Journal of Business Finance & Accounting, 26(5) & (6), June/July 1999, 0306-686X

The Accuracy, Bias and Efficiency of Analysts' Long Run Earnings Growth Forecasts

Richard D.F. Harris*

1. INTRODUCTION

Considerable research has now been undertaken into professional analysts' forecasts of companies' earnings in respect of both their accuracy relative to the predictions of time series models of earnings, and their rationality. The evaluation of the reliability of analysts' earnings growth forecasts is an important aspect of research in accounting and finance for a number of reasons. Firstly, many empirical studies employ analysts' consensus forecasts as a proxy for the market's expectation of future earnings in order to identify the unanticipated component of earnings. The use of consensus forecasts in this way is predicated on the assumption that they are unbiased and efficient forecasts of future earnings growth. Secondly, institutional investors make considerable use of analysts' forecasts when evaluating and selecting individual shares. The quality of the forecasts that they employ therefore has important practical consequences for portfolio performance. Finally, from an academic point of view, the performance of analysts' forecasts is interesting because it sheds light on the process by which agents form expectations about key economic and financial variables.

* The author is from the Department of Economics, University of Exeter. This paper was started while the author was a Visiting Fellow at the University of Warwick, April 1997. Financial Support from the ESRC (project no. R000234759) is gratefully acknowledged. He would also like to thank George Bulkley, Alan Gregory and an anonymous referee for their helpful comments. (Paper received March 1998, revised and accepted December 1998)

Address for correspondence: Richard D.F. Harris, School of Business and Economics, University of Exeter, Exeter EX4 4PU, UK.
e-mail: R.D.F.Harris@exeter.ac.uk

Nearly all of the research to date, however, has been concerned with analysts' forecasts of quarterly and annual earnings per share.[1] While the properties of analysts' short run forecasts are undoubtedly important in their own right, it is long run expectations of earnings growth that are more relevant for security pricing (see, for instance, Brown et al., 1985). A number of papers have suggested that there is substantial mis-pricing in the stock market as a consequence of irrational long run earnings growth forecasts being incorporated into the market expectation of earnings growth (DeBondt, 1992; La Porta, 1996; Bulkley and Harris, 1997; and Dechow and Sloan, 1997). The evaluation of the performance of analysts' long run forecasts is clearly important as corroborating evidence.

This paper provides a detailed study of the accuracy, bias and efficiency of analysts' long run earnings growth forecasts for US companies. It identifies a number of characteristics of forecast earnings growth. Firstly, the accuracy of analysts' long run earnings growth forecasts is shown to be extremely low. So low, in fact, that they are inferior to the forecasts of a naïve model in which earnings are assumed to follow a martingale. Secondly, analysts' long run earnings growth forecasts are found to be significantly biased, with forecast earnings growth exceeding actual earnings growth by an average of about seven percent per annum. Thirdly, analysts' forecasts are shown to be weakly inefficient in the sense that forecast errors are correlated with the forecasts themselves. In particular, low forecasts are associated with low forecast errors, while high forecasts are associated with high forecast errors. The bias and inefficiency in analysts' long run forecasts are considerably more pronounced than in their short run and interim forecasts.

It is investigated whether analysts incorporate information about future earnings that is contained in current share prices. It is demonstrated that consistent with their short run and interim forecasts, analysts' long run earnings growth forecasts can be enhanced by assuming that each individual firm's earnings will evolve in such a way that its price-earnings ratio will converge to the current market average price-earnings ratio. Analysts therefore neglect valuable information about future earnings that is readily available at the time that their forecasts are made.

The source of analyst inaccuracy is explored by decomposing the mean square error of analysts' forecasts into two systematic components, representing the error that arises as a result of forecast bias and forecast inefficiency, and a random, unpredictable component. In principle, the systematic components of analysts' forecast errors can be eliminated by taking into account the bias and inefficiency in their forecasts. However, it is shown that the bias and inefficiency of analysts' forecasts contribute very little to their inaccuracy. Over eighty-eight percent of the mean square forecast error is random, while less than twelve percent is due to the systematic components. This is an important result for the users of analysts' forecasts since it means that correcting forecasts for their systematic errors can potentially yield only a small improvement in their accuracy.

A second decomposition is used to examine the level of aggregation at which forecast errors are made. The mean square forecast error is decomposed into the error in forecasting average earnings growth in the economy, the error in forecasting the deviation of average growth in each industry from average growth in the economy, and the error in forecasting the deviation of earnings growth for individual firms from average industry growth. It is demonstrated that the error in forecasting average earnings growth in the economy contributes relatively little to analysts' inaccuracy. Over half of total forecast error arises from the error in forecasting deviations of individual firm growth from average industry growth. The error in forecasting deviations of average industry growth from average growth in the economy is smaller, but also significant. However, there is evidence that this pattern is changing over time, with increasing accuracy at the industry level, and diminishing accuracy at the individual firm level.

Finally, it is shown that the performance of analysts' long run earnings growth forecasts varies substantially both with the characteristics of the company whose earnings are being forecast and of the forecast itself. The accuracy, bias and efficiency of analysts' forecasts is examined for sub-samples of firms partitioned by market capitalisation, price-earnings ratio, market-to-book ratio and the level of the forecast itself. The most reliable earnings growth forecasts are low forecasts issued for large companies with low price-earnings ratios and high market-to-book ratios. Again, this is of considerable practical importance since it offers users of analysts' forecasts some opportunity to discriminate between good and bad forecasts.

The organisation of this paper is as follows. The following section gives a detailed description of the data sources and the sample selection criteria. Section 3 describes the methodology used to evaluate forecast accuracy, bias and efficiency. Section 4 reports the results, while Section 5 concludes.

2. DATA

The sample is drawn from all companies listed on the New York, American and NASDAQ stock exchanges. Data on long run earnings growth expectations are taken from the Institutional Brokers Estimate System (IBES). The data item used in this paper is the 'expected EPS long run growth rate' (item 0), which has been reported by IBES since December 1981, and is defined as:

> the anticipated growth rate in earnings per share over the longer term. IBES Inc. requests that contributing firms focus on the five-year interval that begins on the first day of the current fiscal year and make their calculations based on projections of EPS before extraordinary items.

The expected long term growth rate is therefore taken to be the forecast average annual growth in earnings per share before extraordinary items, over the five year period that starts at the beginning of the current fiscal year.[2] The measure used in this paper is the median forecast calculated and reported in April of each year, *t*. The analysis was also conducted using the mean forecast, but the quantitative results are virtually identical, and the qualitative conclusions unchanged.[3]

Only December fiscal year end companies are included in the sample and so the use of the consensus forecast reported in April should ensure that the previous fiscal year's earnings are public information at the time that the individual forecasts that make up the consensus forecast are made (see Alford, Jones and Zmijewski, 1994). Restricting the sample to December fiscal year-end companies ensures that observations for a particular fiscal year span the same calendar period, thus allowing the identification of macroeconomic shocks that contemporaneously affect the earnings of all firms.

