Research Papers
in Management Studies

UNIVERSITY OF CAMBRIDGE
Judge Institute of Management
cambridge's business school

PROSPECTIVE EARNINGS PER SHARE

R G Bates, M A H Dempster,
H G Go & Y S Yong

WP 06/2004

The Judge Institute of Management
University of Cambridge
Trumpington Street
Cambridge CB2 1AG, UK
www.jims.cam.ac.uk

PROSPECTIVE EARNINGS PER SHARE

R G Bates, M A H Dempster, H G Go & Y S Yong

WP 06/2004
To appear as a Merrill Lynch Report

R G Bates
Centre for Financial Research
Judge Institute of Management
University of Cambridge
Tel: +44 (0) 1223 339651
Fax: +44 (0) 1223 339652
Email: rgb2@cam.ac.uk

Professor M A H Dempster
Centre for Financial Research
Judge Institute of Management
University of Cambridge
Tel: +44 (0) 1223 339651
Fax: +44 (0) 1223 339652
Email: m.dempster@jims.cam.ac.uk

H G Go
Centre for Financial Research
Judge Institute of Management
University of Cambridge
Tel: +44 (0) 1223 339651
Fax: +44 (0) 1223 339652
Email: hgg21@cam.ac.uk

Y S Yong
Centre for Financial Research
Judge Institute of Management
University of Cambridge
Tel: +44 (0) 1223 339651
Fax: +44 (0) 1223 339652
Email: ysy20@cam.ac.uk

Please address enquiries about the series to:

Research Support Manager
Judge Institute of Management
Trumpington Street
Cambridge CB2 1AG, UK
Tel: 01223 760546 Fax: 01223 339701
E-mail: research-support@jims.cam.ac.uk

Prospective Earnings Per Share

R.G. Bates, M.A.H. Dempster, H.G. Go, Y.S. Yong

email: {rgb2,mahd2,hgg21,ysy20}@cam.ac.uk

Abstract

This report considers the relation between pro-forma and forecast consensus *earnings per share*(EPS) figures in terms of six measures identified, qualitatively, as good indicators for quality of earnings. These are, *return on capital employed* (RoCE), *productive asset reinvestment ratio* (PARR), *cash realization* (CR), *tax rate* (TR), *Standard and Poors* (S&P) *equity rating* and *S&P debt rating* as identified by Merrill Lynch. The choices are thought to capture aspects of and differentiate between long term strategies and short term, non-sustainable earnings through financial engineering.

Analyses are run on ten years of data from 1992 to 2001 for 131 S&P500 companies also provided by Merrill Lynch, using consensus EPS forecasts one year ahead. This sample is smaller than 500 because the index is subject to change, with companies being added and removed based on their market share or other factors. S&P debt ratings were not available over the entire sample length for all companies thus cutting the two ratings from analyses. However, correlations between the two S&P ratings and other indicator measures are found to be high so their removal is not a significant problem. The ten years are split into three phases based on market sentiment:

1992-1995 Bull market, emphasis on sustainable growth

1996-1999 Bull market, emphasis on high earnings per share (bubble)

2000-2001 Bear market, refocussing on sustainability.

Given the radical change in conditions after 2000, the sample is enlarged for 2000 to 2001 and investigated separately. This increases sample size to 366, but does not change results significantly or give any further insights.

The ability of the four indicator measures to connect EPS predictions to released figures appears weak in analyses. Adding lagged information and market proxies improves the situation, but unfortunately not sufficiently for linear regression to be used confidently for out-of-sample prediction. Given a sample of ten observations in time, non-linear regressions were not carried out. The indicator measures are not thought to be effective indicators of companies' future performance.

Contents

1 Introduction

The 1990s witnessed a surge in service, internet and technology stocks, resulting in a record high of the stock market. Many individuals, including research analysts, believed that this signified the start of a 'new economy', also known as the 'virtual economy' [1]. Transformation of business investments and strategies were anticipated following rapid advances in the fields of technology and computing. Many new, high-technology companies emerged and were thought to possess immense growth potential, coupled with easy access to global markets, low barriers to entry and low supply chain costs. Wall Street analysts demonstrated creativity by introducing new valuation methods and abandoning the traditional methods, such as price-earnings ratios. For example, former Merrill Lynch analyst Henry Blodget introduced an unconventional valuation concept whereby Yahoo stocks were highly recommended after the company received large tax breaks, though these resulted from employees exercising their stock options [6]. The market shifted its emphasis to metrics relating to sales or 'hyper-growth' indicators such as price-sale revenue ratio and customer growth rate. Believing that these stocks had large, uninterrupted growth potential, the market failed to recognize that inflated prices were not matched by similar rises in the company earnings. Spurred by market sentiments and speculative activities, fundamentals were outstripped by psychology. The bubble burst, leaving behind a trail of bankrupt companies and unemployment. Investors caught in the euphoria suffered severe losses, and have since learned that the stock price does not necessarily reflect a company's true potential. This has led to greater emphasis on using corporate fundamentals in assessing a company's performance. With increasing reports of fraudulent earnings management, investors have begun to question the accounting techniques being employed, together with critical issues such as corporate governance and compensation incentives. Investors now demand greater transparency and have started to focus on earnings quality, a proxy for the longer term prospects of a company, where sustainability is a top priority. Furthermore, measures of earnings quality may highlight changes in a company even before the effects are reflected in the income statements.

Six key financial ratios, *Return on Capital Employed* (RoCE), *Cash Realization* (CR), *Productive Asset Reinvestment Ratio* (PARR), *Tax Rate* (TR), *Standard & Poor's* (S&P) *Equity Score*[1] and *S&P Debt Score* have been identified by Merrill Lynch as good indicators of qual-

[1] Also known as Earnings and Dividend Quality Ranking

ity of earnings. Patel and Santicchia [12] have shown that a high quality portfolio composed of common stocks with high S&P Equity Scores gives the highest risk-adjusted returns, outperforming the S&P 500 index. In addition, they found that firms with high earnings quality scores have higher earnings predictability, giving rise to more accurate earnings forecasts by analysts. We aim to find which of these measures have the best predictive power in determining whether or not a company will meet its forecast *earnings per share* (EPS). An optimum weighting of these six ratios is to be found, forming a *quality of earnings index* (QEI) where a high quality index indicates increased probability of meeting forecast EPS.

Section 2 introduces earnings per share and explains the six indicators. This is followed by an overview of possible statistical techniques to relate the measures and EPS in Section 3. These methods are applied and discussed in the following chapter.

2 Quality of Earnings

2.1 Motivation

Pro-forma earnings per share and, by implication, *Consensus EPS* (CEPS) forecasts figures are important for various stock valuation models. Investors have used CEPS predictions over the last decade, where a company's failure to meet these estimates could lead to its stock being decimated [14]. Hence, managerial objectives and incentives concentrated on these figures creating a situation where healthy but opaque earnings could hide warning signs, e.g., Parmalat, Enron, WorldCom and Xerox. Better indicators for the quality of earnings numbers published are essential to improving this situation.

The study of quality of earnings originated in fundamental analysis, which was developed in the 1930s to identify over- and under-valued securities by looking into a firm's financial statements to create a better understanding of its operations. If the market value is above or below the intrinsic value, then the security is over- or under-valued respectively. It was not until the 1970s that more comprehensive studies into quality of earnings were undertaken. O'Glove [10] looked into earnings quality as a determinant of a company's future prospects to provide professional commentaries and investment recommendations in his reports. These successfully identified potentially distressed companies as well. Bernstein and Siegel [2] introduced formally the concept of earnings quality in 1979, carrying out a survey [3] to determine the factors affecting earnings quality by interviewing researchers and accounting professionals. Amongst the issues addressed was the increase in earnings manipulation by management, which contributed to misleading earnings figures and a reduction in their use as a proxy for future earnings.

Awareness of earnings quality has risen substantially over the years. In March 2002, the Merrill Lynch Global Strategy team found that 43% of fund managers surveyed considered US equities to have the best earnings quality while 9% thought otherwise. This worsened dramatically over a few months, with 34% believing American equities to have the worst quality of earnings by July [9].

2.2 Formal Definition

To date, there is still no definitive measure of earnings quality due to its broad coverage and it being a relative measure. Several examples of attempts to define this concept are listed below:

- Bernstein and Siegel have defined quality evaluation in the earnings figure, as *"one of comparative integrity, reliability and predictability."* This is in line with the *Securities and Exchange Commission* (SEC)'s view. Accounting Series Release (ASR) No. 159 states that *"the purpose of the explanation of the Summary of Earnings is to enable investors to assess the source and probability of recurrence of net income, and thus of earnings quality"* and Release No 33-5427 that *"if a company's accounting principles are at variance with prevailing accounting practices within the industry, the dollar effect on earnings should be disclosed for there to be a proper assessment of the quality of the registrant"*. Therefore, quality of earnings investigation requires systematic examination of the impact of different accounting techniques on earnings figures and detection of managerial manipulation.

- Brown [4] points out that financial analysts are cautious about using accounting earnings for valuation, and the *"key is to separate economic value-added from 'cosmetic' earnings"*, *"earnings quality may be lacking for obvious reasons – reported earnings include a large, one-shot revenue item that has no bearing on future earnings potential,* *for example, or substantial research and development write-offs that unduly depress current earnings. But a much more subtle factor particularly complicates the assessment: Management teams often can manage the level or trend of reported earnings".*

- Merrill Lynch defines high-quality earnings as those earned by achieving superior returns on total capital [9], close to being realized in cash and repeatable. High quality earnings should not depend on transients, e.g., reported tax rates should not bear addition of risk as a result of high financial leverage and dividend obligations.

The common theme in these definitions is to determine which of the components of earnings figures result in cash generation. The use of consistent and precise measurements of non-transactional items, such as depreciation and provision for doubtful debts, is crucial. Embedded transient items should be identifiable and corrected to reflect future sustainability. Table 1 identifies several accounting treatments and the maxims of reporting that have clouded the true measure of earnings.

It is difficult to define a measure of earnings quality since it depends on a variety of factors such as full disclosure, management strategy and companies' operational actions. To achieve high quality earnings, prompt disclosure of 'bad' news is needed. Furthermore, an understanding of the management strategy is crucial. For example, an exit from a profitable business might signal a possible decrease in future income. Economic slowdowns too might negatively affect a company's earnings, especially in the case of an industry-specific retardation. Therefore, adherence to GAAP is a necessary but insufficient condition for companies to report high quality earnings.

2.3 Indicator Measures

Even though it is difficult to provide an accurate measure of earnings quality, Professor David Hawkins of Harvard Business School has suggested six accounting metrics encompassing the different aspects of good earnings quality. The following sections illustrate these metrics in detail.

Cash	<———————>	Noncash
Based on Fixed Amount	<———————>	Based on Estimate
Recurring	<———————>	Nonrecurring

High Quality		Low Quality	
Characteristic	**Example**	**Characteristic**	**Example**
Cash earnings	Recurring sales for which cash has been received	Non cash earnings	Goods sold in exchange for stock of another company
Result from consistent application of accounting principles	Consistent application of LIFO method of inventory valuation	Result from discretionary changes to existing accounting principles that result in earnings but no cash increase	Switch to the FIFO method of inventory valuation
Result from consistent application of estimation principles and methods	Consistent application of pension expense calculation assumptions	Result from changes in estimation principles and methods that increase earnings but not reliability of the estimate	Change in the assumed rate of return on pension assets
Result from estimates for which the range of possible balances is relatively small	Changes in the accounts receivable reserve that has a \$100,000 range for the possible balance	Result from estimates for which the range of possible balances is relatively large	Changes in the environmental contingencies reserve that has a \$1 million range for the possible balance
Based on transactions that are recurring	Rental income	Based on transactions that may recur but cannot be predicted or are unusual or nonrecurring	Settlement of litigation
Based on assets that are probable of recovery or liabilities that are fixed and certain	Interest expense	Based on assets with uncertainty of recovery or liabilities subject to change	Expenses associated with warranty reserves
Result from arms-length, commonly executed transactions with independent parties	Sales to an independent customer	Result from sales to related parties or uniquely structured transactions	Sales to an off-balance sheet, special purpose entity
Result from assets or liabilities recorded at cost	Interest on investments	Result from assets or liabilities recorded at fair value	Mark-to-market, held for sale investments
Reflect proposed external/internal audit adjustments as presented in the financial statements	Repairs and maintenance expense that reflects the recording of external auditor adjustments for expenses that were originally capitalized	Do not reflect proposed external/ internal audit adjustments as presented in the financial statements	Repairs and maintenance expense that does not reflect the recording of external auditor adjustments for expenses that were erroneously capitalized

Table 1: Comparison of different accounting methods and their impact on quality of earnings (Source: Deloitte & Touche [5])

2.3.1 Return on Capital Employed (RoCE)

Return on Capital Employed (RoCE) measures the rate at which shareholders' funds generate income from operating activities, excluding income and expenses from financing activities. It also measures a firm's ability to pay dividends and service its debt obligations. RoCE can be decomposed into its two drivers: operating profit margin and asset turnover via the *Du Pont model*, where the former is a profitability indicator and the latter is an efficiency measure. Analysts usually penalize a decrease in profit margin, especially a decline in gross margin where the ratio is affected by factors such as competition intensity and operating expenses. Therefore, the shift in this indicator measure is related to the long-term performance of a company. A higher RoCE shows efficient usage of funds and greater profitability, thus indicating better earnings quality.

$$\begin{aligned}
\text{RoCE} &= \frac{\text{Pre-tax Income} + \text{Interest Expense}}{\text{Total Assets} - \text{Current Liabilities}}\ \% \\
&= \frac{\text{Operating Income}}{\text{Common Shareholder Equity}}\% \\
&= \frac{\text{Operating Income}}{\text{Sales}} \times \frac{\text{Sales}}{\text{Common Shareholder Equity}}\ \% \\
&= \text{Profit Margin} \times \text{Asset Turnover}\ \% \qquad (1)
\end{aligned}$$

2.3.2 Cash Realization (CR)

Cash realization (CR) measures a company's cash generation capability by determining the proportion of cash inflow generation in net income. This measure aims to identify earnings manipulation by management through non-transactional income such as mark-to-market accounting valuations and aggressive accounting techniques. In addition, an increase in account receivables, which in turn reduces the cash inflow, may suggest firms trying to boost sales volume and earnings through credit extensions. This signifies a lower persistence of the earnings figure. High quality earnings companies usually have CR above one.

$$\text{CR} = \frac{\text{Cash Flow from Operating Activities}}{\text{Net Income}} \tag{2}$$

2.3.3 Productive Asset Reinvestment Ratio (PARR)

An increase in earnings from a reduction in *Research and Development* (R&D) or capital expenditure is not good. This phenomenon is common in times of recession as companies adopt a more conservative approach. However, a reduction in a company's commitment to investing in capital assets could affect its future growth; the impact is particularly serious in companies heavily involved in research and technology. Furthermore, a decrease in capital expenditure might indicate a company with insufficient funds to maintain its level of investment in capital assets. *Productive Asset Reinvestment Ratio* (PARR) measures the sustainability of a company's growth. Short-term managerial objectives are often shown by a low PARR whereas companies with high quality earnings have PARR above unity.

$$\text{PARR} = \frac{\text{Capital Expenditure}}{\text{Depreciation Expense}} \tag{3}$$

2.3.4 Tax Rate (TR)

Analysts usually perceive a substantial change in a firm's effective tax rate as transient since it might be a result of statutory tax changes. Companies with high earnings quality tend to have a tax rate greater than the average tax rate reported for all companies.

$$\text{TR} = \frac{\text{Tax Expense}}{\text{Pre-tax Income}}\% \tag{4}$$

2.3.5 S&P Equity Score

S&P ranks companies based on their long term growth prospects and the stability in dividend payments and earnings. Companies with high ranking tend to outperform the S&P 500 index and have fewer fundamental risks. These companies are rated as A+, whereas companies likely to face bankruptcy are rated D. Companies with high rating tend to exhibit high earnings quality in terms of fewer cases of earnings manipulation. They are also less susceptible to business and credit cycles. Stocks with high ratings tend to outperform the index even in a bear market. This measure is a strong indicator of earnings quality.

2.3.6 S&P Debt Score

The S&P long term debt score rates the credit-worthiness of a company based on the following principles: independence, objectivity, integrity and disclosure. Business fundamentals are thoroughly investigated in the rating process; growth prospects and industry outlook are assessed. The rating is reviewed frequently to capture changes arising from developments such as mergers and acquisitions, as well as changes in economic conditions. This rating is included as a quality of earning measure to account for the effect of industry-specific balance sheet structures and facilitate comparisons between sectors.

3 Statistical Methods

3.1 Specification of the Problem

We aim to find which of the indicator measures described in Subsection 2.3 has the highest predictive power in determining whether or not a company will meet its consensus EPS forecasts. The group companies of that have met this criterion will be referred to as the 'met group' and the 'not-met group' otherwise. This allows us to compile a quality of earnings index using an optimum combination of the financial ratios based on the criterion that a high index indicates earnings predictability. The composition of possible indices for 1992-2001 is investigated with First Call consensus EPS as the target earnings figure.

Figure 2 depicts possible statistical techniques from which appropriate methods are chosen depending on the problem under investigation. We start our investigation by examining possible differences between the indicator measures of the two groups – those meeting and not meeting their projected EPS. Here, group membership is treated as the dependent variable, whereas the indicator measures are the independent variables. Hence, Multivariate Analysis of Variance (MANOVA) is the appropriate technique to be applied. Our investigation will

Variables				Classification Method
Independent		**Dependent**		
Num	**Cont***	**Num**	**Cont***	
One	Yes	One	Yes	Regression
One	Yes	One	No	Discriminant Analysis / Logistic Regression
One	Yes	Many	Yes	Canonical Correlation
One	Yes	Many	No	Multiple-group discriminant analysis (MDA)
One	No	One	Yes	t-test
One	No	One	No	Discrete Discriminant Analysis
One	No	Many	Yes	Multivariate Analysis of Variance (MANOVA)
One	No	Many	No	Discrete MDA
Many	Yes	One	Yes	Multiple Regression
Many	Yes	One	No	Discriminant Analysis / Logistic Regression
Many	Yes	Many	Yes	Canonical Correlation
Many	Yes	Many	No	MDA
Many	No	One	Yes	ANOVA
Many	No	One	No	Discrete Discriminant analysis / Conjoint analysis (MONANOVA)
Many	No	Many	Yes	MANOVA
Many	No	Many	No	Discrete MDA

*: 'yes' if the variable is continuous, 'no' if not.

Table 2: Classification of different statistical techniques.

also look into univariate dependence tests to identify individual indicator measures with good discrimination power, if any. Fisher discriminant analysis and logistic regression are carried out at a later stage of the investigation to determine the optimum weighting of these measures that yields the highest predictability of earnings forecasts.

