consider dispersion. The findings indicated greater dispersion across forecasts for failing firms, but what specific measure of dispersion should be selected and evaluated for its predictive ability? Seven different measures of dispersion were introduced (e.g. standard deviation, range, log variance, etc.). And within any given forecast year each measure can be computed for up to 24 months prior to the announcement of actual earnings. Thus to represent dispersion in any given forecast year, 7x24 or 168 possible measures exist to choose from.

Then, what if one wishes to create a measure of changes in dispersion within a year (intra-year changes)? Section 3.9 reported tests using 10 versions of measures reflecting intra-year changes. If those 10 measures are computed for each of the seven types of dispersion measures, 70 possible intra-year change in dispersion variables result. And the 10 versions of intra-year changes do not come close to exhausting the possible pairwise combinations of months from which intra-year changes could be calculated.

Further, what if one wishes to create a measure of changes in dispersion across two successive years (inter-year trends)? Should dispersion measures taken immediately prior to announcement be compared between the two years? Should measures in mid year, say month six, be used? There are as many possibilities as there are monthly forecasts within any forecast year.

Similar choices exist among the candidate measures for reflecting bias, accuracy, mean estimate, and revision frequency. When one considers the a) five properties of interest b) the various ways to measure each of those properties c) the different months when forecasts are available for any forecast year, and d) the various pairwise combinations of months that may be used to create measures of intra-year changes or inter-year trends, there are thousands of available measures that present themselves as candidates for consideration in constructing a predictive model. Clearly a narrowing down of the available measures to a manageable subset to be used in constructing predictive models is needed. The following measures were selected for inclusion in the subset:

Notation:

A = Actual earnings

F = Mean forecasted earnings

Numerical subscript = number of months prior to year end within a forecast year.

t subscript = year prior to bankruptcy.

Measures:

1. Mean Forecast (ME) = $F_{0,t}$
2. Intra-year change in forecast (MECHG) = $F_{0,t} - F_{6,t}$
3. Inter-year trend in forecast (METRND)=$F_{0,t} - F_{0,t-1}$
4. Error (ERR)= $F_{0,t} - A_t$
5. Intra-year error change (ERRCHG) = $ERR_{0,t} - ERR_{6,t}$
6. Inter-year error trend (ERRTRND) = $ERR_{0,t} - ERR_{0,t-1}$
7. Bias (BIAS) = $A_t - F_{0,t}$
8. Inter-year trend in Bias = $BIAS_{0,t} - BIAS_{0,t-1}$
9. Dispersion (SD) = Standard deviation of forecasts at month 0.
10. Intra-year change in Dispersion (SDCHG) = $SD_{0,t} - SD_{6,t}$
11. Inter-year trend in Dispersion (SDTRND) = $SD_{0,t} - SD_{0,t-1}$
12 Average frequency of downward revisions (DN) = percentage of forecasts revised downward during a month, averaged over months 0-11.
13. Intra-year change in downward revisions (DNCHG) = Average frequency of downward revisions over months 0-5 less average frequency of downward revisions over months 6-11.
14. Inter-year trend in downward revision (DNTRND) = $DN_t - DN_{t-1}$

Several considerations were taken into account in selecting these measures for further consideration:

a. Completeness: It was desired to have the subset include measures representing each of the primary properties (level, bias, error, dispersion, revisions), as well as both intra-year and inter-year changes in those properties.

b. Availability/Sample size: Generally more forecasts per firm become available as year end approaches and more firms have forecasts. Choosing measures based on month zero forecasts increased sample size.

c. Simplicity: One goal in developing a practical model is making it easy to understand and use. Undeflated measures for ME, BIAS, and ERR are simpler to use and don't require price information for deflation. (Tests were conducted and models constructed using price-deflated measures; findings and conclusions were the same). The standard deviation (SD) is more understandable then dispersion measures such as the log variance.

d. Wilcoxon tests results: High levels of significance on the previously reported wilcoxon tests for group differences suggest the potential for a measure to have some discriminating or predicting ability. Thus measures were selected that tended to be associated with highly significant wilcoxon tests.

e. Correlation: There tended to be high correlations for any given measure across different months within a forecast year. There also tended to be high correlation between different measures of a given primary property taken from the same month (e.g. standard deviation and range as measures of dispersion). Thus much data reduction, at little cost in terms of information, could be accomplished by selecting, even arbitrarily, one measure for each property taken at one month within a given forecast year. As indicated above the simplest measures and month zero were selected.

f. Factor analysis: Using factor analysis is an obvious data reduction technique, but it does not provide definitive guidance in selecting measures. When there are thousands of measures available, computer program limitations are exceeded. Furthermore, extracted factors are in part dependent on the package of measures imputed to a factor program. (For example having seven versions of dispersion measures and only three versions of bias measures may increase the probability of "dispersion" factor being identified by a factor program.) Factor analysis was conducted on several judgementally selected sets of variables. The final subset of measures listed above does adequately reflect the factors identifiable by factor analysis.