Actual growth in earnings is calculated using data on earnings per share, excluding extraordinary items, taken from the Standard and Poor's Compustat database (item EPSFX). Average annual earnings growth is computed as the average change in earnings over each five year period, from December of year $t-1$ to December of year $t+5$, scaled by earnings in December of year $t-1$. The need for five years' subsequent earnings growth data limits the sample period to the eleven years 1982–92. Data on a number of other variables are also used in the analysis. The share price and market capitalisation are both taken at the end of April of year t (Compustat items PRCCM and MKVALM). The market price-earnings ratio, used to test whether information contained in the share price is incorporated in analysts' forecasts, is computed as the price at the end of April in year t (item PRCCM) divided by earnings per share in the fiscal year ending December $t-1$ (item EPSFX). The market-to-book ratio is computed as the market value of the company in April of year t (item MKVALM) divided by the book value of the company in the fiscal year ending December of year $t-1$ (item CEQ).

There are a total of 7,660 firm-year observations that satisfy the data requirements for all the variables used in the analysis, and that have a December fiscal year-end. However, for 658 of these, earnings reported at the end of the preceding fiscal year are zero or negative. These are omitted from the sample since forecast growth has no natural interpretation when earnings in the base year are non-positive.[4] When initial earnings are close to zero, actual growth in earnings may take extreme values, resulting in outliers that have a disproportionately high degree of influence on the least squares regression results. There is no immediately obvious way to circumvent this problem without dropping some observations from the sample. The approach most commonly adopted is to omit observations for which the calculated growth rate, the forecast growth rate or the forecast error is above a certain threshold in absolute value, or for which calculated initial earnings are below a certain level. For instance, Fried and Givoly (1982) truncate observations for which forecast error exceeds 100%. Elton et al. (1984) include in their sample only those companies for which initial earnings are above 0.20 dollars per share. O'Brien (1988), in order to test the robustness of her results to outliers, also uses 0.20 dollars as a threshold value.

Capstaff et al. (1995) omit observations for which forecast earnings growth or forecast error exceeds 100%, while Capstaff et al. (1998) exclude companies for which forecast earnings growth or actual earnings growth exceeds 100%. In this paper, all observations for which actual earnings growth or forecast earnings growth exceeds 100% in absolute value are omitted from the analysis, reducing the sample by a further 336 firm-year observations. The final pooled sample comprises 6,666 firm-year observations.[5]

3. METHODOLOGY

(i) Forecast Accuracy

The metric used to evaluate forecast performance is the forecast error, defined as the difference between actual and forecast earnings growth:

$$fe_{it} = g_{it} - g_{it}^{f} \tag{1}$$

where fe_{it} is the forecast error for firm i corresponding to the forecast made at date t, g_{it} is actual earnings growth over the five year forecast period and g_{it}^{f} is forecast five year earnings growth. Forecast accuracy is evaluated using the mean square forecast error, which is computed in each year t as:

$$\mathrm{MSFE}_t = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f})^2. \tag{2}$$

The mean square forecast error for the pooled sample is computed over all firms and years. The mean square forecast error was chosen in preference to the mean absolute forecast error to maintain consistency with the subsequent analysis which uses the former measure rather than the latter. However, it should be noted that the use of the mean square forecast error is consistent with a quadratic loss function of risk averse economic agents (see Theil, 1964; and Mincer and Zarnovitz, 1969). It can be reported that the conclusions drawn about forecast accuracy are not sensitive to the choice of measure.

As a benchmark against which to compare the accuracy of analysts' long run forecasts, the performance of two 'naïve' forecasts is also considered. The first is the forecast generated by a martingale model of earnings, in which expected earnings growth is zero. The second is the forecast generated by a sub-martingale model, in which expected earnings is equal to a drift parameter that is identical for all firms. In each forecast year, the common drift parameter is set equal to the average growth rate in earnings over all firms, over the previous five year period.[6] This choice of naïve forecasts is motivated by the early evidence on the time series properties of earnings, which suggests that annual earnings follow a random walk, or a random walk with drift (see, for instance, Brooks and Buckmaster, 1976; or Foster, 1977). Although more recent evidence finds that annual earnings may have a mean reverting component (see Ramakrishnan and Thomas, 1992), the martingale and sub-martingale models of earnings nevertheless provide simple alternative models that are approximately consistent with the reported evidence.

(ii) Forecast Bias

In order for a forecast to be unbiased, the unconditional expectation of the forecast error must be zero. If the average forecast error is greater than zero then analysts are systematically over-pessimistic (since their forecasts are on average exceeded) while if the average forecast error is less than zero analysts are systematically over-optimistic (since their forecasts are on average unfulfilled). Unbiasedness is tested using the mean forecast error, which is computed in each year t as:

$$\mathrm{MFE}_t = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^f). \tag{3}$$

The mean forecast error for the pooled sample is computed over all firms and years. The hypothesis that the mean forecast error is zero is tested using the standard error of the mean forecast error across all firms and years for the pooled sample, and across all firms for each of the annual samples.

(iii) Forecast Efficiency

A forecast is efficient if it optimally reflects currently available information, and is therefore associated with a forecast error that is unpredictable. If a forecast is strongly efficient, the forecast error is uncorrelated with the entire information set at time *t*. Strong efficiency is a stringent condition, and so more usually forecasts are instead tested for weak efficiency, which requires that the forecast error is uncorrelated with the forecast itself (see Nordhaus, 1987). Weak efficiency is tested by estimating the following regression:

$$g_{it} = \alpha + \beta g_{it}^{f} + v_{it}. \tag{4}$$

Under the null hypothesis that analysts' forecasts are weakly efficient, the intercept, α, should be zero, while the slope coefficient, β, should be unity. If β is significantly different from one then conditioning on the forecast itself, the forecast error is predictable.[7] If β is significantly less than one then analysts' forecasts are too extreme, in the sense that high forecasts are associated with high forecast errors, while low forecasts are associated with low forecast errors. If β is significantly greater than one then forecasts are too compressed.

(iv) The Incremental Information Content of Price-Earnings Based Forecasts

A stronger form of forecast efficiency can be tested by examining whether analysts' forecasts incorporate particular sources of publicly available information. One such source of information is the current share price. In an efficient market, the share price is the present discounted value of all rationally expected future economic earnings of the company, and hence it should reflect, *inter alia*, the market's expectation of long run earnings growth. To extract the information about future earnings embodied in the share price, some assumption must be made about the company's cost of equity, or risk. The simplest assumption is that all companies face the same constant cost of equity in the long run, so that the earnings of each company evolve in such a way that its price-earnings ratio converges to the current market average price-earnings ratio. The earnings growth forecast that is implicit in this assumption can then be used to supplement the analysts' earnings growth forecast in the following regression:

$$g_{it} = \alpha + \beta g_{it}^{f} + \gamma g_{it}^{p} + v_{it}, \tag{5}$$

where

$$g_{it}^{p} = \frac{p_{it}/pe_{mt} - e_{it}}{e_{it}}, pe_{mt} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_{it}}{e_{it}}$$

and p_{it} is the share price of firm i at time t. If analysts incorporate all information contained in the current share price, the coefficient, γ, should be zero (see Capstaff et al., 1995 and 1998). Naturally, the assumption that all firms have the same long run price-earnings ratio is a strong simplification, and a superior forecast would almost certainly be obtained by assuming that price-earnings ratios differ between industries. Nevertheless, the assumption of a single market-wide long run price-earnings ratio has been shown to forecast earnings growth over shorter horizons (see, for instance, Ou and Penman, 1989).