3.2 Dependence Statistical Analysis

Dependence statistical techniques are used to investigate the significance of differences in mean, variance and the distribution of two or more groups. Groups with sizeable differences in their characteristics can be successfully distinguished. The methods used are Student's t-test, the F-test, the Kolmogorov-Smirnov (KS) test and MANOVA.

3.2.1 Student's t-test

The t-test is used to test for equality of means between normally distributed populations. Samples in the two groups are assumed to be randomly selected and representative of larger populations. This test is robust to moderate deviations from these assumptions provided that the samples are not too small and are of equal size. The t-statistic for populations with equal variance is given by

$$\begin{aligned} t &= \frac{\bar{x}_1 - \bar{x}_2}{S} \qquad (5) \\ \text{degrees of freedom (d.o.f)} &= N_1 + N_2 - 2 \end{aligned}$$

where $\bar{x}_i$ is the sample mean of group i and N_i denotes group i's sample size. The sample 'pooled variance', S, takes unequal sample sizes into account and is given by

$$S = \sqrt{\frac{\sum_{i \in 1}(x_i - \bar{x}_1)^2 + \sum_{i \in 2}(x_i - \bar{x}_2)^2}{N_1 + N_2 - 2} \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}$$

When there is a large difference in the sample variances, the unequal variance t-test is used:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{[Var(x_1)/N_1 + Var(x_2)/N_2]}} \qquad (6)$$

$$\text{d.o.f} = \frac{\left[\frac{Var(x_1)}{N_1} + \frac{Var(x_2)}{N_2} \right]^2}{\frac{[Var(x_1)/N_1]^2}{N_1 - 1} + \frac{[Var(x_2)/N_2]^2}{N_2 - 1}} \qquad (7)$$

If, however, the differences in variance are sufficiently large to warrant the use of this form, it implies that the mean may not be useful in distinguishing between the groups. The F-test, described next, would be used instead.

3.2.2 F-test

The F-test is used to test the null hypothesis of equality of variance. The test statistic is given by the ratio of the two variances, thus values $\gg 1$ or $\ll 1$ indicate significant differences in variance. The F-statistic is given by

$$F = \frac{Var(x_1)}{Var(x_2)} \qquad (8)$$

3.2.3 Kolmogrov-Smirnov Test

The Kolmogrov-Smirnov (KS) test is used to test whether two distributions are similar. It is a simple test that compares the *cumulative distribution functions* (c.d.f) of the two data sets. A simple unbiased estimator of the c.d.f, $C_N(x)$, for the N data points, x_i, where $i = 1, \ldots, N$, can be defined as the total number of observations smaller than x. Hence, $C_N(x)$ is constant between data points, and the jump at each of the data point is $1/N$. If the c.d.f of two data sets are C_{N1} and C_{N2}, the KS-statistic is computed as

$$KS = \max_{-\infty < x < \infty} |C_{N1}(x) - C_{N2}(x)| \tag{9}$$

A disadvantage of the KS-test is its insensitivity to data points near the extreme ends of the distribution. This is because the end values for all c.d.f. are 0 and 1, preventing the KS-test from accurately identifying mismatch in the distribution at either end. Several variants of the test alleviate this problem, but will not be discussed here. See Chapter 14 of Numerical Recipes in C by Vetterling *et al.* [13] for more details.

3.2.4 Multivariate Analysis of Variance

The above univariate tests do not take account of interactions between the dependent variables. A multivariate test is usually preferred to separate univariate tests on each of the dependent variables because the latter is subject to Type I errors – false detection of statistically significant differences – caused by correlation between parameters [16]. Multivariate Analysis of Variance (MANOVA) is the multivariate equivalent of a t-test; it tests the mean differences in the combinations of the dependent variables. The advantage of MANOVA over two univariate t-test is illustrated in Figure 1 where the latter would produce a false negative – Type II error – and the former correctly distinguishes between the two groups. The null and alternative hypotheses for MANOVA are $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ and $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$.

MANOVA assumes that independent samples are obtained from multivariate normal distributions, $\mathbf{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathbf{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. In addition, the covariance matrices are assumed to be equal, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. Similar to univariate t-test, MANOVA is robust to small violations of assumptions as long as the samples are large and equal in size. The associated Hotelling T^2

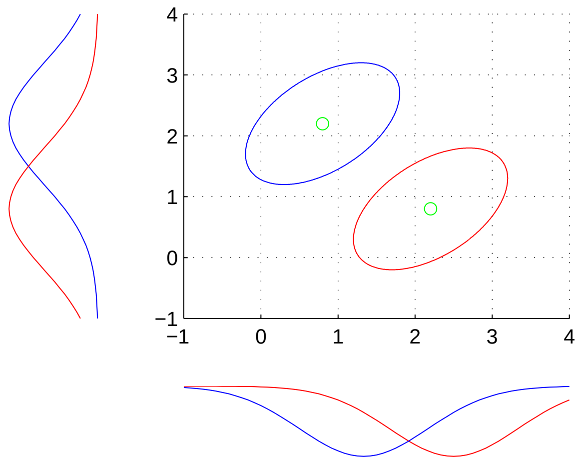

Figure 1: Successful discrimination using multivariate analysis whereas separate t-tests (left and below) fail

test statistic is given by:

$$T^2 = \frac{N_1 N_2 (N_1 + N_2 - 2)}{N_1 + N_2} \mathbf{x}' \mathbf{S}^{-1} \mathbf{x} \tag{10}$$

where $\mathbf{x} = \hat{\mathbf{X}}_1 - \hat{\mathbf{X}}_2$ and $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$, the pooled within-group SSCP.

The Hotelling Trace, U, is more commonly used than Hotelling T^2 and is defined as

$$U = \frac{T^2}{N_1 + N_2 - 2} \tag{11}$$

3.3 Methods for Discrimination

Discriminant analysis (also known as classification) divides and groups of data into categories based on statistical information. This process can be carried out in two ways, each solution attempting to minimize the cost of misclassification:

- **Supervised**: Group memberships are known beforehand and a suitable discrimination function is computed from the samples such that it best distinguishes between the groups. Examples of these discrimination methods are Fisher discriminant analysis, logistic regression and Bayesian classification
- **Unsupervised**: Does not require prior information about group memberships. The

discrimination process automatically identifies a solution that optimally separates the data. This type of classification method includes principal component analysis, decision-tree classification and Gaussian mixture model.

3.3.1 Fisher Linear Discriminant Analysis

Fisher proposed an alternative approach to the two-group problem, his discriminant analysis finds the direction that maximizes the distance between projected means while, at the same time, minimizing the in-group scatter – group variances. Consider a d-dimensional sample $\mathbf{x}$ from two groups, w_1 and w_2. Fisher discriminant analysis finds the optimal Fisher coefficient vector, $\mathbf{w}$, such that the projected value or discrimination score satisfies

$$\mathbf{w} = \mathbf{S}_W^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \tag{12}$$

In some cases, especially during exploratory work, we may want to find a small set of variables that best differentiate between the groups from a larger set. This is accomplished by dynamically selecting the appropriate variables based on certain statistical criteria, such as the Wilk's Λ, Rao's V, Mahalanobis squared distance and between-group F-ratio. Wilk's Λ criterion will be described here and used in subsequent analysis, it is defined as

$$\Lambda = \frac{SS_w}{SS_t} = \frac{SS_w}{SS_b + SS_w} \tag{13}$$

where SS_w, SS_b and SS_t are the sum of squares for within-group, between-group and total scatter respectively. Small values of Λ indicate scatter is low *within-group* but large *between groups*, thus implying good classification properties. Hence variables with the largest (smallest) Wilk's Λ are added (removed) in the variable selection process. The algorithms used in selecting the appropriate variables are as follows:

- **Forward Selection**: This begins with no selected explanatory variables in the discrimination function. The variable that gives the greatest increase in discrimination power is selected and this process is repeated on the remaining variables until the improvement in discrimination is saturated.

- **Backward Selection**: In contrast to the above, this method begins with all variables present in the discrimination function. Variables with the worst discrimination power are removed repeatedly until the decrease in discrimination power falls below a certain threshold.

- **Stepwise Selection**: This is a combination of the above two methods. It begins without any variables present in the discrimination function. Forward and backward selections are alternated until no variable is added to, or removed from, the function.

3.3.2 Logistic Regression

Logistic regression differs from discriminant analysis in taking exponentials of its independent variables and calculating *odds* rather than *probabilities*,

$$\frac{p}{1-p} = e^{\beta_0} + \prod_{i=1}^{N} e^{\beta_0 x_i} \tag{14}$$

This is linearized by taking logs and can be solved using methods such as maximum likelihood estimation. See Sharma [15] for more information.

Logistic regression has several advantages over the conventional discrimination methods. It provides more flexibility by not making any assumption about the distribution of the observation vectors (predictors) as opposed to the normality distribution assumed by the traditional classifiers. Logistic regression accommodates a mixture of dichotomous and continuous variables, in comparison to Fisher discriminant analysis which is only applicable, strictly speaking, to continuous variables. Unlike multiple regression analysis, the exponential form in logistic regression guarantees non-negative probability.

However, logistic regression does have a few limitations. Despite its flexibility, it is less powerful than the alternative methods when none of those strict assumptions are violated. Further, logistic regression assumes linearity between the continuous predictors. In addition, logistic regression is very sensitive to high correlations amongst the predictors, and this is signified by large standard errors in the estimation of the coefficients. 'One-sided' bias is often observed in logistic regression where classification accuracy is very high for a certain group, but low for the other.

3.4 Principal Component Analysis

Principal component analysis (PCA) is an unsupervised classification method since it assumes no prior knowledge of the group membership. It can be used to decorrelate the variables and select the prominent indicator measures. Say we have a group of vectors, $\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_N$ with averages $\bar{U}_i$, we can express their fluctuations as $\mathbf{u}_i = \mathbf{U}_i - \bar{U}_i$. These can be orthogonalized using some matrix, $\mathbf{\Lambda}$:

$$\mathbf{v} = \mathbf{u}\mathbf{\Lambda}^T \tag{15}$$

These linear combinations can be sorted in order of variance. See Papoulis [11] for more information.

4 Discussion

The analysis methods introduced in the previous section are used on a ten-year sample. This sample is based on 131 companies from the S&P 500 index from 1992 to 2001. Further analysis splits the 10 years of data into three periods based on market sentiment where the sample for the last two years is expanded to facilitate sector-based analysis and to investigate further the sudden change in market conditions.

4.1 Data Preparation

4.1.1 Availability

Two overlapping data sets, both drawn from the S&P 500 index between 1992 and 2001, have been provided by Merrill Lynch where one spanned the entire length for a relatively small number of companies and the other spanning only three years but includes a larger number of companies. Table 3 gives more information, including sample sizes after treatment of outliers. The longer period covers fewer companies because the index is not constant, with firms being added and removed depending on their market capitalization and various other factors. This introduces a survivor bias, although the effect is not believed to be significant given that removal indicates divergence rather than, say, bankruptcy. A full investigation of the effects is beyond the scope of this report but may be considered in the future. S&P equity and debt ratings are not available for the entire data series because the figures are unavailable early in the sample. The impact of their removal is studied using the short sample by running correlations of the ratings against the four indicator measures as shown in Table 4. Initially, all analyses will consider the long data set because our aim is to find long-term relations though the short one is equally important because it allows the recent bear market to be covered in greater detail.

Span	**1992 - 2001**	**1999 - 2001**
Cross Section		
Before Cleaning	131	405
After Cleaning	131	366
Indicators Provided		
RoCE	✓	✓
CR	✓	✓
PARR	✓	✓
TR	✓	✓
S&P Equity		✓
S&P Debt		✓

Table 3: Data sets provided by Merrill Lynch

	Correlation							
	S&P Equity				S&P Debt			
	1999	2000	2001	Overall	1999	2000	2001	Overall
RoCE	0.22	0.26	0.40	0.29	0.21	0.30	0.40	0.30
CR	7.4×10^{-4}	-0.039	4.1×10^{-4}	-0.038	-0.037	-0.13	0.012	-0.052
PARR	0.090	0.056	0.16	0.10	-0.020	-0.044	0.013	-0.017
TR	0.13	-0.026	0.21	0.10	0.068	0.063	0.16	0.097

Table 4: Correlation for S&P Ratings showing a weak relation with individual indicator measures

4.1.2 Cleaning

This subsection considers the outliers and other spurious data which can be found in the raw data as supplied initially. For example, visual inspection reveals the problems shown in Table 5.

Outliers can be caused by erroneous data or extreme values and indicate the need for further verification. Filtering methods are likely to misdetect or fail to recognize outliers so data from

Outliers
Ball Corp Tax Rate (1995..1997): [1039, 606, 584]
International Paper Cash Realization Ratio (1997..2001): [0, 7, 9, 17, 0]
Viacom Tax Rate (1998): 1357
Missing Values
Raytheon Total Assets and Current Liabilities (1995-8)
Sprint Total Assets and Current Liabilities (1995-8)
General Mills All data (1995)
Repeated Values
Tenet Healthcare Repeated values in 1998 and 1999

Table 5: Some examples of outliers and missing values in the data before cleaning

a separate source was used for comparisons. Full information – to reproduce the indicator measures – is available from *Datastream* and *Moneyline Telerate* for all companies except Alcan, Inco, Philip Morris, Phillip Petroleum, Transocean and TRW. The records obtained are of similar quality as the original data set in terms of number of outliers. See Table 21 in Appendix B. The results vary significantly for profit and loss data, pre-tax income and tax paid in particular, as depicted in Table 6.

Type	**Financial Statement Items**	**Percentage**	**Average**
Profit & loss	Depreciation	10.8	
	Interest Expense	14.9	
	Net Income	6.4	16.0
	Pretax Income	15.9	
	Income Tax Paid	31.9	
Balance Sheet	Total Assets	3.3	3.3
	Current Liabilities	3.3	
Cash Flow	Cash flow Operations	3.7	3.7

Table 6: Percentage of non-matching data with a difference greater than 10%.

The cleaning algorithm used is

1. Exclude companies without complete data thus reducing occurrences of missing values
2. Replace missing values with time series averages
3. Apply a filtering band of two standard deviations from the mean.

Missing values can also be replaced with cross-sectional means, but these do not allow for the heterogeneity normally expected between firms. Another solution is interpolation, but there is no guarantee that the values would follow a regular path over time. Hence, an arbitrary decision to use the average was made. Applying the filtering bands reduces the impact of erroneous and extreme values on the performance of the statistical analyses carried out. This band has been set to two standard deviations from the cross-sectional mean, where values exceeding the range are capped to the corresponding upper or lower limit. This approach does not account for circumstances where companies have consistently outperformed or underperformed compared to their peers, but this is not believed to cause significant problems since the objective is to find a general model. The percentage of data affected is shown in Table 7.

Indicator Measures	**RoCE**	**CR**	**PARR**	**TR**
Mean(average over 10 years)	17.4	2.7	1.6	36.4
Standard deviation	15.3	7.2	1.0	58.8
Percentage of outliers	5.1	2.8	4.1	3.0

Table 7: Descriptive statistics and the percentage of outliers.

4.1.3 Fundamentals

Before testing predictive power of the indicators, it is useful to get an indication of their connection to the fundamentals behind EPS figures. Table 8 shows the correlation between four indicator measures and pro-forma EPS for the same year and the year ahead where applicable. Appendix B shows yearly results. Table 9 shows high correlation between years for two of the indicators, RoCE and PARR, as well as Consensus and Pro-Forma EPS figures. For RoCE, this trend ends for 2000 onward pointing, in hindsight, toward changing market conditions and the correlation for the two EPS figures drop significantly as well but only in 2001. There is a notable outlier in CEPS between 1994 to 1995, which is investigated further in later sections. Though cointegration was thought to be a viable methodology initially, the 10-year span was found to be insufficient for this.

Indicator	Pro-forma		Consensus
	same year	+1 year	(same year)
RoCE	0.12	0.033	-0.026
CR	4.1×10^{-3}	-0.033	0.033
PARR	-0.10	-0.17	-0.10
TR	0.017	5.0×10^{-3}	3.4×10^{-3}

Table 8: Contemporaneous and 1-year predictive (lagged) correlation between the four indicator measures and pro-forma EPS. Correlation with consensus EPS for the same year is shown in the fourth column.

Year	**RoCE**	**CR**	**PARR**	**TR**	**CEPS**	**EPS**
92-93	**0.86** (0.00)	**0.52** (0.00)	**0.79** (0.00)	**0.34** (0.00)	**0.88** (0.00)	**0.89** (0.00)
93-94	**0.78** (0.00)	0.07 (0.40)	**0.67** (0.00)	**0.40** (0.00)	**0.94** (0.00)	**0.87** (0.00)
94-95	**0.72** (0.00)	**0.40** (0.00)	**0.66** (0.00)	**0.23** (0.008)	**0.26** (0.002)	**0.82** (0.00)
95-96	**0.73** (0.00)	**0.23** (0.007)	**0.77** (0.00)	*0.22* (0.014)	**0.81** (0.00)	**0.85** (0.00)
96-97	**0.75** (0.00)	0.14 (0.10)	**0.78** (0.00)	*0.17* (0.049)	**0.92** (0.00)	**0.95** (0.00)
97-98	**0.70** (0.00)	0.04 (0.63)	**0.74** (0.00)	*0.10* (0.047)	**0.92** (0.00)	**0.81** (0.00)
98-99	**0.77** (0.00)	**0.31** (0.00)	**0.76** (0.00)	**0.22** (0.010)	**0.81** (0.00)	**0.85** (0.00)
99-00	**0.65** (0.00)	0.03 (0.70)	**0.77** (0.00)	*0.22* (0.012)	**0.76** (0.00)	**0.71** (0.00)
00-01	**0.42** (0.00)	0.02 (0.87)	**0.75** (0.00)	**0.36** (0.00)	**0.63** (0.00)	**0.56** (0.00)

bold items are significant at 1% level, *italic* items are significant at 5% level

Table 9: Correlation of indicators and Consensus & Pro-Forma EPS between years

4.1.4 Consensus EPS

Before moving to dependence statistical analyses, it may be of interest to consider the correct prediction rate using consensus EPS values. The average error is positive for overpredictions and, zero, obviously indicates no error.

Year	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001***
Rate	50	52	57	70	67	58	59	66	40
Avg error	-0.0108	-0.106	-0.286	-0.0670	-0.0689	0.112	0.0798	-0.372	0.614
σ	1.11	1.16	1.35	0.71	0.63	0.88	0.78	1.40	2.25

* Merrill Lynch has linked analyst bonuses to forecasts since 2002 to penalize overpredictions

4.2 Data Exploration and Statistical Analysis

This subsection presents the results from dependence statistical tests, used to determine those distinct properties that may be exploitable in predicting whether the achieved pro-forma EPS exceeds or falls short of the previous year's consensus forecasts.