4.3 Univariate Analysis: Classification, Validation and Prediction

As a first step toward using forecast information to predict failure, a univariate analysis was conducted. The approach used follows Beaver (1966). The procedure is straight forward. Sample firms were rank-ordered independently on each of the measures of interest. The rank-ordered values for a given measure were visually observed. A cutoff or threshold value of the measure was selected to divide sample observations into failing and healthy firms. Cutoff values were selected that minimized the percentage of firms misclassified. Results using measures from year 1 are provided in table U1.

Five items relating to errors in classification are provided under the classification column in table U1: The type 1 error is the percentage of failing firms misclassified as healthy. The type 2 error is the percentage of healthy firms misclassified as failing. The average error is a weighted average of the type 1 and type 2 errors and thus represents the overall classification error rate. The percentage in the Naive column is provided as a benchmark for comparison. It represents the frequency of misclassification errors from the following naive classification rule: assign all firms to the group (failing or healthy) with the highest frequency in the sample. (This generally meant classifying all firms as healthy because data limitations were such that healthy firms outnumbered failing firms in the samples used to develop the cutoffs).

The final item in the table is a rough measure of the efficiency (EFF) of using the cutoff value on a variable to classify firms as compared to using the naive approach. It is calculated as the error rate from the naive approach minus the error rate from the cutoff value approach divided by the error rate from the naive approach, and thus measures the percentage of firms that were misclassified by the naive approach that were correctly classified by the cutoff value approach. EFF equals zero when the naive and cutoff approach have the same overall error rate. Higher positive values of EFF indicate increasing superiority of the cutoff approach over the naive rule, with a value of one indicating no errors in classification. Negative values indicate that the cutoff approach was less successful than the naive rule.

Classification results, however, typically overstate the value of an approach or model in discriminating between two groups since the classification rule (cutoff value) is applied to the same sample on which it is developed. Validation is required. Ideally validity should be assessed on a sample unrelated to that used to develop the classification rule, a hold out sample. Operationally this can be achieved by randomly dividing the sample into two subsamples, developing the model or cutoff value on each subsample, and using the cutoff from each subsample to classify the firms in the other subsample. Findings from using this approach are contained in the second set of results under the verification column.

Another approach to validation is to determine validity across time. The two remaining columns in the table, labeled prediction, show the results of applying the cutoffs (developed from year 1 measures) to the measures available for sample firms in year 2 and year 3.

Analogous univariate classification, validation and prediction results using year 2 and year 3 measures to determine cutoff values are provided in tables U2 and U3.

Several broad conclusions can be drawn from the tables.

a. The frequency of type 1 errors is consistently greater than the frequency of type 2 errors. This is unfortunate since the costs associated with type 1 errors are likely to be greater than