(v) Forecast Error Decomposition

In order to analyse the source of analysts' forecast errors, two decompositions of the mean square forecast error are used. The first decomposes the mean square forecast error into systematic and unsystematic components. The systematic component is further divided into a component due to forecast bias and a component due to forecast inefficiency. In each year t, the decomposition of the MSFE is given by:

$$\text{MSFE}_t = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^{f})^2 = (\overline{g}_t - \overline{g}_t^{f})^2 + (1-\beta_t)^2\sigma_{g^f t}^2 + (1-\rho_t^2)\sigma_{gt}^2 \tag{6}$$

where N_t is the sample size in year t, $\overline{g}_t$ and $\overline{g}_t^{f}$ are the average values of g_{it} and g_{it}^{f}, β_t is the slope coefficient from regression (4), above, ρ_t is the correlation coefficient between g_{it} and g_{it}^{f}, and $\sigma_{g^f t}^2$ and σ_{gt}^2 are the variances of g_{it} and g_{it}^{f}. The first term in the decomposition gives the error that is due to the inability of analysts to forecast earnings growth for the whole sample. When computed over all years, it is therefore a measure of the error that is due to forecast bias. The second term captures the error that is due to forecast inefficiency. Together, these two terms capture the systematic error in analysts' forecasts. In contrast, the third term captures the component of the error that is purely random. This decomposition is particularly useful since it reveals to what extent forecasts can be improved through 'optimal linear correction' procedures (see Mincer and Zarnovitz, 1969; and Theil, 1966). For instance, if the main component of mean square error is systematic, rather than random, then assuming that the data generating process for both the actual data and the forecast data remains constant, the accuracy of analysts' forecasts can be substantially improved by using the predicted values from regression (4), above, rather than the forecasts themselves. The extent to which this reduces the inaccuracy of the forecasts depends upon the fraction of the mean square forecast error that is due to the systematic component.

The second decomposition breaks the mean square forecast error into economy, industry and firm components. The decomposition of the MSFE is given each year t by:

$$\begin{aligned} \mathrm{MSFE}_t &= \frac{1}{N}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2 \\ &= (\overline{g}_t - \overline{g}_t^f)^2 + \frac{1}{N_t}\sum_{j=1}^{J_t} N_{jt}[(\overline{g}_{jt} - \overline{g}_t) - (\overline{g}_{jt}^f - \overline{g}_t^f)]^2 \qquad (7) \\ &\quad + \frac{1}{N_t}\sum_{i=1}^{Nt}[(g_{it} - \overline{g}_{jt}) - (g_{it}^f - \overline{g}_{jt}^f)]^2, \end{aligned}$$

where J_t is the number of industries in the sample, N_{jt} is the number of firms in industry j, $\overline{g}_{jt}$ and $\overline{g}_{jt}^f$ are the average values of g_{it} and g_{it}^f in industry j. The decomposition has the following interpretation. As before, the first term measures the error that is due to analysts' inability to forecast the average growth for the whole sample, which in this context may be interpreted as their inability to forecast earnings growth for the economy. The second term measures the error that is due to an inability to forecast the deviation of average growth in an industry from average growth in the economy. The third term measures the error that is due to an inability to forecast deviation of individual firm growth from average growth in its industry. The decomposition for the pooled sample is computed by taking the weighted average of the decomposition for the annual samples, with weights proportional to the sample size each year. Such a decomposition is useful because it reveals the level of aggregation at which forecast errors are made, and may reflect the particular approach used to generate earnings growth forecasts (see Elton, Gruber and Gultekin, 1984). In the present study, each industry is defined by a two digit SIC code. This yields a total of 56 industries, with an average of about twelve firms in each industry. The use of three digit SIC codes yields a large number of industries that comprise only a single firm. In these cases, the firm-specific error and industry specific error are not separately identifiable, and are reflected in the third component of the decomposition. The effect of using two digit, rather than three digit SIC codes is therefore to increase the firm specific error and reduce the industry specific error.

For both decompositions, it is convenient to express each term as a percentage of the total mean square forecast error. For the pooled samples, the mean square forecast error components are averaged over the individual years, with weights proportional to the sample size each year.

(vi) The Performance of Analysts' Forecasts Conditional on Firm and Forecast Characteristics

In order to explore possible heterogeneity in the performance of analysts' long run earnings growth forecasts, the sample is partitioned by various characteristics of the firm whose earnings are being forecast and of the forecast itself. Specifically, the sample is split into equally sized quintiles on the basis of market capitalisation, market-to-book ratio, price-earnings ratio and the level of the forecast itself. Forecast accuracy, bias and efficiency is then examined for each sub-sample. Forecast accuracy is measured by the mean square forecast error given by (2), forecast bias is measured by the mean forecast error given by (3), while forecast efficiency is measured by the estimated slope parameter in regression (4).

In order to identify the marginal effects of each of the firm and forecast characteristics on forecast accuracy, bias and weak form efficiency, the following regressions are estimated:

$$(g_{it} - g_{it}^f)^2 = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}, \qquad (10)$$

$$g_{it} - g_{it}^f = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it} \qquad (11)$$

and

$$(g_{it}^f - \overline{g}_t^f)[(g_{it} - \overline{g}_t) - (g_{it}^f - \overline{g}_t^f)] = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}, \tag{12}$$

where $\ln m_{it}$ is the natural logarithm of the market capitalisation of firm i at the beginning of the forecast period, mb_{it} is the market-to-book ratio and pe_{it} is the price-earnings ratio. The dependent variables in the three regressions are the summands in (a) the mean square forecast error, (b) the mean forecast error and (c) the estimated covariance between $(g_{it} - g_{it}^f)$ and g_{it}^f.[8]

(vii) Estimation Procedure

In order to allow for time specific market wide shocks, each of the regression equations (4), (5), (9), (10), (11) and (12) is estimated by OLS, including fixed time effects. However, inference based on OLS estimates of the variance-covariance matrix of the disturbance term may be misleading since both heteroscedasticity and cross-sectional correlation are likely to be present in the data. One potential solution is to use GLS, in which the heteroscedasticity and cross-section correlation are parameterised and estimated. However, in the present case, GLS is infeasible since the number of cross-section observations is large relative to the number of time series observations. This paper employs instead the non-parametric approach of Froot (1989), which is robust to both contemporaneous correlation and heteroscedasticity. This involves partitioning the data by a two digit SIC code and assuming that the intra-industry correlation is zero. This then allows the consistent estimation of the parameter covariance matrix. The Froot estimator is modified using the Newey-West (1987) procedure in order to allow for the serial correlation in the regression error term that is induced by the use of overlapping data.