4.2.1 Distributions

Figure 2 shows the distributions of the indicator measures combined for all years, with a normal distribution with the same mean and variance superimposed. Return on Capital Employed appears to yield a near symmetric distribution. This is not the case, however, for the other indicator measures; Tax Rate and Cash Realization are strongly skewed. Productive

Asset Reinvestment Ratio is skewed also and has fat tails. These visual observations are supported by the Jarque Bera test [16], conducted on the data on a yearly basis with the results shown in Table 10. Hence, any statistical test with a strong assumption of normality will need to be tested for robustness against violations. This is most likely to impact the discriminant analysis, as discussed later.

Figure 3 illustrates another possible problem for the regressions. Splitting the companies into two groups, those having met or exceeded the consensus EPS forecast and those falling short, we find a large overlap between values for the indicator measures. Amongst these, RoCE has the largest separation and will probably have greatest predictive value while the others are expected to offer only limited discrimination power. Splitting the set into individual years does not improve the result significantly. These results are reported in Appendix C.

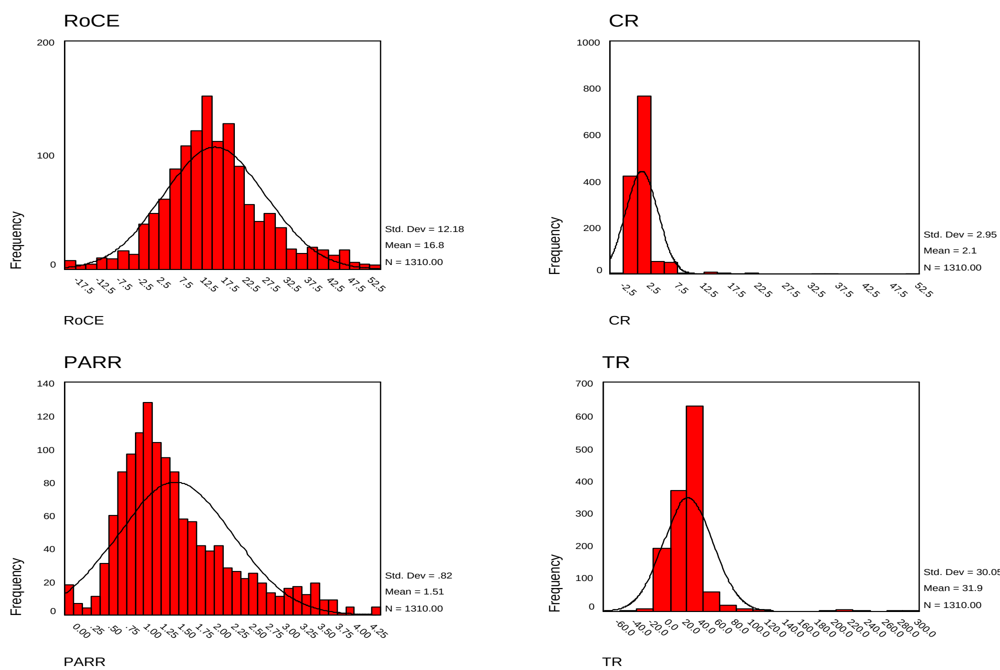

Figure 2: Histograms of the four indicator measures: RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right)

4.2.2 Student's t, F, KS and MANOVA

These four tests were run on individual indicator measures with the results shown in Table 12. Means of the indicator measures do not differ significantly between the two groups, which will

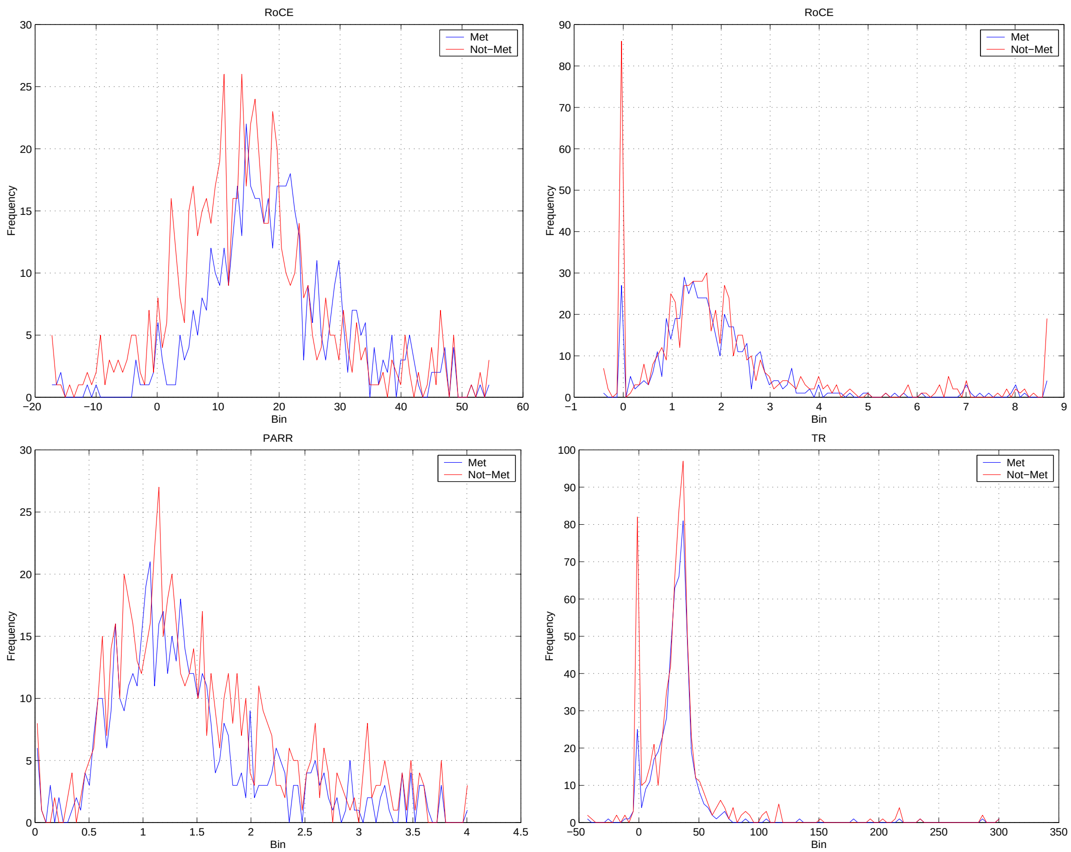

Figure 3: Comparison of the distributions of the four indicator measures split into the met and not-met groups of companies: RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right).

Year	RoCE	CR	PARR	TR
1992	**3.20**	2302.05	45.84	**0.04**
1993	**3.09**	22199.64	19.83	176.66
1994	12.49	249.97	31.09	**2.25**
1995	16.67	180.08	9.98	2337.51
1996	15.43	480.15	16.48	1046.70
1997	**5.55**	478.61	9.49	2247.32
1998	**3.97**	169.97	15.50	2508.11
1999	*9.16*	210.54	38.10	19.80
2000	*6.26*	491.02	22.70	255.34
2001	**0.35**	7372.27	19.80	**3.48**

Table 10: Yearly Jarque Bera test results for the four indicators. Bold results are not rejected at the 95% level and italic ones at 99%

Year	Hotelling Trace	F-statistic	Sig
1993	0.053	1.771	0.139
1994	0.025	0.797	0.529
1995	0.033	1.059	0.380
1996	0.115	4.109	0.004
1997	0.075	2.554	0.042
1998	0.031	1.000	0.410
1999	0.064	2.136	0.080
2000	0.050	1.646	0.167
2001	0.005	0.144	0.965

Table 11: Results of MANOVA test on the four indicator measures.

be problematic as many discrimination methods rely heavily on separation of means. The F-test, on the other hand, shows that the variances of CR and TR indicators differ consistently for the two groups except for 1995 and 1998 where it is insignificant. While linear discriminant analysis relies heavily on separation of mean, quadratic discrimination incorporates differences in variance into the discrimination process. The KS-test and MANOVA (Table 11) also give negative results, indicating a high correlation between the distributions of indicator measures for the two groups, even in higher dimensional spaces.

4.2.3 Variation over Time

The graphs in figure 4 show how the cross-sectional mean of each of the indicator measures changes over time, including the difference observed when splitting the series into two groups – those having and not having met the forecast. RoCE increases steadily to 1999 although the indicator rebounds in 2000 and dips to a record low in the following year. The PARR means decrease consistently over the years while CR does so for the 'met group' only, with a sudden inrease found after 1999 for overall and not-met results. Tax rates are exceptionally high in 1995 and 1998, although it is interesting to note that they were, on average, lower for companies meeting the forecast.

We also note that the RoCE curves retain some separation in all years, which will help to discriminate between groups. PARR displays similar characteristics from 1992 to 1997, so its discriminative power is likely to be just below RoCE's. Curves for CR and TR intersect on several occasions, which may limit their use in the analyses. For instance, TR for the met group was greater than the not met group between 1992 to 1994, but this reversed for the subsequent three years, reverting to its original trend between 1998-2001. We will show in

Year	RoCE			CR		
	t-test	F-test	KS-test	t-test	F-test	KS-test
1993	2.14 (*0.03*)	1.07 (0.79)	0.33 (***0.00***)	0.26 (0.79)	1.63 (*0.05*)	0.17 (0.31)
1994	1.29 (0.20)	1.48 (0.12)	0.22 (0.07)	-0.91 (0.37)	224.11 (***0.00***)	0.14 (0.56)
1995	0.24 (0.81)	1.24 (0.39)	0.12 (0.72)	0.04 (0.97)	1.04 (0.89)	0.19 (0.16)
1996	-1.06 (0.29)	2.53 (***0.00***)	0.11 (0.79)	2.13 (*0.04*)	5.60 (***0.00***)	0.25 (***0.03***)
1997	0.07 (0.94)	5.53 (***0.00***)	0.24 (*0.04*)	-0.73 (0.47)	10.23 (***0.00***)	0.21 (0.11)
1998	1.31 (0.19)	1.69 (*0.05*)	0.25 (*0.04*)	-0.88 (0.38)	28.68 (***0.00***)	0.18 (0.26)
1999	-0.69 (0.49)	3.54 (***0.00***)	0.12 (0.69)	-1.01 (0.31)	5.31 (***0.00***)	0.12 (0.75)
2000	1.34 (0.18)	1.24 (0.41)	0.20 (0.15)	1.39 (0.17)	2.51 (***0.00***)	0.15 (0.46)
2001	-0.38 (0.70)	1.61 (0.14)	0.15 (0.64)	-0.79 (0.43)	8.35 (***0.00***)	0.19 (0.33)

Year	PARR			TR		
	t-test	F-test	KS-test	t-test	F-test	KS-test
1993	0.85 (0.40)	1.96 (***0.01***)	0.13 (0.63)	1.72 (0.09)	1.55 (0.10)	0.22 (0.08)
1994	-0.39 (0.69)	1.14 (0.63)	0.12 (0.68)	-0.05 (0.96)	2.41 (***0.00***)	0.16 (0.33)
1995	-0.88 (0.38)	1.38 (0.20)	0.19 (0.18)	-0.07 (0.94)	1.62 (0.06)	0.11 (0.79)
1996	-2.89 (***0.00***)	2.17 (***0.00***)	0.40 (***0.00***)	-1.08 (0.28)	6.78 (***0.00***)	0.12 (0.76)
1997	-2.03 (*0.04*)	2.55 (***0.00***)	0.15 (0.46)	-1.95 (*0.05*)	17.68 (***0.00***)	0.18 (0.22)
1998	-0.46 (0.64)	1.18 (0.54)	0.19 (0.18)	0.57 (0.57)	1.34 (0.24)	0.18 (0.23)
1999	-2.27 (*0.02*)	2.39 (***0.00***)	0.28 (*0.01*)	-1.07 (0.28)	9.05 (***0.00***)	0.15 (0.44)
2000	0.99 (0.32)	1.85 (*0.01*)	0.09 (0.95)	0.21 (0.83)	2.39 (***0.00***)	0.17 (0.31)
2001	0.10 (0.92)	1.42 (0.20)	0.09 (0.99)	0.28 (0.78)	1.69 (0.06)	0.17 (0.45)

Italic indicates significance at the 5% level; ***bold italic*** indicates significance at the 1% level. This convention will be adopted throughout this report.

Table 12: Student's t, F and KS tests on the four indicator measures

later sections that these observations correspond well with the best single grouping variable experiments.

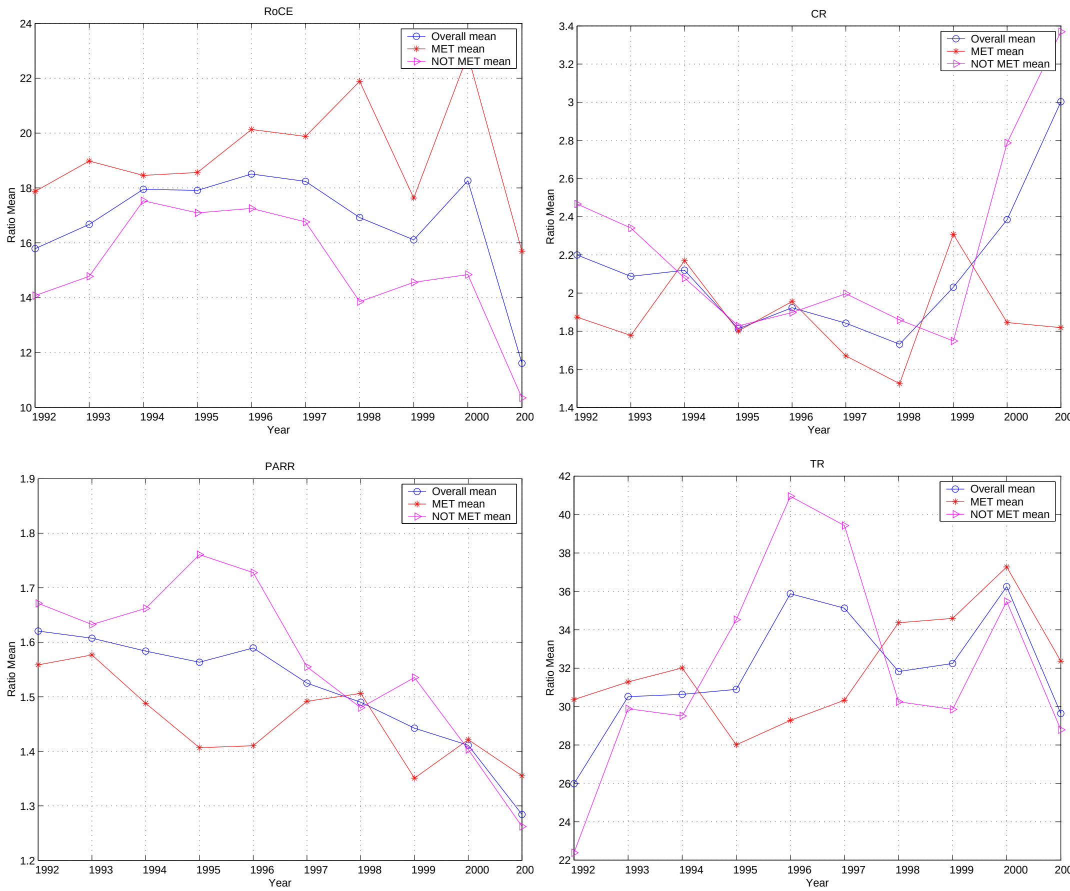

Figure 4: Variations of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right).

4.3 Best Single Grouping Variable

This subsection examines the discriminative power of each of the indicator measures. Predictions are made based on the nearest Euclidean distance[2] where the discrimination point, p, is

$$
\begin{aligned}
(p-\mu_m)/\sigma_m &= (p-\mu_n)/\sigma_n \\
p &= (\sigma_m\mu_n - \sigma_n\mu_m)/(\sigma_m - \sigma_n) \qquad (16)
\end{aligned}
$$

[2]Euclidean distance, d, is defined as $d = (x-\mu)/\sigma$

where (μ_m, σ_m) and (μ_n, σ_n) are the mean and variance of a single indicator measure for the met and not-met groups respectively. Prediction accuracies for the indicators lie within a relatively small range of 45 to 55% with no single measure standing out so they will be assumed to be of equal importance in the regression analyses. However, RoCE does depart from the norm in two years, where discrimination based solely on this indicator yields accuracies as high as 62% in 1998 and 76% in 2001 as illustrated in Figure 5. Combining all ten years of data for each of the indicators produces only a small difference compared to taking an average of the prediction rate for each year independently. RoCE is found to produce the highest discrimination power in both cases, but no single measure is sufficient for accurate predictions.

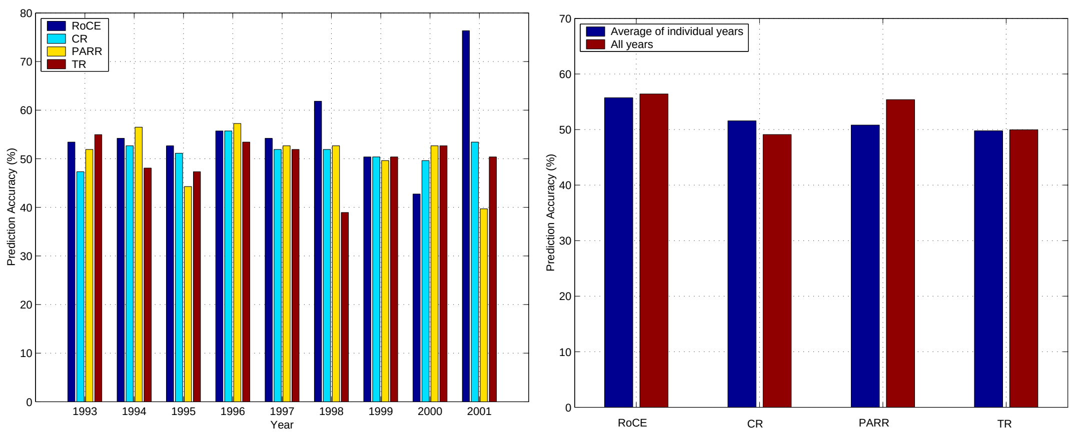

Figure 5: Overall prediction accuracies for each of the indicator measure. (Left: Yearly predictions; Right: Comparison between the average of the yearly prediction results and the prediction accuracies obtained using the ten-year combined data)

4.4 Fisher Discriminant Analysis

Fisher discriminant analysis allows for multiple independent variables, thus allowing all four measures to be used simultaneously unlike the method used in the previous section. A fifth indicator, meeting of consensus EPS in the previous year, is added in subsection 4.4.2 and is found to improve predictive accuracy. Our investigation is based on in-sample discrimination, and with out-of-sample testing if the indicator measures are found to have sufficient discriminatory power.

4.4.1 Using the Four Indicator Measures

Fisher discriminant analysis on the four indicator measures can be formulated as

$$y_t = \beta_{0,t} + \beta_{1,t} RoCE_{t-1} + \beta_{2,t} CR_{t-1} + \beta_{3,t} PARR_{t-1} + \beta_{4,t} TR_{t-1} + \varepsilon_t \tag{17}$$

where $t = 2 \ldots 10$ and the projected value, y_t is

$$y_t = \begin{cases} \geq 0 & \text{for} \quad Met_t = 1 \\ < 0 & \text{for} \quad Met_t = 0 \end{cases}$$

The resulting Fisher direction, given by the Fisher coefficient vector, $\boldsymbol{\beta}$, maximizes the criteria stated in Subsection 3.3.1 (equation 13).