TABLE U1

UNIVARIATE DISCRIMINATION- YEAR 1 MEASURES

VARIABLE	CUTOFF	CLASSIFICATION ON YEAR 1 DATA					VERIFICATION ON YEAR 1 DATA					PREDICTION ON YEAR 2 DATA					PREDICTION ON YEAR 3 DATA				
		ERROR RATE %					ERROR RATE %					ERROR RATE %					ERROR RATE %				
		Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff
ERR	> .565	8	17	14	31	.55	31	15	20	31	.29	61	15	32	38	.16	57	17	33	39	.15
ERRTRND	> 2.88	56	0	14	26	.48	61	0	16	26	.38	92	2	35	36	.03	91	0	33	37	.11
ERRCHG	<- .83	35	14	20	26	.23	60	9	22	26	.15	58	7	23	38	.36	72	6	31	38	.18
BIAS	<-1.57	46	0	16	31	.48	35	8	16	31	.48	78	5	32	38	.16	82	6	35	39	.10
BTRND	<-2.46	50	0	13	26	.50	61	0	16	26	.38	85	7	35	36	.03	77	8	33	37	.11
SD	> 1.09	74	2	25	33	.24	74	13	33	33	.00	89	2	36	40	.10	100	2	39	38	-.03
SDTRND	> .31	63	6	26	35	.26	48	12	25	35	.29	69	7	31	38	.18	83	9	35	35	.00
SDCHG	> .175	60	10	27	33	.18	76	10	32	33	.03	74	11	35	39	.10	75	8	32	35	.09
ME	< .11	37	10	22	43	.49	37	10	22	43	.49	62	5	30	45	.33	74	2	34	45	.24
METRND	< -.46	52	17	31	39	.21	68	14	35	39	.10	60	18	36	42	.14	69	13	38	42	.14
MECHG	<-1.17	56	8	27	40	.33	70	8	33	40	.18	60	10	32	43	.21	78	4	36	44	.18
DN	> .080	85	8	39	40	.03	85	12	45	40	-.13	95	0	38	42	.10	83	2	36	42	.14
DNTRND	< -.05	76	0	30	40	.25	74	16	39	40	.03	92	4	42	43	.02	65	0	28	44	.36
DNCHG	> .041	60	21	37	40	.08	78	5	34	40	.15	61	20	38	42	.10	77	19	42	41	-.02

TABLE U2

UNIVARIATE DISCRIMINATION- YEAR 2 MEASURES

VARIABLE	CUTOFF	CLASSIFICATION ON YEAR 2 DATA: ERROR RATE % Type1	Type2	Ave	Naive	Eff	VERIFICATION ON YEAR 2 DATA: ERROR RATE % Type1	Type2	Ave	Naive	Eff	PREDICTION ON YEAR 1 DATA: ERROR RATE % Type1	Type2	Ave	Naive	Eff	PREDICTION ON YEAR 3 DATA: ERROR RATE % Type1	Type2	Ave	Naive	Eff
ERR	> .70	64	8	29	38	.24	56	17	31	38	.18	19	15	16	31	.48	60	13	31	39	.21
ERRTRND	> 1.26	81	7	32	36	.11	88	9	38	36	-.06	44	14	22	26	.15	68	5	28	37	.24
ERRCHG	< -.63	42	9	21	36	.42	42	9	21	36	.42	35	23	26	26	.00	72	13	36	38	.05
BIAS	< -.305	44	15	26	38	.32	58	17	31	38	.18	19	19	19	31	.39	37	17	25	39	.36
BTRND	<-1.43	73	11	33	36	.08	85	8	36	36	.00	44	10	19	26	.27	73	13	35	37	.05
SD	> .70	72	4	31	40	.23	72	4	31	40	.23	67	9	28	33	.18	83	4	38	38	.00
SDTRND	> .28	50	7	31	38	.18	96	5	40	38	-.05	63	6	26	35	.26	78	9	33	35	.06
SDCHG	> .40	71	2	29	39	.26	68	6	30	39	.23	80	6	31	33	.06	89	3	31	35	.11
ME	< .39	48	6	25	45	.44	50	8	27	45	.40	33	16	24	43	.44	67	9	35	45	.22
METRND	< -.13	40	22	30	42	.29	49	24	36	42	.14	37	32	34	39	.13	55	23	36	42	.14
MECHG	< -.61	42	14	26	43	.40	40	19	28	43	.35	39	25	30	40	.25	65	8	33	44	.25
DN	> .063	51	15	30	42	.29	46	30	37	42	.12	55	27	38	40	.05	67	19	39	42	.07
DNTRND	> .076	86	2	39	43	.09	86	4	40	43	.07	89	9	41	40	-.03	87	3	39	41	.05
DNCHG	> .040	62	20	38	42	.10	87	7	41	42	.02	60	23	38	40	.05	77	17	42	41	-.02