4. RESULTS

(i) Forecast Accuracy

Panel A of Table 1 reports the mean square forecast error, given by (2), for the pooled sample and for each individual year. It also reports the mean square forecast errors for the naïve forecasts of the martingale model, where forecast earnings growth is zero, and the sub-martingale model, where forecast earnings growth is the historical economy wide average earnings growth rate.

The accuracy of analysts' long run earnings growth forecasts is extremely low. In the pooled sample, the mean square forecast error for analysts is 7.15%. For the martingale model, the mean square error is 6.63%, while for the sub-martingale model, it is marginally lower at 6.60%. On average, therefore, a superior forecast of long run earnings growth for individual companies can be obtained simply by assuming that average annual earnings growth will be zero. This is a strong indictment of the accuracy of analysts' long run forecasts, and in view of the additional information available to analysts, is surprising. It also contrasts with the evidence for shorter horizon forecasts where analysts appear to have some advantage over time series models. Furthermore, the alternative models used here are relatively simple. If in fact earnings are stationary, then it is likely that a yet superior forecast could be obtained from an estimated time series model for each firm, and so the relative inferiority of analysts' forecasts is probably understated here.

Turning to the annual samples, the martingale model generates superior forecasts in seven out of eleven years, while the sub-martingale model generates forecasts that are superior to analysts' forecast in nine of the eleven years, and superior to the forecasts of the martingale model in ten out of eleven years. This suggests that one can improve on the zero growth forecast of the martingale model by using the historical economy average earnings growth rate to predict subsequent growth for individual firms. However, the improvement is only marginal, reflecting both considerable variation in average earnings growth between years and considerable dispersion in earnings growth rates across the economy. The time-series pattern of forecast errors suggests that analyst inferiority is not caused by just one or two outlying years. Nor does it suggest that there is any improvement in the accuracy of analysts' forecasts over the sample period, either relative to the forecasts of the martingale and sub-martingale models, or in absolute terms. The (unweighted) average mean square forecast error for the first five years in the sample is 7.02%, while in the last five years it is 7.28%. This is in contrast with evidence reported elsewhere that analyst accuracy has increased over time (see Brown, 1997).

(ii) Forecast Bias

Panel B of Table 1 reports the mean forecast error for analysts' forecasts of long run earnings growth, given by (3), and its standard error. In the pooled sample, the mean forecast error is negative indicating that analysts' long run earnings growth forecasts are over-optimistic. The mean forecast error is very significant both in statistical and economic terms. On average, forecast growth exceeds actual growth by about seven percent per annum. Over-optimism in long run earnings growth forecasts is consistent with evidence reported for analysts' shorter horizon earnings forecasts (see, for instance, Fried and Givoly, 1982; Brown et al., 1985; and O'Brien, 1988). It is also consistent with international evidence on analysts short run and interim forecasts (see Capstaff et al., 1995 and 1998).

The mean forecast error is also negative in each individual year, and significantly negative in all but the last, ranging from 1.50% to 11.82% per annum. This is in contrast with analysts' shorter horizon forecasts where the direction of the reported bias displays considerable year to year variation (see, for instance, Givoly, 1985). It is again notable that the degree of over-optimism has not diminished significantly over time. The (unweighted) mean forecast error for the first five years of the sample is -6.99%, while for the last five years it is -7.20%. It is of course possible that the last year in the sample, where the mean forecast error is less than two percent, marks the start of a reduction in analyst over-optimism. Whether this is borne out by future studies will be of considerable interest.

(iii) Forecast Efficiency

Panel A of Table 2 presents the results of regression (4). The efficiency condition is very strongly rejected for analysts' long run earnings growth forecasts. In the pooled sample, $\hat{\beta}$ is significantly less than unity and at 0.20, only marginally greater than zero. This is a considerably stronger rejection of efficiency than found by other authors for shorter horizon forecasts. For instance,

Table 1

Forecast Accuracy and Forecast Bias

	Panel A: Forecast Accuracy			Panel B: Forecast Bias	
	MSFE of Analysts	*MSFE of Martingale*	*MSFE of Sub-martingale*	*MFE of Analysts*	*Standard Error*
Pooled sample	7.15	6.63	6.60	−7.33	(0.31)
1982	7.34	5.15	6.41	−11.39	(1.01)
1983	6.88	7.01	6.51	−5.48	(1.20)
1984	6.75	7.14	6.40	−4.01	(1.12)
1985	7.19	6.67	6.29	−6.61	(1.08)
1986	6.92	6.47	6.24	−7.44	(1.08)
1987	6.95	5.77	5.75	−10.78	(0.99)
1988	7.38	6.32	6.40	−10.20	(1.00)
1989	6.99	5.22	5.71	−11.82	(0.91)
1990	5.69	5.20	4.95	−7.40	(0.85)
1991	7.58	7.78	7.60	−5.04	(0.99)
1992	8.78	9.62	9.78	−1.50	(1.10)

Notes:
Panel A reports the mean square forecast error for analysts' forecasts and the forecasts of two naïve models.

The MSFE of analysts forecasts is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f})^2$;

the MSFE of the martingale model is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it})^2$;

the MSFE of the sub-martingale model is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it} - \overline{g}_{t-1})^2$;

where g_{it} is five year earnings growth from January year t to December year t+4, is forecast of g_{it} reported at April year t and $\overline{g}_{t-1}$ is the average value over all companies of five year earnings growth from January year t−5 to December year t−1. The MSFE for the pooled sample is computed over all firms and years.

Panel B reports the mean forecast error of analysts, calculated as:

$$\text{MFE} = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f}),$$

and its standard error. The MFE for the pooled sample is computed over all firms and years.

DeBondt and Thaler (1990) find that while they reject the hypothesis that β is equal to unity for one and two year forecasts, their estimated parameters (0.65 for one year forecasts, 0.46 for two year forecasts) are much larger than those reported here, both statistically and economically. For annual earnings forecasts,

Table 2

Forecast Efficiency

	Panel A: Weak Efficiency			Panel B: The Incremental Information Content of Price-Earnings Based Forecasts				
	$\hat{\beta}$	*SE*	$\overline{R}^2$	$\hat{\beta}$	*SE*	$\hat{\gamma}$	*SE*	$\overline{R}^2$
Pooled sample	0.20	(0.08)	0.00	0.05	(0.09)	0.04	(0.01)	0.02
1982	−0.73	(0.26)	0.04	−0.81	(0.28)	0.03	(0.04)	0.05
1983	0.42	(0.25)	0.01	0.08	(0.27)	0.05	(0.02)	0.04
1984	0.19	(0.27)	0.00	0.03	(0.30)	0.04	(0.02)	0.01
1985	0.05	(0.29)	0.00	0.02	(0.33)	0.01	(0.02)	0.00
1986	0.31	(0.23)	0.01	−0.25	(0.22)	0.10	(0.02)	0.06
1987	0.46	(0.22)	0.01	0.41	(0.22)	0.01	(0.02)	0.01
1988	0.42	(0.21)	0.01	0.43	(0.21)	0.00	(0.01)	0.01
1989	0.08	(0.22)	0.00	−0.03	(0.23)	0.03	(0.02)	0.01
1990	0.28	(0.17)	0.01	0.20	(0.20)	0.02	(0.02)	0.01
1991	0.39	(0.17)	0.01	0.11	(0.50)	0.06	(0.03)	0.03
1992	0.09	(0.27)	0.00	−0.20	(0.31)	0.10	(0.03)	0.05