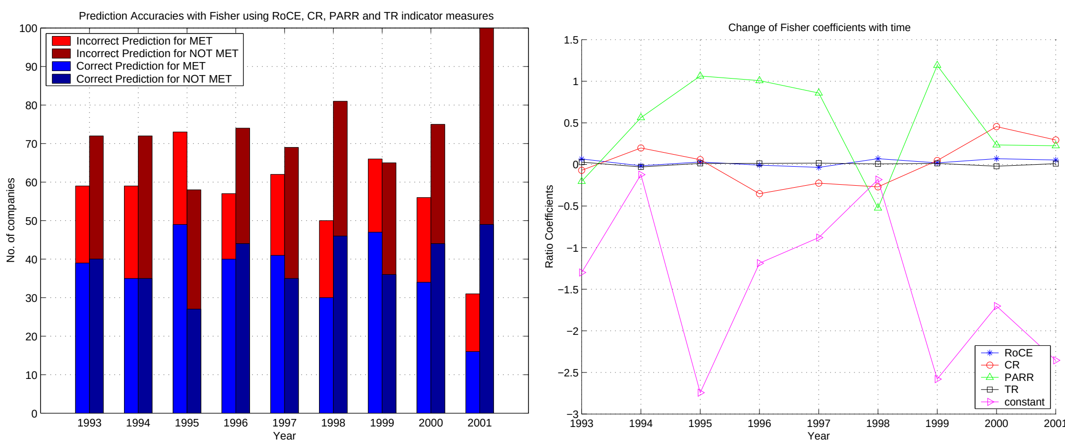

Figure 6: Fisher discriminant analysis using the four indicator measures. (Left: Yearly prediction accuracies; Right: Time variation of the Fisher coefficients)

Fisher discriminant analysis using the four indicators does not provide much advantage over the 50-50 chance of getting a correct prediction through an unbiased random estimation. Insignificant Wilks' Λs at the 5% level are observed in most cases (Table 23 in Appendix D highlights this). Prediction accuracy falls below the 50% level in some years. The proportion of correct predictions for the two groups is approximately constant (Table 13) even though the distribution of the companies into the met and not met groups varies over time, especially in 2001 with only 24% meeting their EPS forecasts. This indicates that the Fisher discriminant method is robust against unequal group sizes. We shall see later that this is not the case for logistic regression.

Met to Not Met Ratio	1993	1994	1995	1996	1997	1998	1999	2000	2001
Number of Companies	0.82	0.82	1.26	0.77	0.90	0.62	1.02	0.86	0.31
Predictive Accuracy	1.19	1.22	1.44	1.18	1.30	1.06	1.29	1.03	1.05

Table 13: Comparison of the performance of Fisher discriminant analysis with the variation of the groups' sample sizes.

Fisher coefficients, especially those for the PARR indicator measure, and the intercept vary over time. The large fluctuation in the constant terms suggests that the selected indicator measures fail to include all the factors affecting the predictability of meeting EPS forecasts. The troughs in the constant terms observed in 1995 and 1999 are related to the change in market conditions: investors shifted their emphasis from mature companies to high growth technology stocks in 1995 during the bull market; whereas 1999 corresponds to the burst of the technology bubble and the reverse process. This intertemporal instability of the Fisher coefficients implies that prediction using the indicator measures will be difficult, even if they were to have excellent discrimination power. However, discriminatory power is found lacking and deteriorates further over time.

4.4.2 Lagged Meeting of Estimates

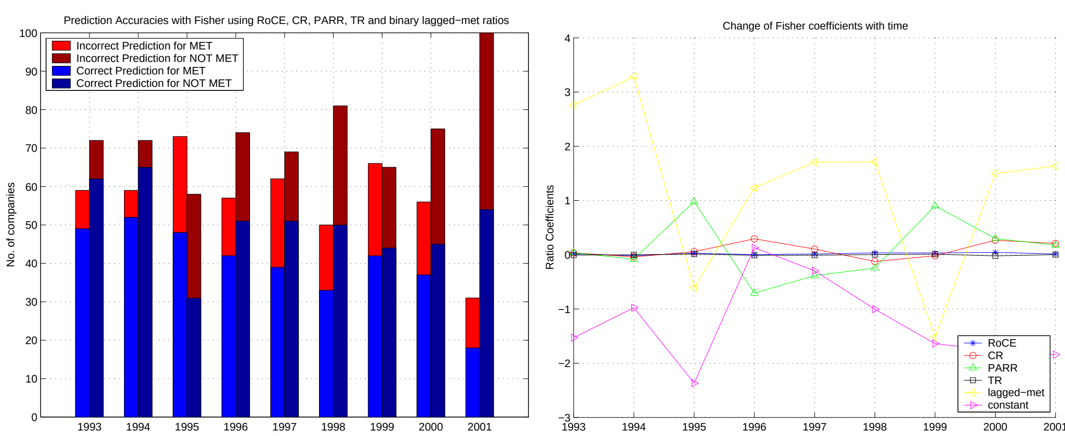

Figure 7: Fisher discriminant analysis with added binary lagged-met term (Left: yearly prediction accuracies; Right: Time variation of the Fisher coefficients)

Equation (17) can be extended to include a lagged-met indicator such that

$$y_t = \beta_{0,t} + \beta_{1,t}RoCE_{t-1} + \beta_{2,t}CR_{t-1} + \beta_{3,t}PARR_{t-1} + \beta_{4,t}TR_{t-1} + \beta_{5,t}Met_{t-1} + \varepsilon_t \quad (18)$$

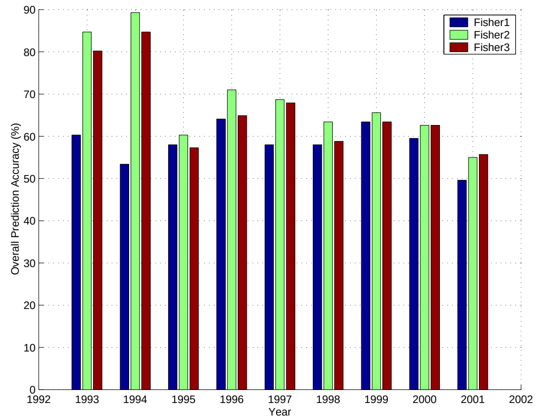

Figure 8: Prediction accuracies using Fisher discriminant analysis (Fisher1: Discrimination using the four indicator measures; Fisher2: With added binary lagged-met indicator; Fisher3: With continuous lagged-met indicator)

The lagged met term indicates whether a particular company has met the consensus EPS projection from the previous year. This term can be defined as:

1. a binary variable. A value of 1 indicates that the company has met or exceeded the forecast in the previous year and a value of 0 shows otherwise

2. a continuous variable of value given by the difference between the Pro-forma and the consensus EPS of the previous year. A positive value indicates meeting a forecast and a negative value indicates not meeting.

Adding a binary lagged met term into the analysis improves predictive accuracy by 20-30% as shown in see figure 8 and some of the time variation of the Fisher constant is absorbed by the term. This can be seen from the proximity of the variation of the coefficient to that of the constant term in the previous analysis (Figure 7). Hence, the added indicator correlates well with market conditions, which explains the improvements in predictive accuracy. A large gain is observed but only for the first two years. Further investigation is required to understand the factors which could have caused this effect.

Incorporating a continuous lagged met variable is expected to improve discrimination accuracy as a continuous variable expresses the degree of over- or undershoot. However, experimental results (Figure 8) show inferior prediction accuracy to using a binary term for most years. In particular, the prediction rate in 1995 falls below that without using any lagged-met indicator.

This is counter-intuitive since Fisher discrimination tends to work better for continuous data. In addition, the Fisher coefficients exhibit increasingly large variation. The deterioration in predictive performance might be caused by noise disrupting the discrimination process. Highly skewed numerical lagged-met indicators might also affect discrimination performance.

4.4.3 Stepwise Discriminant Analysis

Predicting next year's success in meeting an EPS forecast may require data earlier than the current year if information is not initially reflected in the indicator measures. Such dynamics can be detected using stepwise discriminant analysis, which optimally selects independent variables and their lags. Table 14 shows that the one-year lagged met indicator plays a dominant role in the discrimination process in almost all years whereas other measures do not feature consistently, supporting the low predictive power mentioned in subsection 4.3. Overall, indicators are most significant when lagged by one year (in nearly all) cases justifying the lag chosen initially. The selection of 7-year lagged CR and 6-year past Met for 2001 is likely due to spurious correlations since a cash realization from six-years ago is unlikely to be significant.

Year	**Selected predictors**
1993	Met(-1)
1994	Met(-1), Met(-2), CR(-2), RoCE(-2)
1995	No variables entered
1996	Met(-1), PARR(-1)
1997	Met(-1), *RoCE(-4)*, RoCE(-1)
1998	Met(-1), Met(-2)
1999	TR(-2), Met(-2), PARR(-1)
2000	*Met(-4)*, Met(-2)
2001	Met(-2), *Met(-6)*, *CR(-7)*, CR(-4)

Table 14: Stepwise Fisher discriminant analysis with (selected) predictors arranged in order of significance of Wilks' $\mathbf{\Lambda}$ criterion. Italicized predictors are believed to be spurious.

4.5 Logistic Regression

The inclusion of a *binary* lagged met indicator violates the assumption of continuously distributed variables in Fisher discriminant analysis. Logistic regression permits categorical independent variables, thus avoiding the problem. As with Fisher discriminant analysis, logistic regression is performed on the four indicator measures initially, with a lagged met indicator added at a later stage.

We aim to find weights β such that:

$$\ln\left(\frac{p_t}{1-p_t}\right) = \beta_{0,t} + \beta_{1,t}RoCE_{t-1} + \beta_{2,t}CR_{t-1} + \beta_{3,t}PARR_{t-1} + \beta_{4,t}TR_{t-1} + \varepsilon_t \quad (19)$$

where p is the probability of meeting EPS. The cut-off point is set at 0.5, so samples with $p \geq 0.5$ are predicted to meet their EPS forecast and samples with $p < 0.5$ do not.

The results are highly biased – in most cases, there is a high percentage of *either* true positives *or* false negatives – although predictive accuracy is higher than that obtained with Fisher discriminant analysis. For example in 2000, all the indicator measures become insignificant and all companies are predicted to not meet the forecast. This is caused in part by 76% of the companies missing targets as a consequence of the economic downturn (Table 15). Since logistic regression is very sensitive to violation of the equal prior assumption, a proxy for macroeconomic conditions should be included. Alternatively, the prior probability could be altered by assigning a higher cut off when market conditions are good and vice versa.

4.5.1 Logistic Regression with Lagged Met Indicator

When the lagged met indicator is included, the analysis becomes:

$$\begin{aligned}\ln\left(\frac{p_t}{1-p_t}\right) &= \beta_{0,t} + \beta_{1,t}RoCE_{t-1} + \beta_{2,t}CR_{t-1} + \\ &\quad \beta_{3,t}PARR_{t-1} + \beta_{4,t}TR_{t-1} + \beta_{5,t}Met_{t-1} + \varepsilon_{t-1} \quad (20)\end{aligned}$$

The inclusion of the lagged met indicator produces a better fit[3] as confirmed by the increase in R^2 noted in Table 26 in Appendix D. As in Fisher discriminant analysis, the lagged met variable is assigned a large weight in the equation for all years except 1995 where PARR emerges as the most important predictor. Following the inclusion of the lagged met variable, the prediction accuracies are more balanced (Table 15). Despite the improvement, this method

[3]The comparison was done using adjusted R^2, which takes into account the number of explanatory variables.

still lacks the robustness of Fisher discriminant analysis as illustrated by the large fluctuation in the ratio of correct prediction for the met and not met groups in Table 15. Overall, the prediction accuracy obtained through logistic regression is slightly better than Fisher discriminant analysis in many instances, but the tendency to produce 'one-sided biased' results makes it a less desirable method.

Met to Not Met Ratio	1993	1994	1995	1996	1997	1998	1999	2000	2001
Number of Companies	0.82	0.82	1.26	0.77	0.90	0.62	1.02	0.86	0.31
Accuracy (4 var)	0.66	0.40	2.46	0.66	0.87	0.09	1.32	0.34	0.00
Accuracy (5 var)	0.97	0.98	2.30	0.88	0.85	0.55	0.96	0.62	0.00

Table 15: Comparison of the performance of logistic regression with the variation of the groups' sample sizes using the four indicator measures (4 var) and with an additional lagged consensus EPS meeting term (5 var).

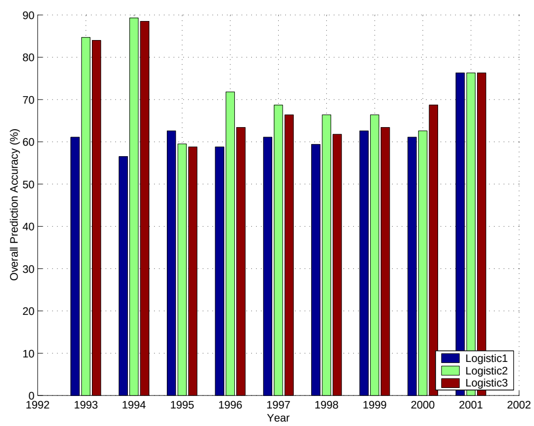

Figure 9: Overall accuracies when using logistic regression (Logistic1: Discrimination using the four indicator measures; Logistic2: With added binary lagged-met indicator; Logistic3: With the continuous lagged-met indicator)

The decline in predictive accuracy over the years is highlighted in Figure 9. Accuracies above 80% are observed in the first two years because of highly separable data as illustrated in Table 16. The overlap in the data for the later years results in poorer discrimination, but usually still over 60%.

4.6 Principal Component Analysis

Principal component analysis (PCA) is often used to reduce data dimensionality thereby improving computational and storage efficiency. However, this is not a major concern in our

```
Frequency | Distribution for 1993
       16 |
          |
          |    01
          |    00                                                 1
       12 |    00                                                 1
          |   000                                                 1
          |   000                                                 1
          |   000                                                11
       08 |   000                                                11
          |   000                                                11
          |   000                                                11
          | 0 0001                                             1 111
       04 | 000000                                             11111  1
          | 0000000       1                                   111111 11
          | 0000000000 0000 10 0    1                11  01111010100010
          | 0000000000 0000 10 0    1                11  01111010100010
     Prob | 0            .25           .5              .75            1
    Group | 000000000000000000000000000000111111111111111111111111111111
```

```
Frequency | Distribution for 2001
       16 |
          |
          |
          |
       12 |
          |
          |
          |                  1
       08 |                  0
          |                  0          1 1
          |                  011        111    1
          |                11010   11   1111   1
       04 |               111000   1111111001 11      1
          |        0      000000   1011000001 1111    0
          |        0   11 00000010 001000000100111   10
          |        000000000000000000000000000001011110 0   1
     Prob | 0            .25           .5              .75            1
    Group | 000000000000000000000000000000111111111111111111111111111111
```

Table 16: Observed Groups and Predicted Probabilities for 1993 (top) and 2001 (bottom).

analysis given the limited number of indicator measures; four, or five if the lagged met indicator is included. PCA can also be applied to decorrelate variables, which may be advantageous as some discrimination methods perform better with uncorrelated variables.

The transformation matrices for 1993 to 2001 are shown in Table 17. The near-diagonal transformation matrices imply low correlation in the original data space. Hence, we only anticipate a slight gain from applying PCA before running the discriminant analysis. The first two components capture more than 90% of the variance as indicated in Table 18. It is not surprising that tax rate turns out as the most prominent component since it has the largest variance amongst all the indicator measures (subsection 4.2). This is followed by RoCE, CR and PARR in all years but 1993 with PARR leads CR.

Fisher discriminant analysis and logistic regression have been carried out on the orthogonalized components. This shows that prediction accuracy deteriorates for the majority of years although there are marginal improvements over non-orthogonalized data in 1994 and 1998. The rather substantial deterioration might be due to loss of information content when orthogonalization was performed, as PCA is an unsupervised transformation that disregards the met and not met groups. Unfortunately, PCA stabilizes neither the Fisher nor regression coefficients and we conclude that it is not worth using this method to pre-process the data.

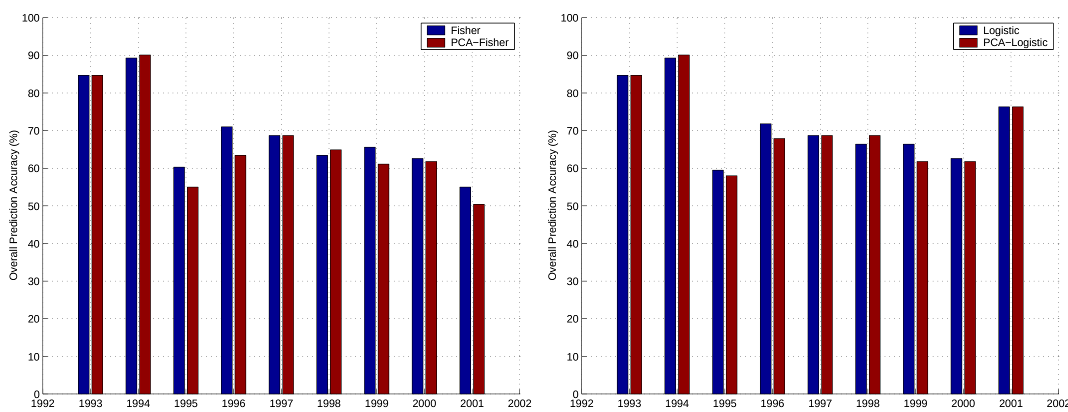

Figure 10: Fisher discriminant analysis (left) and logistic regression (right) on the transformed components

Year	Component	TR	RoCE	CR	PARR
1992	1	**0.955**	0.293	-0.029	-0.024
	2	-0.294	**0.950**	-0.107	0.006
	3	0.006	-0.110	**-0.990**	0.083
	4	-0.025	-0.011	-0.082	**-0.996**
1993	1	**0.990**	0.127	0.065	-0.007
	2	-0.118	**0.985**	-0.119	-0.040
	3	-0.019	-0.017	0.209	**-0.978**
	4	-0.078	0.117	**0.968**	0.206
1994	1	**0.983**	0.183	0.015	-0.000
	2	-0.183	**0.983**	-0.033	0.007
	3	0.021	-0.030	**-0.974**	0.223
	4	-0.003	-0.000	0.223	**0.975**
1995	1	**0.985**	0.172	0.007	0.005
	2	0.172	**-0.985**	0.033	-0.006
	3	0.012	-0.032	**-0.999**	0.039
	4	0.005	0.006	-0.039	**-0.999**
1996	1	**0.974**	0.225	-0.012	0.007
	2	0.226	**-0.974**	0.018	-0.014
	3	-0.008	-0.020	**-0.999**	-0.049
	4	0.004	0.016	0.049	**-0.999**
1997	1	**1.000**	0.023	0.005	0.001
	2	0.023	**-1.000**	0.014	-0.005
	3	0.005	-0.014	**-1.000**	-0.010
	4	0.001	0.005	0.009	**-1.000**
1998	1	**1.000**	0.013	0.007	-0.001
	2	-0.012	**1.000**	-0.023	-0.005
	3	0.007	-0.023	**-0.995**	0.097
	4	0.001	0.007	0.097	**0.995**
1999	1	**1.000**	-0.027	-0.003	-0.002
	2	-0.027	**-0.999**	0.009	0.019
	3	-0.002	-0.001	**-0.926**	0.377
	4	0.003	0.021	0.377	**0.926**
2000	1	**0.999**	-0.041	0.016	-0.001
	2	-0.042	**-0.999**	0.033	-0.012
	3	0.015	-0.034	**-0.999**	0.000
	4	-0.001	0.012	-0.000	**-1.000**
2001	1	**1.000**	-0.009	-0.003	-0.002
	2	-0.009	**-1.000**	0.013	-0.001
	3	-0.003	-0.013	**-0.999**	0.044
	4	-0.002	0.001	-0.044	**-0.999**

Note the deviation from the trend in 1993, where the third and fourth components are dominated by PARR and CR respectively

Table 17: PCA Transformation coefficients.