TABLE U3

UNIVARIATE DISCRIMINATION - YEAR 3 MEASURES

VARIABLE	CUTOFF	CLASSIFICATION ON YEAR 3 DATA					VERIFICATION ON YEAR 3 DATA					PREDICTION ON YEAR 1 DATA					PREDICTION ON YEAR 2 DATA				
		ERROR RATE %					ERROR RATE %					ERROR RATE %					ERROR RATE %				
		Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff	Type1	Type2	Ave	Naive	Eff
ERR	> .42	40	20	28	39	.28	60	20	36	39	.08	8	25	19	31	.39	50	27	35	38	.08
ERRTRND	> .745	64	11	30	37	.19	68	13	33	37	.11	44	25	30	26	-.15	73	11	33	36	.08
ERRCHG	< -.97	72	7	32	38	.16	72	7	32	38	.16	45	14	22	26	.15	65	7	28	36	.22
BIAS	< -.29	37	17	25	39	.36	66	13	34	39	.13	19	19	19	31	.39	44	17	27	38	.29
BTRND	<-1.10	64	11	30	37	.19	77	13	37	37	.00	44	12	20	26	.23	73	17	38	36	-.06
SD	> .30	67	11	32	38	.16	67	18	36	38	.05	41	21	28	33	.15	61	20	36	40	.10
SDTRND	> .095	50	12	25	35	.29	76	17	36	35	-.03	44	24	31	35	.11	58	24	37	39	.03
SDCHG	> .175	75	8	32	35	.09	95	8	39	35	-.11	68	10	28	33	.12	68	10	33	39	.15
ME	< .52	54	11	30	45	.33	54	13	31	45	.31	30	21	25	43	.42	44	13	27	45	.40
METRND	<-1.71	76	3	33	42	.21	76	5	35	42	.17	92	5	38	39	.00	86	6	39	42	.07
MECHG	< -.35	51	15	31	44	.30	64	15	34	44	.23	93	0	37	40	.08	89	0	38	43	.12
DN	> . 8	83	2	36	42	.14	57	36	44	42	-.05	85	10	40	40	.00	87	2	38	42	.10
DNTRND	> .08	87	3	39	41	.05	94	3	42	41	-.02	89	9	41	40	-.03	86	2	39	43	.10
DNCHG	NONE	100	0	41	41	.00	93	0	38	41	.07	100	0	40	40	.00	100	0	42	42	.00

those associated with type 2 errors. But such results are likely to occur if, as in reality, the frequency of healthy firms in a sample is greater than the frequency of failing firms. This tendency is apparent for all models in the study.

b. The classification accuracy for failing firms is generally not good. Type 1 errors are frequently above 50%.

c. EFF values are generally positive indicating that this univariate approach does have some ability to identify group membership. However, the superiority of using cutoff values as compared to the naive approach is frequently marginal.

d. Regardless of which year the cutoffs are developed on, there is a tendency for discrimination to improve as the cutoffs are applied to years closer to bankruptcy. This is not surprising since, if the measures have any ability to identify failing firms, the properties of failing firms should be more evident as bankruptcy approaches.

e. The cutoffs developed from year 3 measures tend to perform most poorly.

f. Across the classification, validation and prediction tests, and across the three years, the ME measure appears to perform best. Average error rates tend to be low and efficiency rates relatively high. The ability of a cutoff based on ME to outperform the naive approach tends to be the most consistent across the validation and prediction tests and across the years.

In summary, one would like to have a measure (or measures) that a) is valid in that it performs well on the verification tests and b) is consistent, in that it performs well in more than one year prior to bankruptcy, i.e., performs well in the prediction tests. Overall, measure ME performs best. Regardless of the year (1,2, or 3) in which the cutoff value is determined, use of ME allows for a discrimination of firms in the two years prior to bankruptcy which is markedly better than a naive rule. Efficiency indicators suggest that about 40% of firms misclassified by the naive rule can be correctly classified using a cutoff based on ME.

4.4 Multiple Discriminant Models

By far the most popular methodological approach to developing bankruptcy prediction models has been multiple discriminant analysis. (See Zavgren [1983] for a review.) Discriminant analysis is a statistical procedure that creates a linear combination of several discriminating variables that can then be used to assign a score to an individual observation or firm. The discriminant score is then used to classify firms into groups. Stepwise selection procedures can be used to identify the particular set of discriminating variables, from a larger set of potential discriminating variables, that in combination most effectively predicts group membership. Stepwise multiple discriminant analysis was used to develop some multivariate models for predicting failure.

Stepwise discriminant analysis is an approach involving both analytic and heuristic aspects. The model that results from applying the stepwise discriminant procedure depends on the set of variables the procedure is allowed to select from and the selection criteria. Both are choices of the model builder. Hence the individual model that results may not necessarily be optimal. However, by creating various models iteratively, models that apparently make the best use of the data to distinguish groups can be identified. In this study models were developed by allowing the stepwise procedure to select from among three different sets of measures.