Notes:

Panel A reports the results of the test of the weak efficiency of analysts' forecasts. The regression for the pooled sample is $g_{it} = \alpha_t + \beta g^f_{it} + u_{it}$ where g_{it} is five year earnings growth from January year t to December year t+4 and g^f_{it} is the median forecast of g_{it} reported in April of year t. The regression for the annual samples is $g_{it} = \alpha_t + \beta_t g^f_{it} + u_{it}$. The Panel reports the estimated slope parameter, its Froot-Newey-West adjusted standard error and the adjusted R-squared statistic.

Panel B reports the results of the test for the incremental information content of price-earnings based forecasts. The regression for the pooled sample is $g_{it} = \alpha_t + \beta g^f_{it} + \gamma g^p_{it} + u_{it}$ where g_{it} is five year earnings growth from January year t to December year t+4, is the median forecast of g_{it} reported in April of year t,

$$g^p_{it} = \frac{p_{it}/pe_{mt} - e_{it}}{e_{it}}, \ pe_{mt} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_{it}}{e_{it}},$$

e_{it} is the earnings reported in December of year $t-1$, and p_{it} is the price in April of year t. The regression for the annual samples is $g_{it} = \alpha_t + \beta_t g^f_{it} + \gamma_t g^p_{it} + u_{it}$. The Panel reports the estimated slope parameter, its Froot-Newey-West adjusted standard error and the adjusted R-squared statistic.

Givoly (1985) cannot reject the hypothesis that β is unity. Using UK data on the forecasts of individual analysts, Capstaff et al. (1995) find that the estimated coefficient declines with the forecast horizon, with an estimated value of around 0.5 for 20 month forecasts (their longest horizon). The results of this paper therefore strongly support the view (first offered by DeBondt and Thaler, 1990) that forecast earnings growth is too extreme, and that the longer the horizon, the more extreme it becomes. In the annual regressions, β is significantly less than unity in all years, and significantly greater than zero in only three years. In one year, it is actually significantly negative.

(iv) The Incremental Information Content of Price-Earnings Based Forecasts

The results of regression (5), which supplements analysts' forecasts with forecasts that are derived from the assumption that earnings will evolve in such a way that each firm's price-earnings ratio will converge to the current market price-earnings ratio, are reported in Panel B of Table 2. Under the null hypothesis that analysts make optimal use of information about future earnings that is contained in share prices, the coefficient on the price-earnings based forecast, $\hat{\gamma}$, should be zero. In the pooled sample, the estimated coefficient is significantly greater than zero, implying that analysts do not make full use of information that is readily available at the time that their forecasts are made. However, there is much year to year variation in both the statistical and economic significance of the coefficient, with six years in which the coefficient is not statistically different from zero.

The marginal contribution of price-earnings based forecasts can be gauged by comparing the two Panels of Table 2. The inclusion of the price-earnings forecast explains an additional two percent of the variation in actual earnings growth in the pooled sample, while in individual years, this figure varies between zero and five percent. However, the price-earnings based forecast used in the present analysis is derived under the somewhat unrealistic assumption that all firms have a common long run price-earnings ratio. Undoubtedly, more accurate earnings growth forecasts could be imputed by making more sophisticated assumptions about how price-earnings ratios evolve over time. The results presented here therefore almost certainly understate the extent to which analysts neglect information embodied in share prices. The fact that analysts appear to neglect information contained in share prices when forming their long run earnings growth forecasts is consistent with analogous results for their forecasts over shorter horizons (see, for instance, Ou and Penman, 1989; Abarbanell, 1991; Elgers and Murray, 1992; and Capstaff et al., 1995 and 1998).

(v) Forecast Error Decomposition

The preceding results demonstrate that the accuracy of analysts' long run earnings forecasts is extremely low, and that they are very significantly biased and inefficient. In this sub-section, the source of analysts' forecast error is investigated using the two decompositions of mean square forecast error described in Section 3. The first decomposes forecast error into systematic and non-systematic components. The results of this decomposition are given in Panel A of Table 3. It can be seen that by far the largest component of mean square forecast error is random. In the pooled sample, less than twelve percent of the forecast error is the result of the systematic component of analysts' forecast errors. Of the systematic component, about seven percent is due to bias, and about four percent due to inefficiency. A similar pattern holds for the annual samples, although there is considerable year to year variation, with as much as ninety-five percent of mean square forecast error accounted for by the random component in some years. In principle, knowledge of the systematic error in analysts' forecasts permits the use of 'optimal linear correction' techniques in order to improve forecast accuracy. This involves employing the predicted values calculated using the estimated coefficients from regression (4), above, in place of the forecasts themselves. The effect of the ordinary least squares regression is to adjust the forecasts by compensating for their bias and inefficiency. The degree to which accuracy can be enhanced in this way depends upon the proportion of the mean square forecast error that is systematic. The results reported here imply that, assuming that the underlying data generating process for actual earnings growth and the method by which analysts form the expectations of earnings growth remain constant, optimal linear correction of the forecasts will reduce the forecast error only by about twelve percent. This is clearly an important result for the users of analysts' forecasts.

The second decomposition divides the mean square forecast error into the error in forecasting average earnings growth in the economy, the error in forecasting the deviation of average growth in each industry from average growth in the economy, and the error in forecasting the deviation of earnings growth for

Table 3

Forecast Error Decomposition

	Panel A : Decomposition by Error Type			Panel B: Decomposition by Level of Aggregation		
	Bias	*Inefficiency*	*Random*	*Economy*	*Industry*	*Firm*
Pooled sample	7.51	4.07	88.45	9.21	35.53	55.25
1982	17.67	15.41	67.23	17.67	46.06	36.27
1983	4.37	2.12	93.92	4.37	40.21	55.42
1984	2.38	4.64	93.34	2.38	52.27	45.34
1985	6.07	6.68	87.57	6.07	36.45	57.48
1986	8.00	2.96	89.37	8.00	40.59	51.41
1987	16.73	1.86	81.69	16.73	30.15	53.11
1988	14.10	2.04	84.13	14.10	29.77	56.13
1989	20.02	5.32	74.89	20.02	27.45	52.53
1990	9.62	4.49	86.13	9.62	31.68	58.69
1991	3.35	2.63	94.27	3.35	33.05	63.60
1992	0.26	4.78	95.24	0.26	32.13	67.61

Notes:
Panel A reports the results of the decomposition of mean square forecast error for each year t by error type, given by:

$$\text{MSFE} = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2 = (\overline{g}_t - \overline{g}_t^f)^2 + (1-\beta_t)^2\sigma^2_{g^f t} + (1-\rho_t^2)\sigma^2_{gt}$$

where N_t is the sample size in year t, g_{it} is five year earnings growth from January year t to December year t+4, g_{it}^f is the median forecast of g_{it} reported in April of year t, $\overline{g}_t$ and $\overline{g}_t^f$ are the average values of g_{it} and g_{it}^f, β_t is the slope coefficient reported in Panel A of Table 2, ρ_t is the correlation coefficient between g_{it} and g_{it}^f, and σ^2_{gt} and $\sigma^2_{g^f t}$ are the variances of g_{it} and g_{it}^f. The decomposition for the pooled sample is computed over all firms and years.