	Percentage variance				Cumulative percentage			
Year	**1st**	**2nd**	**3rd**	**4th**	**1st**	**2nd**	**3rd**	**4th**
1992	67.10	22.47	7.00	3.43	67.10	89.57	96.57	100.00
1993	73.51	20.18	3.81	2.50	73.51	93.70	97.50	100.00
1994	75.96	22.99	0.62	0.43	75.96	98.95	99.57	100.00
1995	74.89	24.58	0.42	0.11	74.89	99.48	99.89	100.00
1996	76.79	22.65	0.42	0.14	76.79	99.44	99.86	100.00
1997	88.46	11.37	0.13	0.05	88.46	99.83	99.95	100.00
1998	85.78	13.92	0.24	0.06	85.78	99.70	99.94	100.00
1999	91.31	8.41	0.17	0.10	91.31	99.73	99.90	100.00
2000	87.09	12.52	0.35	0.04	87.09	99.61	99.96	100.00
2001	91.05	8.53	0.39	0.02	91.05	99.58	99.98	100.00

Table 18: Percentage importance of each orthogonal component.

4.7 Ranked Data Analysis

Lachenbuch *et al.* [7] found rank discrimination to be more effective for non-normal variables with small differences in their means. The indicator measures appear to have these characteristics. We thus extend the investigation to using rank data for both Fisher discrimination analysis and logistic regression, categorizing the companies into quintiles according to the yearly ranking in each of the indicator measures, with one assigned to the top 20%. A marginal improvement in predictive accuracy using Fisher discriminant analysis is found in only three of the ten years with a substantial deterioration in the other years, which may be explained by the analysis method's requirement for continuous data. A similar decline when using logistic regression suggests data loss as another likely cause; categorization of the companies is similar to discretization. These results suggest that ranked data is not suitable for this analysis, see figure 11.

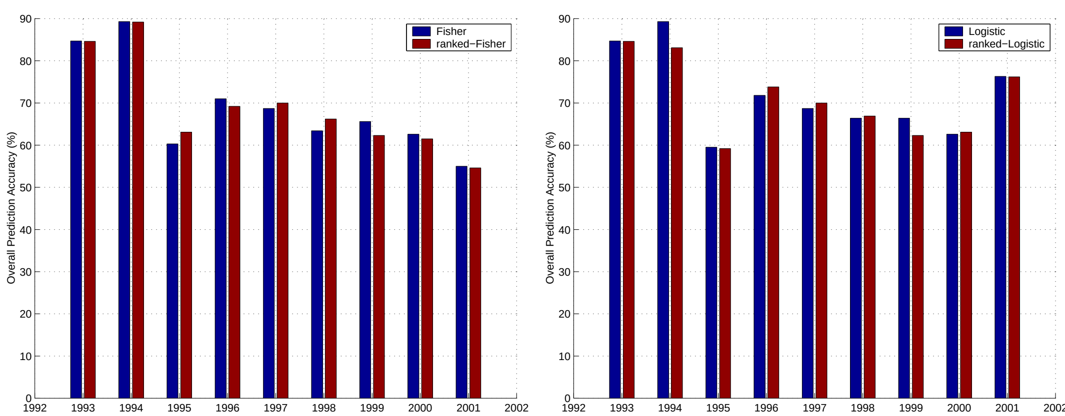

Figure 11: Fisher discriminant analysis (left) and logistic regression (right) using ranked-data.

4.8 Evidence of Market Sentiments

Companies from various industries react differently to fluctuations in market conditions. Analyzing the data based on sectors allows the flexibility to encode sector-specific characteristics into the discrimination process. The companies have been categorized into seven S&P industry sectors, namely Consumer Staples, Consumer Discretionary, Materials, Industrials, Information Technology, Communications, Health Care and Energy. Since there is insufficient yearly sector-based data for a statistically significant analysis, the experiments have been carried out on data grouped into three distinct phases according to the following market conditions and sentiments:

1992-1995 Bull market, emphasis on sustainable growth

1996-1999 Bull market, emphasis on high earnings per share and creation of a bubble

2000-2001 Bear market, refocussing on sustainability after the bubble burst.

Figure 12 shows large variations across different sectors and phases and, consequently, we expect the above refinement to give greater predictive power as shown in Figure 13. However, it is important to note that sector-based analysis may suffer from spurious fitting and unstable prediction because the sample is too small. Table 25 and Table 28 in the appendix illustrate that larger sectors have lower discrimination significance – these are given by Wilk's Λ in the case of Fisher discriminant analysis and R^2 for logistic regression. Therefore, any firm conclusions from sector-based analysis await further work with a larger sample set. Prediction accuracy is lower in the period 1996-9 than in the period 1993-5, which may be explained by the shift to valuing companies on the basis of their short-term growth rather than on traditional measures of sustainable growth. Even though investors returned to sustainability indicators after the bubble burst we do not observe an increase in prediction accuracy. This is most likely due to the application of consensus EPS predictions which have been made during the bubble – thus geared for high growth – but the economy entering a recession in the later phase.

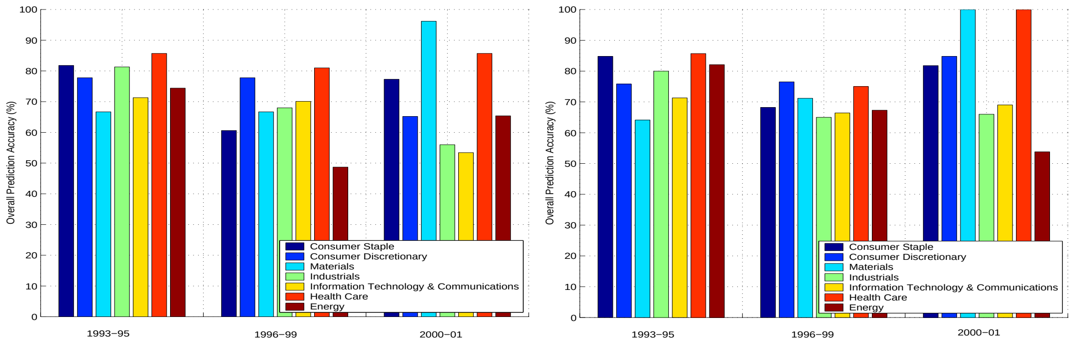

Figure 12: Sector-based Fisher discriminant analysis (left) and logistic regression (right)

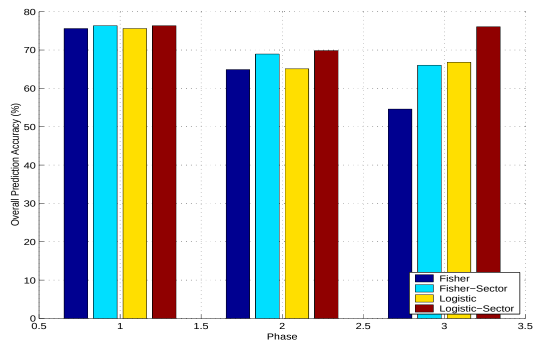

Figure 13: Prediction accuracies using the full cross-sectional data and sector-based data

4.9 The Bear Market

Many companies are left out of the sample because they are not part of the S&P index over the entire sample period. Removal of a company does not necessarily indicate bankruptcy or malperformance. The S&P website[4] states that their indices "*are designed to differentiate between fast growing companies and slower growing or undervalued companies.*" Standard & Poor's and Barra cooperated to develop a Price to Book value calculation that equally divides market capitalization between growth and value. The indices are rebalanced twice per year. Hence, the cross-section is too small to be split into sectors while analyzing on a yearly basis for the 10 years. A solution is to use a shorter period. We have chosen the two most recent years (2000 and 2001), increasing the number of companies from 133 to 366, because of the change to a bear market.

Tables 19 and 20 summarize the results obtained *before* and *after* expansion of the data set where the full cross-sectional results for the expanded data set are consistent with the previous

[4] `http://www.standardandpoors.com`

results. The slight deterioration in prediction accuracy is mainly due to the increased sample size limiting the likeliness and effect of over-fitting which might have occurred in the smaller data set. This problem is revealed clearly in the sector-based analysis for the unexpanded data where small sectors recorded abnormally high prediction accuracy.

Year	Sector	**Unexpanded data set 131 companies**				**Expanded data set 366 companies**			
		N	**Met (%)**	**Not Met(%)**	**Overall (%)**	**N**	**Met (%)**	**Not Met(%)**	**Overall (%)**
2000	Staples	11	66.7	80.0	72.7	31	60.0	68.4	64.7
	Discretionary	33	62.5	60.0	60.6	73	51.9	73.2	66.3
	Materials	13	100.0	100.0	100.0	30	66.7	69.6	68.8
	Industrials	25	76.9	75.0	76.0	59	55.6	64.7	60.7
	IT & Telcos	29	75.0	82.4	79.3	77	50.0	57.1	53.4
	Health Care	7	100.0	100.0	100.0	36	69.6	81.0	75.0
	Energy	13	90.0	100.0	92.3	26	85.7	60.0	80.8
	Full cross section	131	58.1	54.0	55.0	366	55.4	49.8	51.6
2001	Staples	11	60.0	83.3	72.7	31	71.4	70.0	70.6
	Discretionary	33	100.0	82.1	84.8	73	52.4	69.4	65.1
	Materials	13	N/A	N/A	N/A	30	4.1	60.0	71.9
	Industrials	25	80.0	65.5	68.0	59	45.5	58.0	55.7
	IT & Telcos	29	100.0	73.9	79.3	77	50.0	62.3	58.9
	Health Care	7	80.0	100.0	85.7	36	71.4	73.9	72.7
	Energy	13	80.0	87.5	84.6	26	92.3	69.2	80.8
	Full cross section	131	66.1	60.0	62.6	366	49.5	58.7	54.3

N: Sample size N/A: Discrimination failed

Table 19: Fisher discriminant results before and after expanding data set.

5 Conclusions and Future Work

This report investigated the use of four indicator measures, RoCE, CR, PARR and TR, to differentiate between companies meeting and not meeting consensus EPS figures. Several dependence statistical methods were employed to test the significance of the ratios in achieving this objective, namely Student's t, a measure of distance between the means of two distributions, F, the difference in variances, and Kolmogorov-Smirnov for the overall distributions. Both Student's t and KS tests produced non-significant results suggesting likely difficulty in discriminating between the two groups. The F-test, however, showed large differences between the variances for CR and TR, suggesting that quadratic discriminant analysis could work well on CR and TR. This was not attempted because of lack of data, with only 10 observations in the time series for each company. Linear discriminant analyses and logistic regressions were

Year	Sector	Unexpanded data set 131 companies				Expanded data set 366 companies			
		N	Met (%)	Not Met(%)	Overall (%)	N	Met (%)	Not Met(%)	Overall (%)
2000	Staples	11	66.7	60.0	63.6	31	53.3	73.7	64.7
	Discretionary	33	25.0	100.0	81.8	73	29.6	96.4	74.4
	Materials	13	N/A	N/A	N/A	30	11.1	95.7	71.9
	Industrials	25	76.9	75.0	76.0	59	59.3	73.5	67.2
	IT & Telcos	29	75.0	88.2	82.8	77	63.2	37.1	50.7
	Health Care	7	81.0	69.6	75.0	36	82.6	81.0	81.8
	Energy	13	100.0	100.0	100.0	26	90.5	80.0	88.5
	Full cross section	131	46.4	74.7	62.6	366	68.4	57.0	62.5
2001	Staples	11	60.0	83.3	72.7	31	71.4	80.0	76.5
	Discretionary	33	100.0	100.0	100.0	73	14.2	98.4	77.1
	Materials	13	N/A	N/A	N/A	30	40.0	100.0	90.6
	Industrials	25	40.0	90.0	80.0	59	0.0	100.0	82.0
	IT & Telcos	29	50.0	91.3	82.8	77	5.0	94.3	69.9
	Health Care	7	N/A	N/A	N/A	36	66.7	82.6	75.0
	Energy	13	80.0	87.5	84.6	26	69.2	69.2	69.2
	Full cross section	131	0.0	100.0	76.3	366	0.0	100.0	67.4

N: Sample size N/A: Regression failed

Table 20: Logistic regression results before and after expanding data set.

run, given the mean seperability found in RoCE and PARR. The indicators were found to be mostly non-normal, which was a problem for the regressions.

The failure to detect a separation between companies using statistical dependence test was reinforced further by low prediction accuracies in regressions when using the indicator measures. Predictive accuracy was only slightly over 50%. The two methods used were Fisher discriminant analysis and logistic regression, where the latter was found to be more robust and stable when groups had different sizes. Logistic regression using all four measures produced slightly better performance than Fisher discrimination in most cases, but suffered from prediction bias; prediction accuracy would be very high for one group and very low for the other. This was most apparent when there was a significant sample size mismatch between the two groups, hence Fisher discriminant analysis was preferred. RoCE displayed the highest discrimination power amongst the four indicator measures, as expected. However, only marginal differences in prediction rate were observed for each of the indicator measures. Hence, the ratios were taken to be equally important in predicting the success of achieving EPS projections.

Inclusion of a lagged-met term in its binary and numerical forms was found to improve the results significantly, although the former performed best. It seems probable that a numerical lagged-met term introduces noise affecting the overall discrimination process. However, the lagged-met term was unable to stabilize the coefficients sufficiently to reliably predict the meeting of consensus EPS estimates for the next year. Further refinements were implemented to improve the above techniques. Firstly, sector-based analysis was conducted according to three different market phases. The companies were categorized into their specific industries, with classification results being obtained using sector-based discrimination, since this captured the heterogeneity across firms. Further, continuous metrics were replaced with ranked data, which was expected to perform better for non-normally distributed data. In fact it performed worse. It is likely that the ranking process, in effect, discretized the variables, unintentionally eliminating some valuable information. Finally PCA was applied to decorrelate the dependent variables. The PCA transformation coefficients were found to be close to diagonal, suggesting a lack of correlation in the original variables. In this case, decorrelation might result in loss of information since the entire data set was transformed without differentiating between the met and not met classes. Thus, the large overlap between met and not met groups was retained even though PCA had been applied to separate and sort the components in order of overall variance. Stepwise discriminant analysis was applied to find the set of variables most influential in the prediction process. Unsurprisingly the best variables for predicting the following year's EPS came from the current year's indicators, suggesting that long term dependencies are negligible.

Looking at the analysis from a broader perspective we can conclude that the four indicator measures have limited discriminating capability in predicting whether a company will, or will not meet, the EPS expectation. The low predictive power suggests that there exist other economic and financial factors that either directly or indirectly contribute to the process. For example, conventional accounting ratios such as liquidity[5] and gearing[6] might play important roles in this respect. In view of the inconclusive results from these experiments, where classification was possible for the early years (1993-95) with a subsequent deterioration in accuracy, many refinements and extensions are possible to further explore the predictive power of these indicator measures. Linear discrimination may not be suitable for our investigation since the separation of the means of the combined indicator measures was found to be insignificant in MANOVA tests.

[5]Usually measured by Current Ratio $= \frac{\text{Current Assets}}{\text{Current Liabilities}}$.

[6]Computed as $\frac{\text{Total Debt}}{\text{Total Equity}}$.

Given more data, quadratic discrimination methods might be better, as differences in variance between the two groups were found during the classification process. This is especially so since large differences in variance were observed for the CR and TR indicators. Finally, given more data – either in terms of years or frequency – we could investigate whether co-integration would address the inter-temporal instability problems, since all the techniques introduced so far have failed to handle this issue. More complex models may be required to fully capture the variability across different sectors and time periods.

Overall, we do not recommend that these indicators be used further since longer data series are not likely to change the problem of low predictability. Rather, basic accounting data should be considered and grouped to find measures more indicative of earnings per share.

References

[1] Baker, R. (2001). *Determining Value: Valuation Models and Financial Statements.* Financial Times Prentice Hall.

[2] Bernstein, L. A. and J. G. Siegel (1979). The Concept of Earnings Quality. *Financial Analyst Journal* **35** 72–75.

[3] Bernstein, L. A. and J. G. Siegel (1982). Quality of Earnings Concept- A survey. *Financial Analyst Journal* **38**.

[4] Brown, P. R. (1999). Earnings Management: A Subtle (and Troublesome) Twist to Earnings Quality. *Journal Of Financial Statement Analysis* **4** 61–63.

[5] Deloitte and Touche (2002). Quality of Earnings. *Integrity & Quality* .

[6] Futrelle, D. (2002). Merrill: The new earnings cops? *CNN/Money Contributing Columnist* .

[7] Lachenbuch, P. A., C. Sneeringer and L. T. Revo (1973). Robustness of the Linear and Quadratic Discriminant Functions to Certain Types of Non-Normality. *Comm. Statistic* **1** 39–56.

[8] Lev, B. and S. R. Thiagarajan (1993). Fundamental Information Analysis. *Journal of Accounting Research* **31** (2).

[9] Merrill Lynch (2002). Quality of Earnings: Toward a 360^o View of Reality. *Merrill Lynch Special Report* .

[10] O'Glove, T. (1987). *Quality of Earnings.* The Free Press.

[11] Papoulis, A. (1965). *Probability, Random Variables and Stochastic Processes.* McGraw-Hill Book Company.

[12] Patel, S. A. and M. Santicchia (2003). Standard & Poors Earnings and Dividend Ranking System: Portfolio Performance, Risk, and Fundamental Analysis. *Standard & Poors* .

[13] Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery (1992). *Numerical Recipes in C*, 2nd edition. Cambridge University Press.