1. All variables (i.e. the 14 listed in Section 4.2)

2. Primary Variables only (i.e. ERR, BIAS, SD, ME, DN, but no intra-year change or inter-year trend measures)

3. Factor Variables: A subset of all variables identified by factor analysis as representing the major dimensions of variability in the data.

Given a particular discriminant model, there are various indicators of its ability to discriminate groups. For example:

1. Classification error rates. The percentage of firms mis-classified by the model.

2. Cannonical Correlation. A measure of association between the linear discriminant function and the groups to be discriminated.

3. Significance of the model.

A "best" model developed from each of the three different sets of allowable variables mentioned above was developed from measures from years 1-3, resulting in nine models. However the three models constructed using year 3 measures were all one variable models. They are in essence analogous to the univariate approach and consequently are not reported here. Also the "factor" model developed from year 2 measures was the same as the "all variables" model and hence is redundant. The remaining five models are reported in tables M1 through M5.

As with the univariate approach each model was validated by splitting the sample in half, fitting the model to each subsample to generate coefficients and then classifying the firms in the opposite subsample. Also as with the univariate approach, models developed on measures from one year were applied to measures from the other two years to predict the group membership from observations in those years.

The all-variables model from year 1 data (Table M1) appears at first glance to be very good. No firms are misclassified when the model is fit to the full sample and the split-sample validation results still show only a 2% overall error rate. (Consequently efficiency is quite high, 1.00 and .91 respectively.) However, when the model is applied to measures from the other years prediction of group membership is no better (or worse) than the naive classification rule. There are apparently some relationships between the variables included in the model that are unique to the measures taken from year 1. This is a problem with multiple discriminant analysis. Correlation or inter-relationships between the measures that are unique to the data on which the model is developed, and which prove useful in distinguishing that set of observations, may not hold when the model is used to classify a

TABLE M 1

YEAR 1 MODEL - ALL VARIABLES

MEASURE	STANDARDIZED COEFFICIENT	UNSTND. COEFFICIENT
ERRCHG	.53	.50
SDTRND	1.02	5.14
BIAS	-4.18	-.94
BTRND	3.04	.65
ME	.55	.28
MECHG	-1.68	-1.68
DNTRND	-.56	-14.67
DNCHG	.53	12.37
CONSTANT	X	-2.49

Cannonical Correlation : .88

Likelihood Ratio (Wilks Lambda): .217

F Statistic : 17.99

Alpha Level : .0000

	YEAR 1 Classification	YEAR 1 Verification	YEAR 2 Prediction	YEAR 3 Prediction
FAILED FIRMS				
% Correct	100	92	25	0
% Incorrect	0	8	75	100
HEALTHY FIRMS				
% Correct	100	100	87.5	96
% Incorrect	0	0	12.5	4
ALL FIRMS				
% Correct	100	98	67	67
% Naive Correct	73	73	67	69
EFFICIENCY	1.00	.91	.00	-.06

TABLE M 2

YEAR 1 MODEL - PRIMARY VARIABLES

MEASURE	STANDARDIZED COEFFICIENT	UNSTND. COEFFICIENT
ERR	.48	.12
SD	.71	1.25
BIAS	-.28	-.07
ME	-.36	-.18
CONSTANT	X	-1.84

Cannonical Correlation : .68

Likelihood Ratio (Wilks Lambda): .533

F Statistic : 12.47

Alpha Level : .0000

	YEAR 1		YEAR 2	YEAR 3
	Classification	Verification	Prediction	Prediction
FAILED FIRMS				
% Correct	57	64	19	0
% Incorrect	43	36	81	100
HEALTHY FIRMS				
% Correct	100	85	100	98
% Incorrect	0	15	0	2
ALL FIRMS				
% Correct	90	81	83	63
% Naive Correct	77	77	66	65
EFFICIENCY	.56	.17	.50	-.06

TABLE M 3
YEAR 1 MODEL- VARIABLES FROM FACTOR ANALYSIS

MEASURE	STANDARDIZED COEFFICIENT	UNSTND. COEFFICIENT
SD	2.54	4.25
SDCHG	-1.95	-3.57
BIAS	-3.14	-.70
ERRTRND	-2.12	-.46
DNTRND	-.12	-3.05
CONSTANT	X	-2.58

Cannonical Correlation	.81
Likelihood Ratio (Wilks Lambda)	.343
F Statistic	16.09
Alpha Level	.0000