Panel B reports the results of the decomposition of mean square forecast error for each year t by the level of aggregation, given by:

$$\text{MSFE} = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2$$

$$= (\overline{g}_t - \overline{g}_t^f)^2 + \frac{1}{N_t}\sum_{j=1}^{J_t}N_{jt}[(\overline{g}_{jt} - \overline{g}_t) - (\overline{g}_{jt}^f - \overline{g}_t^f)]^2 + \frac{1}{N_t}\sum_{i=1}^{N_t}[(g_{it} - \overline{g}_{jt}) - (g_{it}^f - \overline{g}_{jt}^f)]^2$$

where J_t is the number of industries in the sample, N_{jt} is the number of firms in industry j, $\overline{g}_{jt}$ and $\overline{g}_{jt}^f$ are the average values of g_{it} and g_{it}^f in industry j. The decomposition for the pooled sample is the weighted average of the decompositions for the annual samples, with weights proportional to the sample size each year. The table reports each of the components of mean square forecast error as a percentage of total mean square forecast error.

individual firms from average industry growth. The results of this decomposition are reported in Panel B of Table 3. The results demonstrate that analysts' forecast inaccuracy derives mainly from an inability to forecast deviations of individual firm growth from the average growth rate in its industry. The error in forecasting deviations of industry growth from the average growth rate in the economy is also important, but somewhat smaller than the error in forecasting individual firm growth. In contrast, analysts' inability to forecast average earnings growth in the economy contributes relatively little to their inaccuracy. An interesting feature of this decomposition is that the proportion of forecast error generated at the industry level appears to be diminishing over time, while the proportion generated at the individual firm level is increasing. This is potentially related to changes in the methods used by analysts to forecast earnings growth, or changes in accounting standards.

(vi) The Performance of Analysts' Forecasts Conditional on Firm and Forecast Characteristics

The foregoing analysis has considered analysts' long run earnings growth forecasts as a homogenous group. However, it is likely that forecast performance will vary with the characteristics of the firm whose earnings are being forecast. For instance, one would expect that firms with highly variable cash flows, or those for which little information is available about future earnings prospects, would be associated with lower forecast accuracy. Additionally, forecast performance is likely to vary with the size of the forecast itself since the efficiency results indicate that low forecasts are less overly-optimistic than high forecasts.

In order to investigate this issue, the accuracy, bias and efficiency results are reproduced for sub-samples of companies, partitioned on the basis of market capitalisation, price-earnings ratio, market-to-book ratio and the level of the forecast itself. For each variable, the sample is sorted into ascending order of the partitioning variable and split into quintiles, with equal numbers of firms in each quintile.[10] For all the results of this section, results are reported for quintiles pooled across all years only.

Table 4 presents the results for forecast accuracy, with the mean square forecast error for each quintile reported in Panel A.

There is substantial variation in forecast accuracy across market capitalisation, price-earnings ratio and forecast earnings growth, while there is no obvious systematic variation in forecast accuracy across market-to-book. Forecast accuracy increases with market capitalisation, with forecasts for the quintile of largest firms more than twice as accurate as those for the quintile of smallest firms. There is an inverse relationship between forecast accuracy and price-earnings ratio, with forecasts for the lowest quintile almost three times as accurate as those for the highest quintile. The largest variation in forecast accuracy is with the level of the forecast itself, with low forecasts being five times more accurate than high forecasts. In all three cases, variation in forecast accuracy is monotonic (almost monotonic in the case of price-earnings and forecast size), although it does not appear to be linear, with the largest differences occurring in the lowest and highest quintiles.

The results of Panel A show that forecast accuracy varies substantially with market capitalisation, price-earnings ratio and the forecast itself. However, these variables are not independent, and so variation in forecast accuracy with one variable may merely reflect variation with another. In order to identify the marginal effects of firm and forecast characteristics on forecast accuracy, Panel B of Table 4 reports the regression of the squared forecast error on the natural logarithm of market capitalisation, market-to-book, price-earnings and forecast earnings growth. Interestingly, all four variables independently contribute to the explanation of forecast accuracy, with the most influential, in terms of statistical significance, being the price-earnings ratio, followed by the level of the forecast itself. The most accurate forecasts are therefore low forecasts issued for large companies with low price-earnings ratios and high market-to-book ratios. The four variables together explain more than thirteen percent of the variation in forecast accuracy.

The variation of forecast accuracy with market capitalisation is not surprising. Information about future earnings prospects is likely to be more readily available, and of a higher quality, for larger firms. The variation of forecast accuracy with the forecast itself is consistent with the results on forecast efficiency. The inverse relationship between forecast accuracy and price-earnings ratio is harder to explain, but may be driven by the fact that very

Table 4

Forecast Accuracy Conditional on Firm and Forecast Characteristics

Panel A: Forecast Accuracy by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	11.52	8.24	6.35	5.19	4.47
Market-to-Book	7.84	6.51	6.36	7.18	7.88
Price-Earnings	5.30	4.53	5.02	6.13	14.79
Forecast Size	2.77	6.56	5.70	7.46	13.38

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Accuracy

	Estimated Coefficient	*Standard Error*
Capitalisation	−103.18	(14.39)
Market-to-Book	−17.02	(6.80)
Price-Earnings	24.47	(3.55)
Forecast Growth	42.67	(6.17)
$\overline{R}^2$	0.13	

Notes:
Panel A reports the MSFE in percent for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth.

Panel B reports the estimated slope coefficients from the regression:

$$(g_{it} - g_{it}^f)^2 = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t + 4$, g_{it}^f is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t - 1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t - 1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

high price-earnings ratios arise partly as a result of very low, but transitory earnings, the trajectory of which is likely to be difficult to forecast accurately. The positive relationship between forecast accuracy and market-to-book ratio is potentially explained by the fact that high market-to-book companies, *ceteris paribus*, should on average have high earnings growth. Since forecast earnings growth is generally too optimistic, the size of the forecast error for these companies should on average be lower.