[14] SEC. Securities and Exchange Commission, Plaintiff, v. Xerox Corporation, Defendant. *United States District Court for the Southern District of New York.* Civil Action No. 02-272789 (DLC).

[15] Sharma, S. (1996). *Applied Multivariate Techniques.* John Wiley & Sons, Inc.

[16] Tabachnick, B. G. and S. L. Fidell (2001). *Using Multivariate Statistics*, 4th edition. Allyn & Bacon.

A Table of Abbreviations

ASR	Accounting Series Release
CAPM	Capital Asset Pricing Model
CR	Cash Realization
EMH	Efficient Market Hypothesis
EPS	Earnings per Share
FASB	Financial Accounting Standard Board
FD	Full Disclosure
GAAP	Generally Accepted Accounting Principles
KS	Kolmogrov-Smirnov
MANOVA	Multivariate Analysis of Variance
PARR	Productive Asset Reinvestment Ratio
PCA	Principal Component Analysis
QEI	Quality of Earnings Index
RoCE	Return on Capital Employed
R&D	Research and Development
SEC	Securities Exchange Commission
S&P	Standard and Poor's
TR	Tax Rate

B Data Preparation

Source/Percentage of outliers	RoCE	CR	PARR	TR	Overall
Merrill Lynch	5.1%	2.8%	4.1%	3.0%	3.7%
CFR	8.6%	2.8%	4.0%	2.1%	4.3%

Table 21: Comparison of standard deviations of the Merrill Lynch and locally produced datasets.

t	RoCE			CR		
	EPS_t	EPS_{t+1y}	CEPS	EPS_t	EPS_{t+1y}	CEPS
1992	0.335	0.257	0.042	0.004	-0.033	-0.062
1993	0.201	0.108	-0.102	-0.018	-0.067	0.052
1994	0.065	-0.097	-0.064	0.111	0.157	0.010
1995	-0.043	-0.108	-0.055	0.082	0.154	0.144
1996	-0.033	-0.045	-0.148	0.149	0.101	0.173
1997	-0.019	0.083	-0.113	0.011	-0.078	0.089
1998	0.104	0.027	-0.101	0.099	0.052	0.194
1999	0.054	-0.049	0.015	0.058	0.128	-0.053
2000	0.093	0.119	-0.050	-0.011	-0.126	0.060
2001	0.400		0.145	0.152		0.183

t	PARR			TR		
	EPS_t	EPS_{t+1y}	CEPS	EPS_t	EPS_{t+1y}	CEPS
1992	-0.121	-0.135	-0.092	0.341	0.256	-0.051
1993	-0.122	-0.159	-0.159	0.178	0.127	-0.051
1994	-0.176	-0.300	-0.152	0.135	0.052	-0.028
1995	-0.269	-0.322	-0.185	-0.030	-0.032	0.077
1996	-0.257	-0.274	-0.281	-0.020	-0.022	0.047
1997	-0.164	-0.219	-0.161	-0.081	-0.071	-0.006
1998	-0.141	-0.192	-0.150	0.015	0.007	0.048
1999	-0.044	0.016	-0.082	0.190	0.053	0.121
2000	0.074	-0.127	-0.059	0.045	0.140	-0.017
2001	0.152		0.175	0.404		0.200

Table 22: Correlation coefficients between Consensus & Pro-Forma EPS and the four indicator measures on a yearly basis

C Distribution of the Indicator Measures

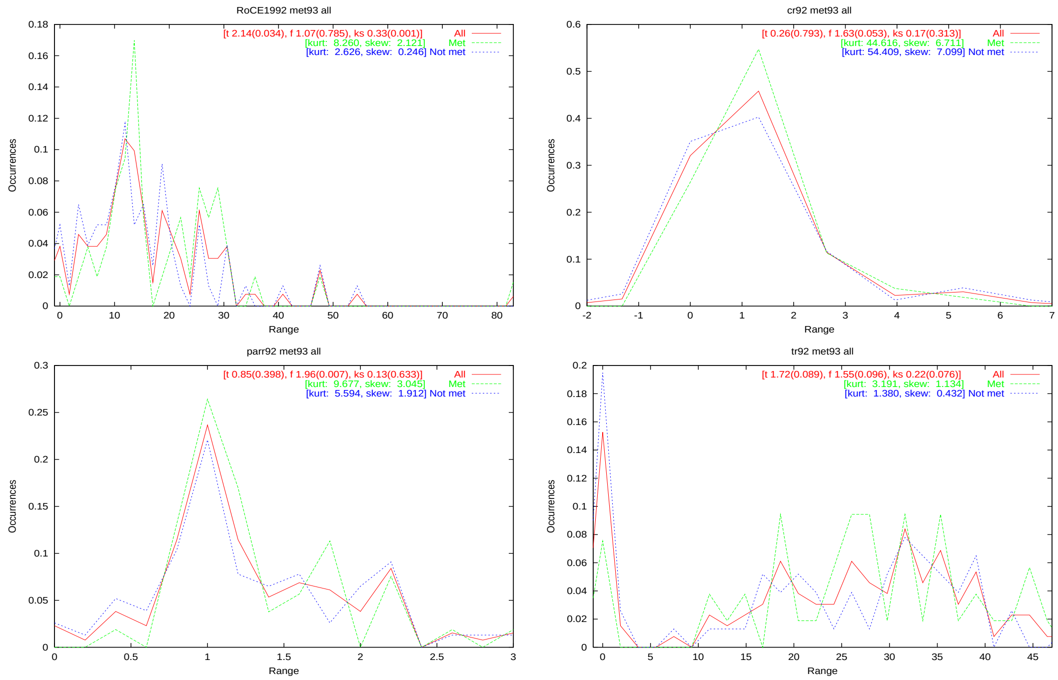

Figure 14: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1993.

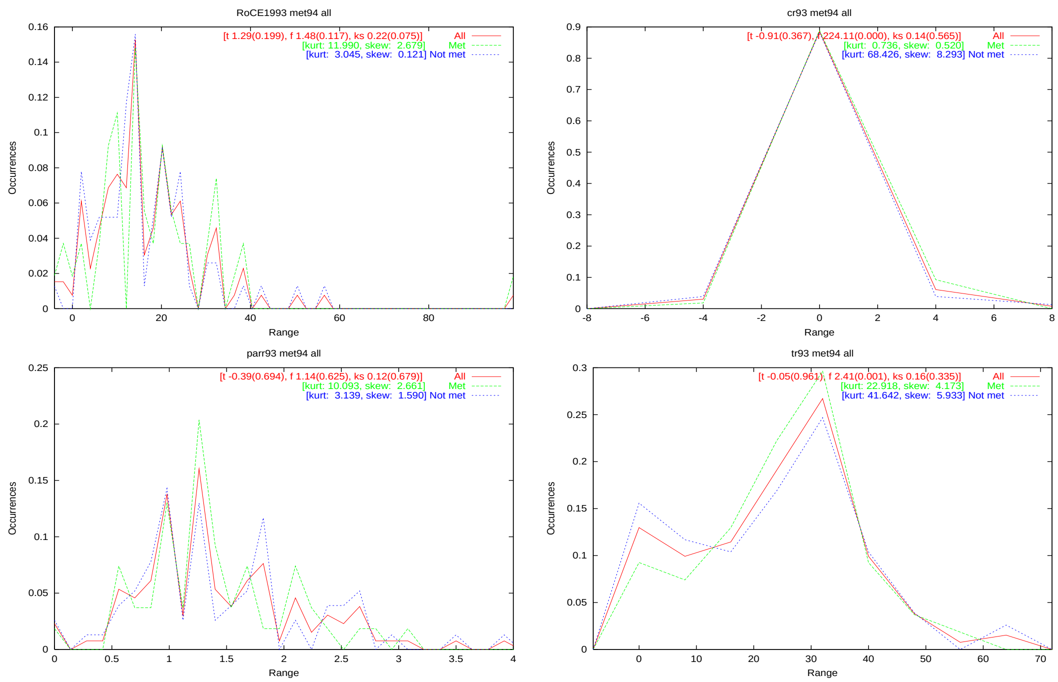

Figure 15: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1994.

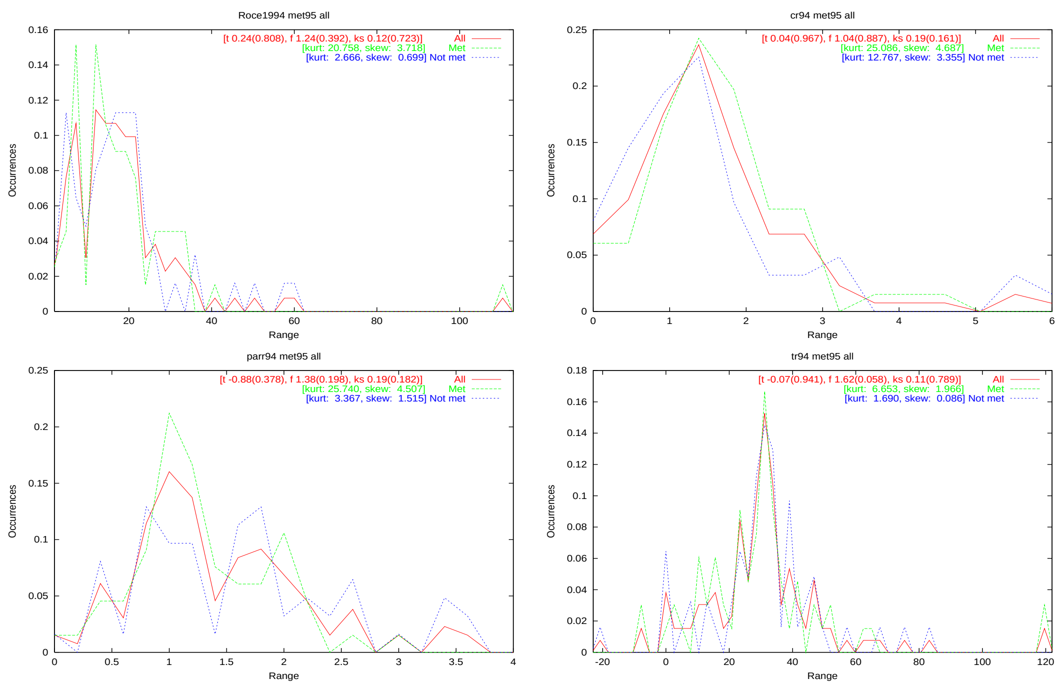

Figure 16: Comparison of distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1995.

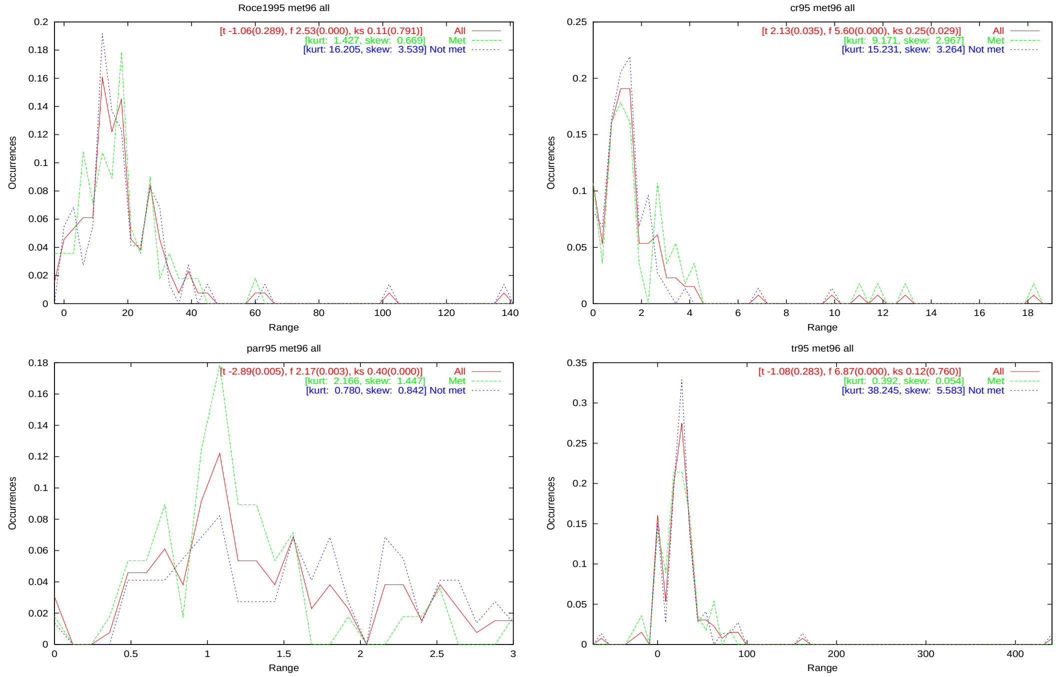

Figure 17: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1996.

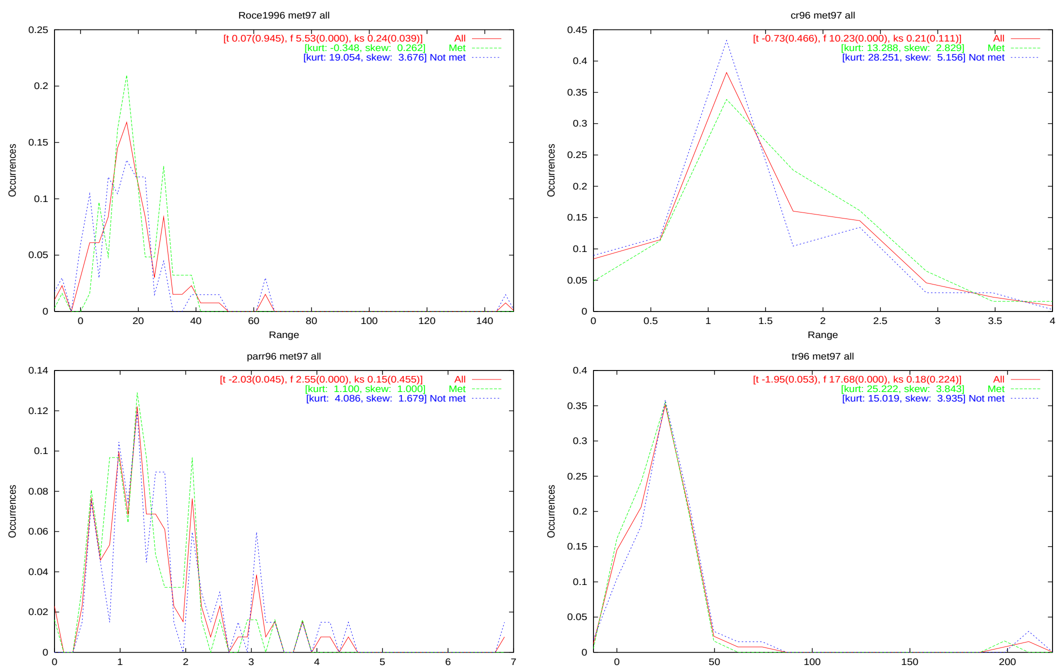

Figure 18: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1997.

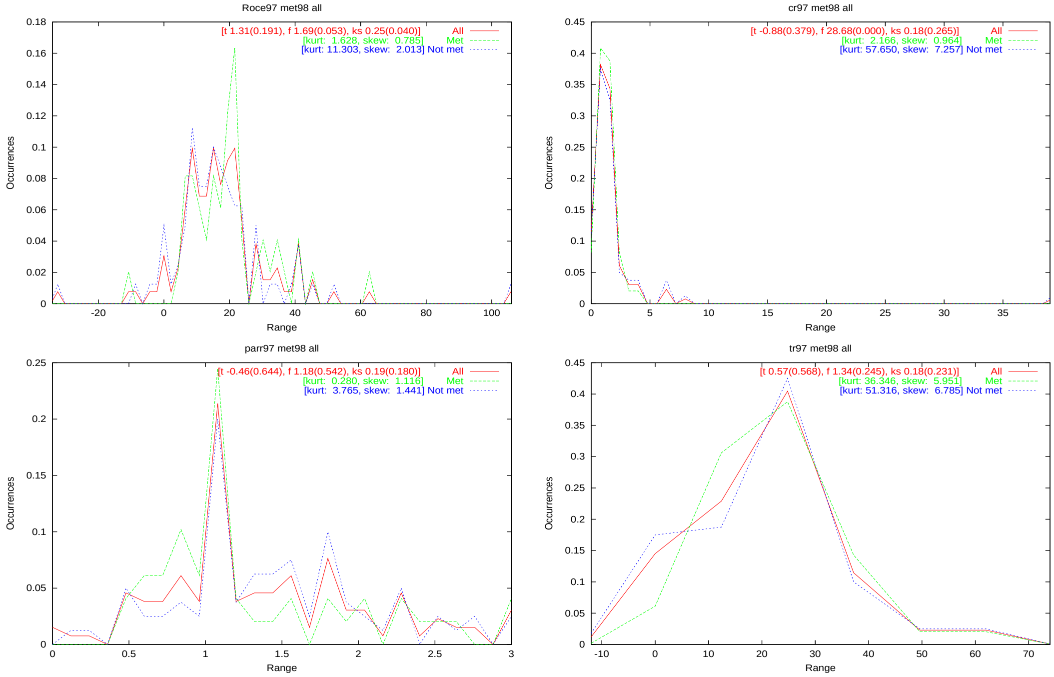

Figure 19: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1998.

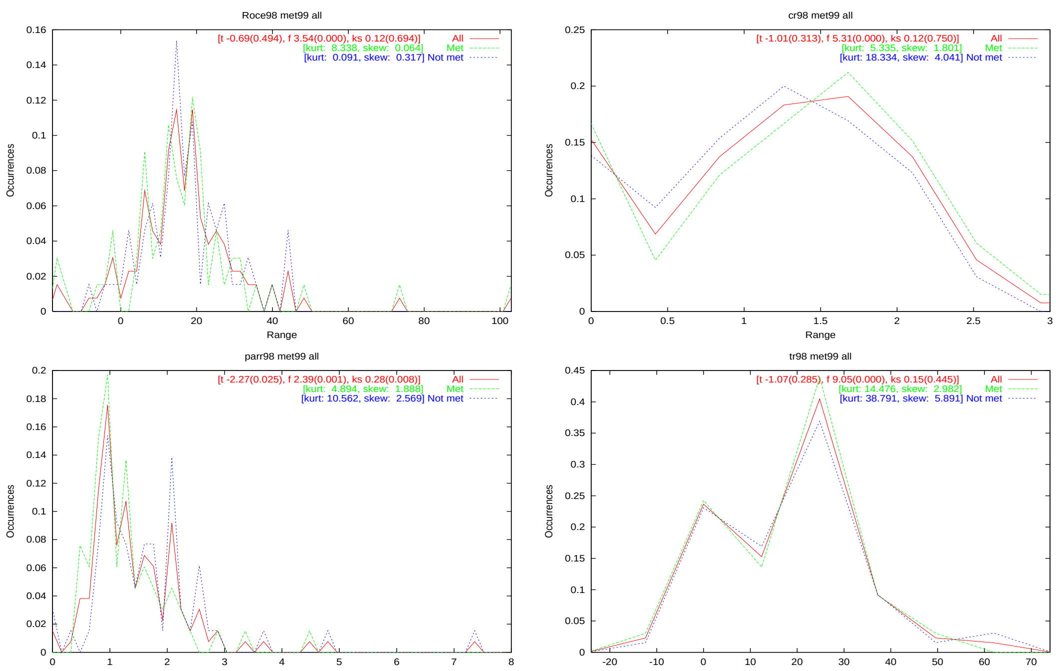

Figure 20: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 1999.