	YEAR 1		YEAR 2	YEAR 3
	Classification	Verification	Prediction	Prediction
FAILED FIRMS				
% Correct	92	83	20	7
% Incorrect	8	17	80	93
HEALTHY FIRMS				
% Correct	97	94	86	87
% Incorrect	3	6	14	13
ALL FIRMS				
% Correct	96	92	63	60
% Naive Correct	75	75	65	67
EFFICIENCY	.84	.68	-.06	-.21

different set. The model is "overfit" to the data. Use of factor analysis to reduce correlation between the independent variables might be expected diminish the problem, but the results for the factor model (Table M3) are also poor. When applied to the off-years (prediction on years 2 and 3) the factor model is poorer then the naive prediction rule.

Overall error rates increase and efficiency relative to the naive rule decreases for all models in the off-years (except for model M2 when applied to year 2). In fact all models have a higher error rate than the naive rule when applied to year 3 data.

Both models fitted on year 2 data (Tables M4 and M5) are no better, and perhaps worse, than the univariate approach using variable ME. If we discount them for that reason, we are left with the three models developed on year one data, and their ability to discriminate firms well is apparently limited to the year immediately prior to bankruptcy. Yet even that conclusion may be questioned. The classification results are good (particularly for M1 and M3) but the verification results are poorer. And there is a bias toward achieving high verification results. Unlike the univariate approach the discriminant approach uses the information in the data itself to <u>select</u> the best discriminating variables. The split-sample validation procedure forms new linear combinations of the variables in each subsample, which is then used to classify firms in the opposite subsample. The variable coefficients of the models developed in each subsample are of course unaffected by the information contained in the opposite subsample but the specific

TABLE M4
YEAR 2 MODEL - ALL VARIABLES

MEASURE	STANDARDIZED COEFFICIENT	UNSTND. COEFFICIENT
SDCHG	-1.09	-.92
MECHG	-2.03	-1.25
DNTRND	-.44	-10.03
CONSTANT	X	-.93

Cannonical Correlation : .42

Likelihood Ratio (Wilks Lambda): .83

F Statistic : 4.38

Alpha Level : .007

	YEAR 2 Classification	YEAR 2 Verification	YEAR 1 Prediction	YEAR 3 Predictio
FAILED FIRMS				
% Correct	48	41	59	12
% Incorrect	52	59	41	88
HEALTHY FIRMS				
% Correct	92.5	95	82	82.5
% Incorrect	7.5	5	18	12.5
ALL FIRMS				
% Correct	75	73	74	61
% Naive Correct	60	60	67	65
EFFICIENCY	.38	.33	.21	-.11

TABLE M5
YEAR 2 MODEL-PRIMARY VARIABLES

MEASURE	STANDARDIZED COEFFICIENT	UNSTND. COEFFICIENT
SD	-.07	-.59
ME	-1.12	-.06
CONSTANT	X	.24

Cannonical Correlation : .40

Likelihood Ratio (Wilks Lambda): .837

F Statistic : 8.54

Alpha Level : .0004

	YEAR 2 Classification	YEAR 2 Verification	YEAR 1 Prediction	YEAR 3 Prediction
FAILED FIRMS				
% Correct	42	56	63	11
% Incorrect	58	44	37	89
HEALTHY FIRMS				
% Correct	93	85	86	91
% Incorrect	7	15	14	9
ALL FIRMS				
% Correct	73	74	78	61
% Naive Correct	60	60	67	63
EFFICIENCY	.33	.35	.33	-.05

variables to be included in the model are not. The fact that the full sample is used originally to select the variables implies that the variables should have some discriminating ability in both subsamples even if the specific way in which the variables are joined in a linear combinations is subsample specific.

Another problem is also evident. The results are reported in the tables such that higher discriminant scores are associated with failing firms. But coefficient signs for individual variables are not always as one would expect for some of the measures. For example in model M1, variable ME has a positive sign, while the univariate results, and common sense, indicate that failing firms should have lower forecasted earnings. This suggests that ME is in the model because of some unique interrelationship among the variables that may be peculiar to the year 1 data.

In summary there is reason to question the value of the discriminant approach. Models developed on year 3 data were single variable models and hence no better then the univariate approach. Models developed on year 2 data did not perform better than individual measures reported from the univariate approach. And models developed from year 1 data can be criticized. While the classification error rates are low, they don't perform well when applied to data from other years and their validity can be questioned.