Table 5 presents the results for forecast bias. Again, there is strong variation in forecast bias with market capitalisation, price-earnings ratio and the level of the forecast itself. Consistent with the results for forecast accuracy reported in Table 4, forecast bias decreases (in absolute value) with market capitalisation and increases with forecast size. However, while forecast inaccuracy increases with price-earnings ratio, forecast bias *decreases* with price-earnings ratio, implying that while forecasts become less biased as the price-earnings ratio increases, they nevertheless become less accurate. However, this merely implies that the random component of forecast inaccuracy decreases more rapidly with price-earnings ratio than does the systematic component. The largest variation in forecast bias is again with forecast size, with forecasts in the highest quintile being more than four times as biased as those in the lowest quintile. This is consistent with the results on efficiency reported earlier that demonstrate a significant negative relationship between forecast error and the level of the forecast. There is some variation in forecast bias with market-to-book value of equity, although it is not monotonic across quintiles, and the difference between the lowest and highest quintile is not large. There is no quintile of companies for which it can be concluded that analysts' forecasts are unbiased.

Panel B reports the results of the regression of forecast error on market capitalisation, market-to-book value of equity, price earnings ratio and forecast earnings growth. There is again independent variation in forecast bias with market capitalisation, price-earnings ratio and the level of the forecast itself, with the latter being the strongest factor, statistically speaking. There is no significant variation with market-to-book. The four variables together explain about six percent of the variation in forecast error.

These results are broadly consistent with Frankel and Lee (1996), who investigate the performance of analysts' shorter horizon forecasts in order to operationalise an accounting valuation model based on book value of equity and the market's expectation of earnings growth. They find that analyst over-optimism is associated with low book-to-price ratio (the inverse of the market-to-book ratio used in the present analysis) and high past sales growth. They also find that analyst over-optimism is

Table 5

Forecast Bias Conditional on Firm and Forecast Characteristics

Panel A: Forecast Bias by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	−12.28	−8.15	−5.99	−5.34	−5.00
	(0.87)	(0.75)	(0.67)	(0.60)	(0.56)
Market-to-Book	−5.32	−6.35	−8.61	−8.08	−8.38
	(0.75)	(0.68)	(0.65)	(0.70)	(0.73)
Price-Earnings	−11.66	−6.87	−7.42	−5.48	−5.32
	(0.54)	(0.55)	(0.58)	(0.66)	(1.04)
Forecast Size	−3.98	−3.56	−5.49	−7.59	−16.12
	(0.44)	(0.69)	(0.64)	(0.71)	(0.90)

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Bias

	Estimated Coefficient	*Standard Error*
Capitalisation	0.76	(0.28)
Market-to-Book	0.05	(0.05)
Price-Earnings	0.23	(0.05)
Forecast Growth	−0.93	(0.09)
$\overline{R}^2$	0.06	

Notes:

Panel A reports the MFE in percent for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth. Standard errors are reported in parentheses.

Panel B reports the estimated slope coefficients from the regression:

$$(g_{it} - g_{t}^{f})^2 = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^{f} + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t+4$, g_{it}^{f} is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t-1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t-1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

associated with forecasts that are high relative to the current level of earnings (i.e. optimistic forecasts). Since forecast earnings growth and actual earnings growth are largely uncorrelated in the present sample, this is consistent with the finding reported above that analyst over-optimism is associated with high forecast earnings growth.

Table 6

Forecast Efficiency Conditional on Firm and Forecast Characteristics

Panel A: Forecast Efficiency by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	0.01	0.25	0.12	0.56	1.15
	(0.10)	(0.09)	(0.09)	(0.12)	(0.13)
Market-to-Book	0.05	0.01	0.00	−0.08	0.28
	(0.14)	(0.12)	(0.11)	(0.11)	(0.09)
Price-Earnings	−0.31	0.24	0.08	−0.04	−0.21
	(0.09)	(0.10)	(0.11)	(0.12)	(0.11)
Forecast Size	0.84	0.59	0.57	0.60	−0.11
	(0.26)	(0.86)	(0.98)	(0.84)	(0.13)

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Efficiency

	Estimated Coefficient	*Standard Error*
Capitalisation	3.87	(2.30)
Market-to-Book	1.99	(1.14)
Price-Earnings	0.12	(0.63)
Forecast Growth	−12.47	(2.31)
$\overline{R}^2$	0.11	

Notes:

Panel A reports the estimate of β in the regression $g_{it} = \alpha_t + \beta g_{it}^f + u_{it}$ for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth. Froot-Newey-West adjusted standard errors are reported in parentheses.

Panel B reports the estimated slope coefficients from the regression:

$$(g_{it}^f - \overline{g}_t^f)[(g_{it} - \overline{g}_t) - (g_{it}^f - \overline{g}_t^f)] = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t+4$, g_{it}^f is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t-1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t-1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

Table 6 presents the results for forecast efficiency. Panel A reveals that there is considerable variation in forecast efficiency across both market capitalisation and the level of the forecast, with some variation across market-to-book. The estimated slope parameter, β, is close to zero for the quintile of smallest firms, and rises monotonically with firm size. For the quintile of largest firms, the efficiency condition that $\beta = 1$ cannot be rejected. The estimated slope parameter decreases with the level of forecast, and for the quintile of firms with the lowest forecasts, the null hypothesis that $\beta = 1$ cannot be rejected either. There is no systematic variation with price-earnings ratio. The most efficient forecasts are therefore low forecasts for large firms with high market-to-book ratios.

Panel B of Table 6 reports the marginal contribution of each of the independent variables to forecast efficiency. Consistent with results of Panel A, there is positive independent variation in forecast efficiency with market capitalisation and market-to-book ratio, although the significance is marginal. Also consistent with the quintile results, the relationship between forecast efficiency and forecast growth is very significantly negative. There is no significant variation in forecast efficiency with price-earnings ratio. The four variables together explain eleven percent of the variation in forecast efficiency.

5. SUMMARY AND CONCLUSIONS

This paper has undertaken a detailed study of the accuracy, bias and efficiency of analysts' forecasts of long run earnings growth for US companies. The results of the paper can be summarised as follows.

(i) The accuracy of analysts' long run earnings growth forecasts is extremely low. Superior forecasts can be achieved simply by assuming that long run earnings growth is zero.
(ii) Analysts' forecasts are excessively optimistic. Forecast earnings growth, on average, exceeds actual earnings growth by about seven percent per annum.
(iii) Analysts' forecasts are weakly inefficient. Forecast errors are not independent of the forecasts themselves. In particular, high forecasts are associated with high forecast errors, while low forecasts are associated with low forecast errors.
(iv) Analysts' forecasts do not incorporate all information contained in current share prices. A superior forecast can be obtained by assuming that each firm's earnings will

evolve in such a way that its price-earnings ratio will converge to the current market-wide price-earnings ratio.

(v) Despite the bias and inefficiency identified in (ii) and (iii) above, the systematic components of analysts' forecast errors contribute relatively little to their inaccuracy. More than eighty-eight percent of the mean square forecast error is random. This is an important result for the users of analysts' long run earnings growth forecasts, since it means that the accuracy of analysts' forecasts cannot be significantly improved using linear correction techniques.