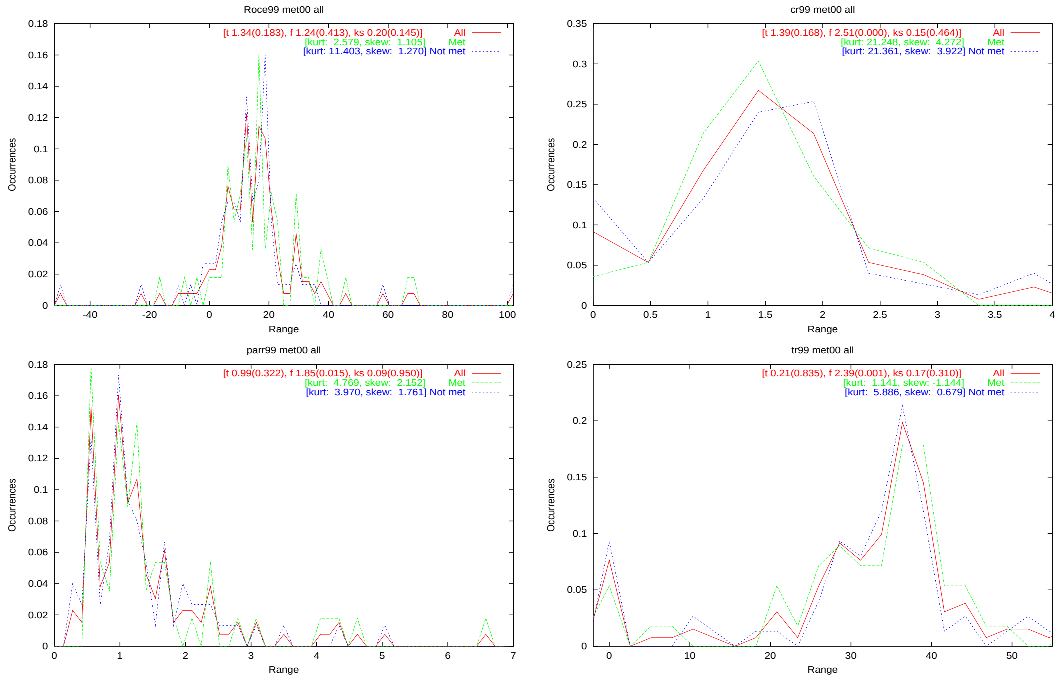

Figure 21: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) for 2000.

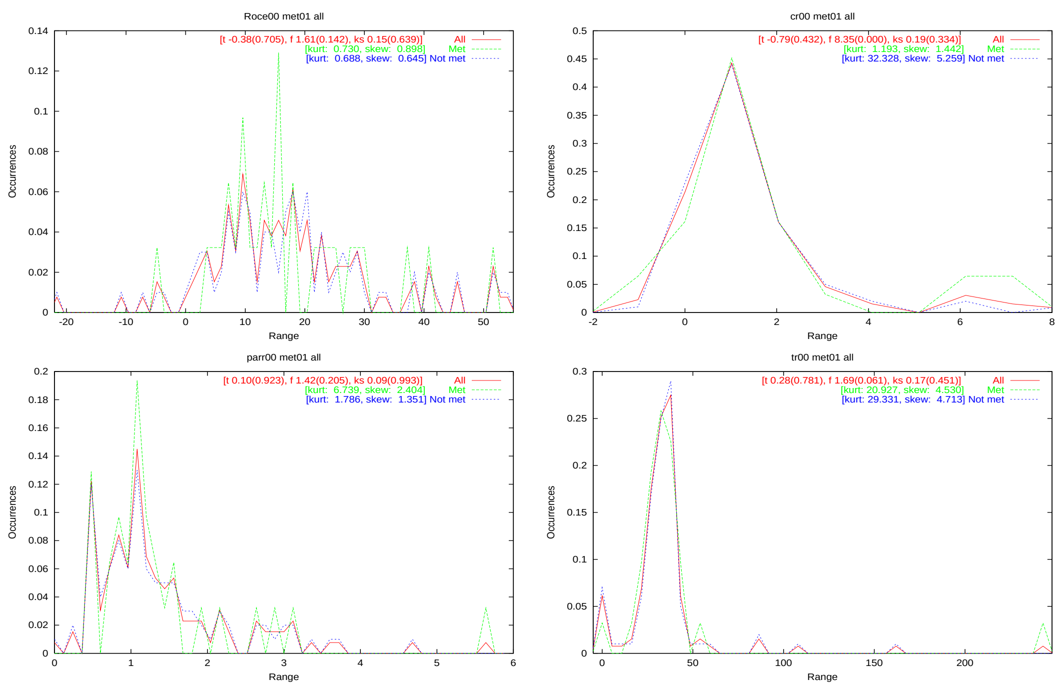

Figure 22: Comparison of the distributions of RoCE (top left), CR (top right), PARR (bottom left) and TR (bottom right) histograms for 2001.

D Tables of Results

Index

Year	Met (%)	Not Met (%)	Overall (%)	Wilk's Λ	Fisher Coefficients $(t-1)$ RoCE	CR	PARR	TR	Met	Constant	Standardized Coefficients[a] RoCE	CR	PARR	TR	Met
1993 to 2001	6.1	55.6	60.3	0.947	0.066	-0.070	-0.203	0.028		-1.299	0.736	-0.200	-0.198	0.482	
	9.3	48.6	53.4	0.971	-0.016	0.198	0.562	-0.030		-0.122	-0.177	0.743	0.472	-0.491	
	67.1	46.6	58.0	0.967	0.030	0.057	1.063	0.013		-2.743	0.317	0.109	0.898	0.211	
	70.2	59.5	64.1	***0.885***	-0.009	-0.352	1.008	0.013		-1.187	-0.116	-0.526	0.827	0.446	
	6.1	50.7	58.0	*0.925*	-0.035	-0.225	0.859	0.016		-0.879	-0.405	-0.314	0.715	0.683	
	60.0	56.8	58.0	0.969	0.069	-0.270	-0.524	0.006		-0.182	0.835	-0.392	-0.392	0.285	
	71.2	55.4	63.4	0.936	0.019	0.047	1.191	0.013		-2.580	0.250	0.067	0.839	0.566	
	60.7	58.7	59.5	0.950	0.069	0.455	0.234	-0.021		-1.702	0.840	0.786	0.189	-0.309	
	51.6	49.0	49.6	0.995	0.053	0.294	0.225	0.010		-2.354	0.686	0.885	0.176	0.185	
1993 to 2001	381	86.1	84.7	***0.511***	0.018	0.036	0.037	-0.006	2.760	-1.529	0.204	0.103	0.036	-0.094	0.999
	88.1	90.3	89.3	***0.371***	-0.017	-0.040	-0.081	0.000	3.283	-0.976	-0.194	-0.150	-0.068	0.000	1.022
	65.8	53.4	60.3	0.964	0.030	0.057	0.980	0.014	-0.603	-2.367	0.321	0.110	0.828	0.222	-0.301
	377	68.9	71.0	***0.836***	0.003	0.295	-0.707	-0.010	1.233	0.131	0.037	0.440	-0.580	-0.322	0.591
	62.9	73.9	68.7	***0.831***	0.014	0.105	-0.385	-0.009	1.709	-0.295	0.168	0.147	-0.320	-0.357	0.793
	6.0	61.7	63.4	*0.911*	0.034	-0.123	-0.244	0.005	1.713	-1.003	0.409	-0.179	-0.182	0.230	0.832
	6.6	67.7	65.6	***0.887***	0.032	-0.018	0.902	0.011	-1.516	-1.638	0.422	-0.026	0.635	0.504	-0.730
	6.1	60.0	62.6	*0.905*	0.044	0.268	0.299	-0.021	1.496	-1.771	0.528	0.462	0.242	-0.302	0.732
	58.1	54.0	55.0	0.989	0.015	0.210	0.177	0.003	1.636	-1.842	0.776	0.448	0.286	0.164	0.098
1993 to 2001	381	77.8	80.2	***0.686***	0.005	-0.051	0.044	-0.001	0.957	0.120	0.056	-0.146	0.043	-0.016	0.993
	977	88.9	84.7	***0.569***	-0.016	-0.037	-0.110	-0.006	1.183	0.794	-0.181	-0.140	-0.092	-0.129	1.032
	6.0	50.0	57.3	0.959	0.034	0.095	0.933	0.015	-0.426	-2.751	0.362	0.183	0.788	0.245	-0.471
	70.2	60.8	64.9	***0.883***	-0.009	-0.339	0.975	0.014	-0.366	-1.184	-0.114	-0.505	0.800	0.462	-0.125
	72.6	63.8	67.9	***0.848***	0.009	0.123	-0.643	-0.009	1.852	1.104	0.110	0.172	-0.534	-0.393	0.766
	70.0	51.9	58.8	0.948	0.044	-0.156	-0.344	0.007	1.739	-0.155	0.527	-0.226	-0.257	0.332	0.680
	71.2	55.4	63.4	0.936	0.021	0.040	1.191	0.013	-0.104	-2.616	0.271	0.058	0.839	0.565	-0.054
	0.6	57.3	62.6	***0.882***	0.040	0.217	0.117	-0.015	1.597	-0.700	0.484	0.374	0.095	-0.228	0.793
	61.3	54.0	55.7	0.993	0.032	0.243	0.104	0.003	0.660	-1.431	0.413	0.732	0.081	0.062	0.634

Table 23: Fisher discriminant analysis. (Top panel: Using the four indicator measures; Middle panel: With binary lagged-met indicator; Bottom panel: With numerical lagged-met indicator)

[a]The coefficients are standardized by its standard error

Year	Met (%)	Not Met(%)	Overall (%)	Wilk's Λ	Fisher Coefficients $(t-1)$ RoCE	CR	PARR	TR	Met	Constant	Standardized Coefficients RoCE	CR	PARR	TR	Met
1993 to 2001	60.0	55.6	61.5	0.949	0.650	0.046	-0.106	0.195		-2.362	0.904	0.066	-0.151	0.275	
	58.6	52.8	55.4	0.984	-0.466	0.416	0.399	-0.190		-0.466	-0.661	0.592	0.563	-0.271	
	59.7	67.2	63.1	0.960	-0.143	-0.398	0.508	0.361		-0.983	-0.204	-0.564	0.715	0.508	
	61.4	61.6	61.5	*0.917*	-0.059	-0.084	0.729	0.043		-1.879	-0.083	-0.119	0.988	0.061	
	6.1	67.6	66.9	***0.898***	0.545	0.324	-0.390	-0.442		-0.111	0.762	0.457	-0.545	-0.616	
	54.0	58.8	56.9	0.949	0.592	0.237	-0.557	0.095		-1.093	0.829	0.335	-0.781	0.134	
	60.6	60.9	60.8	0.935	0.194	-0.140	0.682	-0.016		-2.149	0.274	-0.198	0.936	-0.023	
	58.9	51.4	54.6	0.987	0.655	0.320	-0.173	0.113		-2.738	0.928	0.453	-0.245	0.159	
	58.1	61.6	60.8	0.992	0.527	0.002	-0.196	0.497		-2.490	0.744	0.002	-0.277	0.701	
1993 to 2001	82.8	681	84.6	***0.514***	-0.161	-0.082	0.055	0.083	2.718	-0.899	-0.223	-0.116	0.078	0.117	0.987
	87.9	90.3	89.2	***0.368***	0.123	0.167	0.019	0.029	3.342	-2.504	0.174	0.237	0.027	0.041	1.043
	62.5	63.8	63.1	0.958	-0.131	-0.392	0.475	0.354	0.497	-1.138	-0.186	-0.556	0.668	0.498	0.249
	75.4	64.4	69.2	***0.854***	0.055	-0.048	0.470	-0.044	1.471	-2.125	0.078	-0.068	0.645	-0.062	0.704
	6.1	375	70.0	***0.826***	-0.296	-0.175	0.191	0.244	1.520	-0.559	-0.415	-0.247	0.267	0.341	0.707
	68.0	65.0	66.2	***0.898***	-0.304	-0.150	0.340	-0.149	1.530	0.054	-0.426	-0.212	0.476	-0.210	0.744
	59.1	65.6	62.3	***0.892***	0.331	-0.116	0.511	0.019	1.462	-2.786	0.467	-0.164	0.701	0.027	0.706
	64.3	59.5	61.5	0.940	-0.193	-0.058	-0.052	0.008	1.899	-0.085	-0.273	-0.082	-0.073	0.012	0.930
	61.3	52.5	54.6	0.988	0.271	-0.041	-0.143	0.360	-1.248	-0.806	-0.787	0.532	0.527	-0.179	-0.056

Table 24: Fisher discriminant analysis using rank data. (Top panel: Without lagged-met indicator; Bottom panel: With binary lagged-met indicator)

Yr	Sec[a]	No.[b]	Met (%)	Not Met(%)	Overall (%)	Wilk's Λ	Fisher Coefficients $(t-1)$						Standardized Coefficients				
							RoCE	CR	PARR	TR	Met	Const	RoCE	CR	PARR	TR	Met
1993 to 1995	1	33	90.9	63.6	81.8	***0.629***	-0.017	-0.502	-0.437	-0.015	2.449	-0.831	-0.167	0.440	-0.246	-0.086	0.997
	2	99	80.0	75.9	77.8	***0.661***	0.011	0.114	-0.250	-0.021	2.322	-0.387	0.086	0.138	-0.245	-0.248	0.974
	3	39	57.9	75.0	66.7	***0.761***	0.078	0.041	0.550	-0.042	0.522	-0.926	0.625	0.261	0.544	-0.943	0.255
	4	75	71.8	91.7	81.3	***0.583***	-0.033	-0.036	0.243	0.017	2.720	-1.446	-0.270	-0.072	0.205	0.346	1.059
	5	87	60.0	80.9	71.3	***0.813***	0.015	-0.080	-0.066	-0.003	2.181	-0.801	0.176	-0.172	-0.051	-0.056	0.972
	6	21	71.4	92.9	85.7	***0.535***	0.002	0.234	-0.112	-0.020	2.838	-0.442	0.024	0.385	-0.123	-0.474	1.000
	7	39	73.7	75.0	74.4	***0.697***	0.004	-0.012	-0.573	0.014	1.646	-0.310	0.018	-0.052	-0.440	0.343	0.738
	All	363	71.2	79.7	75.6	***0.735***	0.002	-0.024	-0.131	-0.001	2.298	-0.774	0.017	-0.070	-0.116	-0.016	0.985
1996 to 1999	1	44	70.0	56.5	60.6	0.957	-0.014	0.749	-0.712	0.000	-0.075	0.407	-0.183	0.803	-0.416	0.003	-0.038
	2	132	73.8	80.7	77.8	***0.683***	0.042	0.343	-0.103	0.006	2.047	-2.240	0.297	0.306	-0.091	0.134	0.868
	3	52	62.5	67.7	66.7	0.900	0.052	-0.122	-0.590	0.008	1.104	-0.272	0.540	-0.194	-0.441	0.621	0.511
	4	100	70.5	64.5	68.0	*0.827*	0.111	0.103	-0.934	-0.006	0.185	-0.397	0.838	0.173	-0.668	-0.321	0.089
	5	116	68.4	71.4	70.1	*0.871*	0.022	0.330	-0.761	-0.012	1.365	-0.158	0.345	0.457	-0.629	-0.308	0.670
	6	28	83.3	80.0	81.0	***0.638***	0.010	0.296	-0.130	-0.012	2.288	-0.940	0.087	0.670	-0.047	-0.810	1.019
	7	52	52.4	4.4	48.7	0.909	0.103	-0.134	0.636	0.019	-0.457	-2.122	0.834	-0.226	0.523	0.421	-0.221
	All	524	60.9	68.2	64.090	*0.898*	0.002	0.104	-0.667	-0.007	1.605	0.317	0.019	0.150	-0.518	-0.311	0.773
2000 to 2001	1	22	72.7	81.8	77.3	***0.670***	-0.095	-0.961	2.335	0.038	1.824	-1.338	-1.434	-0.743	1.139	0.394	0.904
	2	66	61.5	66.0	65.2	*0.832*	0.080	-0.364	-0.442	0.030	-0.464	-1.063	0.724	-0.536	-0.385	0.480	-0.231
	3	26	100.0	95.8	96.2	***0.672***	-0.114	-0.027	0.001	0.013	1.137	-0.116	-0.858	0.116	0.001	0.256	0.536
	4	50	61.1	53.1	56.0	0.968	-0.004	0.304	0.168	-0.034	-0.705	0.655	-0.036	0.836	0.146	-0.537	-0.357
	5	58	50.0	55.0	53.4	0.967	0.024	-0.200	0.737	0.011	0.370	-1.680	0.360	-0.495	0.517	0.153	0.188
	6	14	90.0	75.0	85.7	***0.547***	0.044	0.229	-0.276	0.026	-0.107	-1.057	0.602	0.515	-0.256	0.418	-0.055
	7	26	60.0	72.7	65.4	*0.855*	0.074	-0.529	1.121	-0.010	0.454	-1.329	0.635	-0.975	0.926	-0.182	0.225
	All	262	54.0	54.9	54.6	0.987	0.015	0.079	0.306	-0.014	1.838	-1.235	0.191	0.195	0.242	-0.236	0.915

Table 25: Fisher discrimination analysis by sector for the three phases.