4.5 Conclusions

The purpose of this study was to determine if various measures developed from financial analysts forecasts of earnings for firms could be exploited in predicting future bankruptcy. The study consisted of two major parts.

In the first part (Chapter 3) the properties of analysts forecasts were discussed and measures were developed to reflect these properties. Five properties were investigated: forecast level, forecast dispersion, forecast error, forecast bias and forecast revisions. Various tests were conducted to determine if there were systematic differences in the properties for failing firms as compared to healthy firms in years prior to the bankruptcy of the failing firms. Several statistically significant differences were apparent. Failing firms tended to be associated with lower forecasted earnings, higher dispersion in earnings forecasts across forecasters, greater error in forecasts, over-optimistic forecasts, and perhaps greater frequency of downward revisions in forecast estimates. Some differences between failing and healthy firms in both intra-year and inter-year changes in the measures were also noted.

In the second part of the study (Chapter 4) measures reflecting aspects of the five properties were used to discriminate failing from healthy firms. Both univariate and multivariate approaches were attempted. Measures and linear combinations of measures were able to out-predict a naive model in distinguishing between groups. However, overall results were not encouraging. It was possible to develop multivariate models that were highly successful in classifying firms but these models were less successful in predicting group membership when applied to data taken from different years prior to bankruptcy, and were of questionable validity. Individual measures were more consistent across time in identifying failing firms from healthy firms. Perhaps the single best approach identified for using measures taken from analyst forecasts for predicting failure was to look at the mean forecasted value of future earnings. Simply put failing firms were predicted to have lower earnings. An approach where firms with forecasted earnings below a particular cutoff point were classified as facing bankruptcy was able to correctly classify from 22% to 40% (depending on the year prior to bankruptcy) of firms that were misclassified using the naive rule (which classified all firms as healthy). While the results indicate some ability of forecast data to assist in the prediction of failure, the ability to exploit analyst forecast data in predicting failure does not appear to be great. The fact that there is an association between lower forecasted future earnings and future bankruptcy is not surprising and does not appear to provide any novel insight that may be exploited for predicting bankruptcy.

Two directions for future research could change these conclusions. First, if the multiple discriminant models that had high classification success in this study were validated on other sets of data and found not to be unique to the observations in this study, then the conclusion that analyst forecasts do contain information particularly suitable to predicting failure would follow. Second, perhaps other approaches to model building, such as the construction of a "failure index" (e.g. Moses and Liao [1986]) may better exploit the information available in forecast data for predicting failure.

References

. Abdel-Khalik, A. and B. Ajinkya, "Returns to Informational Advantages: The Case of Analysts' Forecast Revisions," Accounting Review (October 1982), pp. 661-80.

. Abdel-Khalik, A. and J. Espejo, "Expectations Data and the Predictive Value of Interm Reporting," Journal of Accounting Research (Spring 1978), pp. 1-13.

. Ajinkya. B. and M. Gift, "Corporate Managers' Earnings Forecasts and Symmetrical Adjustments of Market Expectations," Journal of Accounting Research (Autumn 1984), pp. 425-444.

. Altman, E., Corporate Financial Distress, New York: John Wiley, 1983.

. Bao, D., M. Lewis, W. Lin, J. Manegold, "Applications of Time-series Analysis in Accounting: A Review," Journal of Forecasting, (Oct/Dec. 1983), pp 405-424.

. Barefield, R. and E. Comiskey, "The Accuracy of Analysts' Forecasts of Earnings Per Share," Journal of Business Research (July 1975), pp. 241-52.

. Beaver, W. "Financial Ratios as Predictors of Failures," in Empirical Research in Accounting, selected studies, 1966, Supplement to Journal of Accounting Research.

. Brown, L. and M. Rozeff, "Adaptive Expectations, Time-Series Models and Analyst Forecast Revisions" Journal of Accounting Research (Autumn 1979), pp. 341-51.

. Brown, P., G. Foster and E. Noreen, Security Analyst Multi-year Earnings Forecasts and the Capital Market, Sarasota, Florida: American Accounting Association, 1985.

. Castanias, R. and P. Griffin, "The Effects of Foreign Currency Translation Accounting on Security Analysts Forecasts," Working paper, U.C. Davis, November 30, 1984.

. Chang, L. and K. Most, "Financial Statements and Investment Decisions," (manuscript, Florida International University, Miami, 1980).