[a] sector 1: Consumer Staples; 2: Consumer Discretionary; 3: Materials; 4: Industrials; 5: Information Technology & Communications; 6: Health Care; 7: Energy

[b] Number of samples in the sector

Year	Met (%)	Not Met(%)	Overall (%)	R^2	Coefficients $(t-1)$						Wald's Statistic[a]					
					RoCE	CR	PARR	TR	Met	Constant	RoCE	CR	PARR	TR	Met	Constant
1993 to 2001	47.5	72.2	61.1	0.054	0.032	-0.037	-0.098	0.014		-0.841	3.038	0.240	0.251	1.435		2.761
	30.5	77.8	6.5	0.034	0.004	-0.126	-0.183	0.013		-0.130	0.054	1.597	0.672	1.026		0.058
	84.9	34.5	62.6	0.032	-0.011	-0.021	-0.390	-0.005		1.247	0.344	0.041	3.199	0.178		3.662
	45.6	68.9	58.8	0.120	0.100	0.297	-0.771	-0.012		0.548	0.409	*3.943*	*9.258*	2.772		1.019
	6.5	65.2	61.1	0.079	0.022	0.152	-0.491	-0.012		0.371	1.737	1.043	*4.655*	3.481		0.400
	8.0	91.4	9.4	0.031	0.025	-0.106	-0.198	0.002		-0.537	2.524	0.546	0.598	0.293		0.880
	71.2	53.8	62.6	0.064	-0.010	-0.025	-0.621	-0.007		1.367	0.502	0.038	5.376	2.261		*5.924*
	28.6	85.3	61.1	0.050	0.033	0.215	0.111	-0.009		-1.135	3.676	3.575	0.249	0.442		3.563
	0.0	100.0	673	0.005	-0.008	-0.050	-0.035	-0.002		-0.806	0.238	0.406	0.016	0.018		1.318
1993 to 2001	381	86.1	84.7	0.420	0.035	0.069	0.070	-0.009	-3.513	1.049	2.173	0.694	0.072	0.389	***44.670***	2.071
	88.1	90.3	983	0.529	-0.050	-0.405	-0.127	0.009	-4.861	3.726	2.729	3.279	0.125	0.179	***44.586***	***12.303***
	975	34.5	9.5	0.035	-0.012	-0.022	-0.376	-0.005	-0.234	1.383	0.382	0.044	2.949	0.214	0.414	4.043
	6.7	75.7	71.8	0.167	0.007	0.307	-0.660	-0.012	-1.062	0.848	0.156	*4.005*	*6.506*	2.468	***6.984***	2.164
	62.9	73.9	68.7	0.165	0.014	0.102	-0.345	-0.009	-1.378	1.069	0.665	0.421	2.106	2.285	***12.266***	2.654
	44.0	80.2	6.4	0.088	0.022	-0.090	-0.160	0.003	-1.058	-0.078	1.780	0.347	0.355	0.581	***7.682***	0.016
	65.2	67.7	6.4	0.112	-0.023	0.014	-0.636	-0.008	-1.067	2.250	2.133	0.012	*5.473*	3.020	***6.604***	***0.8193***
	46.4	74.7	62.6	0.095	0.030	0.184	0.198	-0.014	-0.951	-0.547	2.800	2.410	0.733	0.843	*6.090*	0.700
	0.0	100.0	673	0.011	-0.004	-0.053	-0.044	-0.001	0.410	-1.145	0.040	0.481	0.025	0.003	0.833	2.079
1993 to 2001	673	90.3	84.0	0.413	0.014	-0.011	0.120	-0.007	2.589	-0.357	0.376	0.015	0.243	0.221	***27.166***	0.236
	84.7	91.7	88.5	0.542	-0.047	-0.346	-0.083	-0.008	4.446	1.575	2.737	2.379	0.075	0.318	***27.020***	2.381
	78.1	34.5	58.8	0.040	-0.014	-0.038	-0.381	-0.007	0.180	1.372	0.540	0.122	3.065	0.300	1.084	*4.257*
	52.6	71.6	6.4	0.122	0.010	0.289	-0.749	-0.013	0.325	0.558	0.392	3.678	***8.557***	3.051	0.297	1.047
	67.7	65.2	6.4	0.160	0.008	0.109	-0.520	-0.009	1.884	0.789	0.226	0.461	*4.956*	2.109	***8.150***	1.545
	6.0	90.1	61.8	0.057	0.022	-0.094	-0.169	0.004	1.077	-0.570	1.765	0.348	0.421	0.705	2.963	0.881
	71.2	55.4	6.4	0.064	-0.011	-0.021	-0.622	-0.007	0.056	1.388	0.502	0.026	*5.379*	2.257	0.021	*5.719*
	51.8	81.3	68.7	0.133	0.032	0.171	0.113	-0.013	1.685	-0.911	3.202	2.031	0.232	0.703	***7.695***	1.981
	0.0	100.0	673	0.007	-0.006	-0.051	-0.016	-0.001	-0.145	-0.912	0.119	0.435	0.003	0.002	0.339	1.599

Table 26: Logistic Regression. (Top panel: Using the four indicator measures; Middle panel: With binary lagged-met indicator; Bottom panel: With numerical lagged-met indicator)

[a] Standardized by the standard error.

Year	Met (%)	Not Met(%)	Overall (%)	R^2	Coefficients $(t-1)$						Wald's Statistic					
					RoCE	CR	PARR	TR	Met	Constant	RoCE	CR	PARR	TR	Met	Constant
1993 to 2001	44.8	68.1	57.7	0.051	-0.302	-0.027	0.049	-0.093		0.894	4.476	0.036	0.135	0.401		1.514
	3.8	81.9	51.5	0.016	-0.120	0.108	0.104	-0.049		-0.345	0.803	0.680	0.619	0.136		0.230
	75.0	41.1	60.0	0.040	-0.059	-0.165	0.209	0.150		-0.179	0.173	1.394	2.482	1.292		0.047
	50.9	71.2	62.3	0.082	-0.037	-0.052	0.431	0.026		-1.366	0.068	0.145	***9.736***	0.034		2.903
	58.1	69.1	6.8	0.102	-0.372	-0.222	0.261	0.304		-0.015	*6.308*	2.414	3.696	*4.732*		0.000
	22.0	87.5	62.3	0.051	-0.286	-0.118	0.267	-0.050		0.063	3.817	0.698	3.701	0.139		0.006
	6.7	60.9	6.8	0.064	0.102	-0.073	0.353	-0.008		-1.085	0.607	0.309	***7.048***	0.003		2.175
	5.4	90.5	3.8	0.013	-0.153	-0.076	0.040	-0.027		0.364	1.227	0.303	0.097	0.037		0.253
	0.0	100.0	672	0.008	0.108	0.000	-0.041	0.102		-1.681	0.498	0.000	0.074	0.460		3.507
1993 to 2001	82.8	86.1	84.6	0.418	-0.307	-0.164	0.101	0.156	-3.450	2.231	2.568	0.707	0.329	0.599	***44.490***	*5.001*
	81.0	84.7	381	0.388	0.042	0.228	0.038	0.144	-3.388	0.197	0.051	1.612	0.044	0.596	***42.257***	0.044
	75.0	39.7	9.2	0.042	-0.055	-0.168	0.201	0.152	-0.212	-0.045	0.150	1.435	2.274	1.313	0.337	0.003
	68.4	78.1	378	0.143	0.043	-0.038	0.386	-0.037	-1.178	-0.833	0.083	0.072	***7.289***	0.064	***8.596***	0.970
	6.1	73.5	70.0	0.168	-0.272	-0.166	0.172	0.224	-1.247	0.717	3.072	1.227	1.473	2.385	***9.611***	0.589
	38.0	85.0	6.9	0.101	-0.214	-0.112	0.234	-0.110	-1.019	0.613	1.979	0.597	2.715	0.613	***6.744***	0.518
	6.6	60.9	62.3	0.107	0.234	-0.075	0.348	0.010	-1.018	-0.879	2.516	0.315	***6.569***	0.006	*5.794*	1.354
	60.7	64.9	6.1	0.059	-0.101	-0.033	-0.027	0.005	-0.940	0.633	0.493	0.052	0.040	0.001	***6.034***	0.703
	0.0	100.0	672	0.012	0.067	-0.012	-0.037	0.092	0.332	-1.699	0.169	0.006	0.060	0.363	0.531	3.557

Table 27: Logistic regression on the rank data, categorized into quintile.

Yr	Sec [a]	N[b]	Met (%)	Not Met(%)	Overall (%)	R^2	Coefficients $(t-1)$ RoCE	CR	PARR	TR	Met	Const	Wald's Statistic RoCE	CR	PARR	TR	met
1993 to 1995	1	33	95.5	63.6	84.8	0.327	-0.025	0.717	-0.683	-0.027	-3.139	3.427	0.253	1.063	0.537	0.086	*6.412*
	2	99	75.6	75.9	75.8	0.308	0.015	0.163	-0.351	-0.027	-2.634	2.040	0.204	0.562	1.892	1.652	***26.189***
	3	39	63.2	65.0	64.1	0.259	0.091	-0.126	0.726	-0.050	-0.350	-0.466	1.791	0.339	2.383	2.479	0.159
	4	75	71.8	88.9	80.0	0.376	-0.049	-0.037	0.319	0.024	-3.573	1.935	1.249	0.060	0.922	1.966	***02377***
	5	87	60.0	80.9	71.3	0.176	0.014	-0.090	-0.064	-0.002	-1.868	1.026	0.386	0.415	0.034	0.021	***13.529***
	6	21	71.4	92.9	85.7	0.384	0.006	0.392	-0.244	-0.041	-3.664	2.547	0.009	0.126	0.119	0.295	*6.150*
	7	39	78.9	85.0	82.1	0.297	-0.012	-0.013	-1.083	0.021	-1.845	1.967	0.016	0.021	2.259	1.340	*4.796*
	All	393	71.2	79.7	75.6	0.244	0.002	-0.032	-0.159	-0.001	-2.256	1.511	0.026	0.430	1.338	0.024	***87.590***
1996 to 1999	1	44	0.0	93.8	68.2	0.074	0.022	-0.694	-0.512	0.046	-0.236	-1.184	0.429	0.774	0.507	0.726	0.106
	2	132	68.9	83.1	76.5	0.279	0.059	0.301	-0.326	-0.004	-2.105	0.123	3.224	1.795	1.589	0.335	***23.929***
	3	52	12.5	97.2	71.2	0.132	-0.049	-0.146	-0.922	-0.004	-0.825	2.226	1.436	0.453	2.480	0.483	1.038
	4	100	78.2	48.9	65.0	0.124	0.057	0.113	-0.856	-0.004	-0.377	0.569	3.974	0.522	7.057	0.804	0.631
	5	116	58.2	73.8	66.4	0.100	0.008	0.177	-0.552	-0.009	-0.883	0.967	0.333	1.052	4.062	0.725	4.707
	6	28	50.0	85.0	75.0	0.380	0.012	1.151	-0.567	-0.055	-2.973	0.905	0.011	2.645	0.083	3.589	3.470
	7	52	71.4	62.5	67.3	0.130	-0.082	-0.034	-0.570	-0.010	0.022	2.414	2.749	0.038	2.213	0.966	0.001
	All	524	55.7	72.7	65.1	0.101	0.002	0.076	-0.457	-0.005	-1.031	1.041	0.038	1.258	***13.326***	*4.532*	9.908
2000 to 2001	1	22	81.8	81.8	81.8	0.316	0.124	1.163	-3.115	-0.042	2.545	-0.695	3.918	0.733	4.015	0.219	3.085
	2	66	30.8	98.1	84.8	0.173	-0.123	0.319	0.696	-0.042	-0.694	0.766	3.455	2.050	2.454	1.962	0.904
	3	26	100.0	100.0	100.0	0.419	-6.252	10.076	14.948	1.840	-68.978	-140.569	0.000	0.000	0.000	0.000	0.000
	4	50	5.6	100.0	66.0	0.033	0.001	-0.130	-0.055	0.013	-0.264	-0.535	0.002	0.745	0.023	0.306	0.164
	5	58	0.0	100.0	69.0	0.037	0.010	-0.152	0.266	0.007	-0.142	-1.315	0.228	0.453	0.408	0.096	0.050
	6	14	100.0	100.0	100.0	0.698	-6.002	91.195	21.772	6.556	35.256	-160.720	0.000	0.000	0.000	0.000	0.000
	7	26	60.0	45.5	53.8	0.157	-0.059	0.612	-1.109	0.007	0.534	1.000	0.611	1.727	1.945	0.050	0.154
	All	262	0.0	100.0	66.8	0.013	0.004	0.020	0.076	-0.004	0.452	-1.009	0.118	0.133	0.202	0.195	2.765

Table 28: Logistic regression by sectors.

[a] Sectors: 1. Consumer Staples; 2. Consumer Discretionary; 3. Materials; 4. Industrials; 5. Information Technology & Communications; 6. Health Care; 7. Energy

[b] Number of samples in the sector

Year	Met (%)	Not Met(%)	Overall (%)	Wilk's Λ	Fisher Coefficients ($t-1$)						Standardized Coefficients				
					1st	2nd	3rd	4th	Met	Const	1st	2nd	3rd	4th	Met
1993 to 2001	381	86.1	84.7	***0.512***	-0.018	0.005	0.014	0.006	2.691	-1.494	-0.073	0.030	0.145	0.101	0.975
	988	90.3	90.1	***0.355***	-0.010	-0.065	-0.012	-0.006	3.356	-1.286	-0.039	-0.305	-0.136	-0.132	1.044
	49.3	62.1	55.0	0.986	0.047	-0.085	-0.047	0.035	0.959	-1.004	0.068	-0.145	-0.494	0.661	0.480
	64.9	62.2	6.4	***0.886***	0.653	0.013	0.016	-0.002	1.685	0.211	0.517	0.021	0.202	-0.049	0.808
	62.9	73.9	68.7	***0.859***	0.030	0.021	-0.015	-0.003	2.066	-0.919	0.026	0.032	-0.167	-0.055	0.959
	72.0	60.5	64.9	***0.892***	0.007	-0.041	-0.033	0.014	1.728	-1.887	0.005	-0.054	-0.407	0.480	0.839
	9.1	63.1	61.1	0.924	0.620	0.031	0.044	-0.002	-1.685	-1.084	0.536	0.053	0.566	-0.067	-0.811
	60.7	62.7	61.8	*0.916*	0.162	-0.263	-0.027	-0.005	1.707	-1.925	0.205	-0.433	-0.316	-0.155	0.835
	64.5	46.0	50.4	0.986	0.625	-0.006	-0.019	0.015	1.310	-0.631	0.470	-0.013	-0.246	0.500	0.651

Year	Met (%)	Not Met(%)	Overall (%)	$\mathbf{R}^2$	Coefficients ($t-1$)						Wald's Statistic					
					1st	2nd	3rd	4th	Met	Const	1st	2nd	3rd	4th	Met	Const
1993 to 2001	381	86.1	84.7	0.418	-0.032	0.014	0.028	0.011	-3.375	1.019	0.346	0.085	1.364	0.656	***46.887***	3.208
	988	90.3	90.1	0.531	-0.183	-0.126	-0.012	-0.026	-4.725	3.156	0.932	1.957	0.110	1.646	***46.091***	***12.170***
	87.7	20.7	58.0	0.014	0.011	-0.022	-0.012	0.009	-0.233	0.220	0.008	0.045	0.462	0.809	0.419	0.184
	64.9	70.3	67.9	0.113	0.480	0.007	0.012	-0.002	-1.177	1.065	3.827	0.003	0.606	0.041	***9.247***	3.542
	62.9	73.9	68.7	0.135	0.024	0.017	-0.012	-0.002	-1.524	0.731	0.012	0.018	0.507	0.058	***15.562***	1.616
	50.0	80.2	68.7	0.107	-0.011	-0.036	-0.025	0.010	-1.202	-0.738	0.002	0.060	2.503	3.199	***9.026***	1.497
	60.6	63.1	61.8	0.076	-0.370	-0.017	-0.025	0.001	-0.962	1.631	2.681	0.025	2.713	0.048	*5.727*	*6.223*
	3.6	68.0	61.8	0.082	0.108	-0.159	-0.017	0.001	-1.020	-0.598	0.459	1.918	1.045	0.039	***7.160***	1.060
	0.0	100.0	673	0.014	-0.176	-0.001	0.005	-0.005	0.382	-1.376	0.427	0.000	0.091	0.415	0.727	*4.353*

Table 29: Discriminant analysis (top panel) and logistic regression (bottom panel) on the transformed components.

Judge Institute of Management 2003-2004 Working Paper Series

NUMBER	TITLE AND AUTHOR
1/2003	Applying a CBA Model to the 3 Gorges Project in China (R Morimoto & C Hope)
2/2003	Global Strategic Alliance Membership: Market versus institutional criteria (M Warner, E Yin & C J Choi)
3/2003	Culture and Management in China (J Child & M Warner)
4/2003	Nature, Nurture & Economic Growth (B Dowling)
5/2003	Portfolio Management for Pension Funds (S Arbeleche, M A H Dempster, E A Medova & G W P Thompson)
6/2003	Evolutionary Reinforcement Learning in FX Order Book and Order Flow Analysis (R G Bates, M A H Dempster & Y S Romahi)
7/2003	Structured Products for Pension Funds (M A H Dempster, M Germano, E A Medova & M Villaverde)
8/2003	Is the Long Wave Getting Shorter? (B Dowling)
9/2003	A Framework to Measure Integrated Risk (E A Medova & R G Smith)
10/2003	The Marginal Impacts of CO2, CH4 and SF6 Emissions (C Hope)
11/2003	Review: Structuration Theory and Information Systems Research (M Jones & M Karsten)
12/2003	Global Fund Management Using Stochastic Optimization (M Villaverde)
13/2003	Econometric Modelling for Global Asset Liability Management (S Arbeleche & M A H Dempster)
14/2003	Hedging European and Barrier Options Using Stochastic Optimization (M Villaverde)
15/2003	Adaptive Systems for Foreign Exchange Trading (M P Austin, R G Bates, M A H Dempster & S N Williams)
01/2004	Confessions of an IS Consultant or the Limitations of Structuration Theory (M P A Thompson)
02/2004	Changing Patterns of Human Resource Management in Contemporary China: WTO accession and enterprise responses (M Warner, Y Zhu)
03/2004	Reputation Resources, Commitment and Performance of Film Projects in the USA and Canada (1988-1997) (A L Hadida)
04/2004	The Management of Human Resources in Shanghai: A case study of policy responses to employment and unemployment in the People's Republic of China (G O M Lee & M Warner)
05/2004	Chance, Choice and Determinism in Strategy (M de Rond & R-A Thiétart)
06/2004	Prospective Earnings per Share (R G Bates, M A H Dempster, H G Go & Y S Yong)

Working Papers Order Form

Name: ____________________

Address: ____________________

____

Email: ____________________ Phone: ____________________

Paper(s) required:

Number	Author(s)	Title

Price per publication = £5.50, including postage and packing.
For orders of 5 papers or more, postage and packing will be billed separately.
Credit card payments up to £20.00 will incur an additional standard charge of £1.00.
Credit card payments over £20.00 will incur an additional standard charge of £3.00.

Total payment enclosed: £____________________

Payment method:

❑ Cheque payable to "University of Cambridge"

❑ Draft/Money order

❑ Credit card Card type ☐ Visa ☐ Mastercard

Name on card ____________________

Card number ____________________

Expiry (mm/yy) ____________________

Signature ____________________

Date ____________________

Please send this completed form, together with your payment, to:

Research Support Manager
Judge Institute of Management, University of Cambridge
Trumpington Street, Cambridge CB2 1AG, UK
Tel: + 44 1223 760546 Fax: + 44 1223 339701
E-mail: research-support@jims.cam.ac.uk