. Collins, W. and W. Hopwood, "A Multivariate Analysis of Annual Earnings Forecasts Generated from Quarterly Forecasts of Financial Analysts and Univariate Time Series Models," Journal of Accounting Research, (Autumn, 1980), pp. 340-406.

. Crichfield, T., T. Dyckman, and J. Lakonishok, "An Evaluation of Security Analysts' Forecasts," Accounting Review (July, 1978), pp. 651-68.

. Cukierman, A. and D. Givoly, "Heterogeneous Earnings Expectations and Earnings Uncertainty-Theory and Evidence," working paper, Tel Aviv University (1982).

. Elton, E. and M. Gultekin. "Expectations and Share Prices," Management Science (September 1981). pp. 975-87.

. Elton, E., M. Gruber, and M. Gultekin, "Professional Expectations: Accuracy and Diagnosis of Errors," Journal of Financial And Quantitative Analysis, (December 1984), pp. 351-363.

. Foster, G., Financial Statement Analysis Second Edition, Englewood Cliffs, New Jersey, Prentice-Hall, 1986.

. Fried, D. and D. Givoly, "Financial Analysts' Forecasts of Earnings: A Better Surrogate for Earnings Expectations," Journal of Accounting and Economics, (October 1982), pp. 85-107.

. Givoly, D., "The Formation of Earnings Expectations," The Accounting Review, (July, 1985), pp. 372-386.

. Givoly, D. and J. Lakonishok. "The Information Content of Financial Analysts' Forecasts of Earnings," Journal of Accounting and Economics (Winter 1979), pp. 165-85.

. Givoly, D. and J. Lakonishok, "Financial Analysts' Forecasts of Earnings Their Value to Investors," Journal of Banking and Finance, (September, 1980), pp.221-33.

. Givoly, D. and J. Lakonishok, "Divergence of Earnings Expectations: The Effect on Market Response to Earnings Signals," working paper 768/83, Faculty of Management, Tel Aviv University (February, 1983).

. Givoly, D. and J. Lakonishok, "Properties of Analysts Forecasts of Earnings: A Review and Analysis of the Research, " Journal of Accounting Literature (Spring, 1984), pp. 117-152.

. Gonedes, N. "Capital Market Equilibrium and Annual Accounting Numbers: Empirical Evidence," Journal of Accounting Research (Spring, 1974), pp. 26-62.

. Malkiel, B. "The Valuation of Public Utility Equities," The Bell Journal of Economics and Management, (Spring 1970), pp. 143-60.

. Malkiel, B. "Risk and Return: A New Look," working paper, Lynch Jones, and Ryan. New York (April 1981).

. Malkiel, B. and J. Cragg, "Expectations and the Structure of Share Prices," American Economic Review (1970), pp. 601-17.

. Malkiel, B. and J. Cragg, "Expectations and the Valuation of Shares," working paper 471 (April 1980), National Bureau of Economic Research.

. Mincer, J. and V. Zarnowitz, "The Evaluation of Economic Forecasts," ed. by J. Mincer, Economic Forecasts and Expectations (New York, NY: National Bureau of Economic Research 1969).

. Muth, J. "Rational Expectations and Theory of Price Movement," Econometrica (July 1961), pp.315-35.

. Neiderhoffer, V. and P. Regan, "Earnings Changes, Analysts Forecasts and Stock Prices" Financial Analysts Journal (May-June 1972), pp. 65-71.

. Penman, S. "An Empirical Investigation of Voluntary Disclosure of Corporate Earnings Forecasts," Journal of Accounting Research (Spring, 1980), pp. 132-160.

. Theil, H. Applied Economic Forecasting (Amsterdam, Holland: North-Holland Publishing Company, 1966).

. Zavgren, C., "The Prediction of Corporate Failure: The State of the Art," Journal of Accounting Literature, (Spring, 1983), pp. 1-38.

DISTRIBUTION LIST

	No. of Copies
Dudley Knox Library CODE 0142 Naval Postgraduate School Monterey, CA 93943	2
Library CODE 54 Naval Postgraduate School Monterey, CA 93943	2
Research Administration CODE 012 Naval Postgraduate School Monterey, CA 93943	1
Professor O. Douglas Moses CODE 54Mo Naval Postgraduate School Monterey, CA 93943	10
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Chief of Naval Research 800 N. Quincy Street Arlington, VA 22217-5000	2