(vi) The largest part of analysts' forecast error is made at the individual firm level. The inability of analysts to forecast average earnings growth in the economy does not contribute substantially to their inaccuracy. However, there is evidence that the level of aggregation at which analysts' errors are being made is changing over time, with increasing accuracy at the industry level, and decreasing accuracy at the individual firm level.

(vii) There is significant heterogeneity in the performance of analysts' forecasts. The most reliable earnings growth forecasts are low forecasts issued for large companies with low price-earnings and high market-to-book ratios. The least biased forecasts are those for low forecasts for companies with low price-earnings ratios, while the most efficient forecasts are low forecasts for large companies with high market-to-book ratios. This is again an important result for the users of analysts' forecasts since it offers some opportunity to discriminate between good and bad forecasts.

(viii) There is very little evidence to suggest that the inaccuracy, bias or inefficiency of analyst' forecasts have diminished over time.

The idea that analysts systematically make over-optimistic forecasts, is not necessarily an indictment of their rationality *per se* since they may have considerable incentives to do so. An earnings growth forecast is not generally the final product delivered by an analyst to the client. In particular, earnings growth forecasts will be typically provided as part of a package of services, including brokerage, advice on mergers and acquisitions, and underwriting, and these related activities may influence the forecasts that an analyst makes (see Schipper, 1991). Sell-side analysts, for instance, have a vested interest in their clients' reaction to earnings forecasts. If earnings forecasts are used to support stock recommendations then high forecasts will tend to generate more business than low forecasts, since there is a larger potential client base for buy recommendations than for sell recommendations. Francis and Philbrick (1993) provide evidence that suggests that analysts may be intentionally over-optimistic in order to cultivate and maintain good management relations.

The decomposition of mean square forecast error by error type revealed that by far the largest component of analysts' forecast errors is random, with the systematic component accounting for less than twelve percent. Inevitably, at such long forecasting horizons, the potential to make accurate forecasts of earnings growth is limited. However, the fact that such a large component of actual earnings growth is random may explain why analysts' forecasts are so biased. The larger the component of the forecast error that is random, the lower the impact of forecast bias on forecast error. Assuming that analysts do have conflicting objectives — one to produce *accurate* earnings growth forecasts, the other to produce *high* earnings growth forecasts — then if analysts know that the first objective is largely unattainable, they will use the forecasting process to satisfy the second. If analysts are also producing short term and interim forecasts for the same company, then the bias in their long term forecasts may be compounded.

A number of papers have now concluded that there is substantial mis-pricing in the stock market as a consequence of irrational long run earnings growth forecasts being incorporated into the market expectation of earnings growth. The results of this paper support the hypothesis that analysts' consensus long run earnings growth forecasts are indeed irrational if they are to be interpreted as optimal forecasts of future earnings growth. However, given the uncertainty over analysts' incentives, it is by no means inevitable that these forecasts will be incorporated without modification into the market expectation of earnings growth. An interesting topic for future research will be to examine to what extent the market recognises the characteristics in forecast long run earnings growth identified in this paper.

NOTES

1 A partial list would include Brown and Rozeff (1978), Brown et al. (1987a and 1987b) and O'Brien (1988) who consider the performance of analysts' quarterly earnings forecasts, and Collins and Hopwood (1980), Fried and Givoly (1982) and Brown et al. (1985), who consider analysts' annual forecasts. International evidence on analysts' forecasts is provided by Capstaff et al. (1995), who analyse the performance of UK analysts, and Capstaff et al. (1998), who consider the forecasts of European analysts. For a comprehensive survey of the literature on analysts' earnings forecasts, see Brown (1993).

2 This was confirmed in conversation with IBES staff.

3 The correlation between the mean and the median forecast in the sample is 0.98. This is accounted for by the fact that most stocks have long term forecasts originating from only one or two analysts.

4 IBES have confirmed that they do receive earnings growth forecasts for companies whose earnings are currently negative. This may be explained by the fact that while analysts use the latest reported earnings as a base for earnings growth when earnings are positive, they use some other unspecified base measure of earnings, such as forecast annual earnings or average historical annual earnings, when earnings are negative.

5 In order to establish the robustness of the results, the analysis was conducted using maximum earnings growth threshold values in the range 50% to 1,000%, and by trimming the sample instead on the basis of initial earnings per share, using a minimum earnings threshold of between 0.10 and 1.00 dollars. The sensitivity of the results to changes in the threshold values was low, and none of the qualitative conclusions were altered. The regressions were additionally estimated using the minimum absolute deviation estimator, which is considerably less sensitive to outliers. This produced results that were almost completely invariant with respect to the choice threshold values. As a further test of the robustness of the results, the analysis was conducted using the change in earnings scaled by price, with the corresponding forecast change in earnings computed using the forecast growth rate. The results of these robustness tests are not reported here, but are available from the author on request.

6 The average growth rate is taken over all firms for which earnings data are available, using the same sample selection criteria as for subsequent earnings growth, namely excluding observations for which earnings are negative at the beginning of the five year period, and those for which the calculated growth rate exceeds 100% in absolute value.

7 This can be seen by subtracting forecast earnings growth, g_{it}^f, from each side so that the regression becomes one of forecast error on forecast earnings growth — the constant remains the same while the slope parameter becomes $\beta-1$.

8 Taking the conditional expectation of equations (10) and (11) gives the mean square forecast error and the mean forecast error, respectively, as a function of the independent variables. Regressions (10) and (11) thus measure the marginal contribution of each of the independent variables to forecast accuracy and forecast bias. Taking the conditional expectation of equation (12) gives the covariance between $(g_{it} - g_{it}^f)$ and g_{it}^f as a function of the independent variables. This covariance is the numerator of the estimated slope coefficient in a regression of $g_{it} - g_{it}^f$ on g_{it}^f. Under the

null hypothesis that forecasts are weakly efficient, this covariance should be equal to zero. If it is less than zero, forecasts are too extreme, while if it is greater than zero, forecasts are too compressed. Regression (12) thus measures the marginal contribution of each of the independent variables to forecast efficiency.

9 See, for example, Brown et al. (1987a) and O'Brien (1988), who consider the accuracy of analysts' quarterly earnings forecasts relative to the forecasts of different time series models, and Fried and Givoly (1982), who consider the relative accuracy of analysts' annual earnings forecasts.

10 Except for the largest quintile, which has an additional observation.

REFERENCES

Abarbanell, J. (1991), 'Do Analysts' Earnings Forecasts Incorporate Information in Prior Stock Price Changes?' *Journal of Accounting and Economics*, Vol. 14, pp. 147–66.

Alford, A., J. Jones and M. Zmijewski (1994), 'Extensions and Violations of the Statutory SEC Form 10-K Filing Requirements', *Journal of Accounting and Economics*, Vol. 17, pp. 229–54.

Brooks, L. and D. Buckmaster (1976), 'Further Evidence on the Time Series Properties of Accounting Income', *Journal of Finance*, Vol. 31, pp. 1359–73.

Brown, L. (1993), 'Earnings Forecasting Research: Its Implications for Capital Markets Research', *International Journal of Forecasting*, Vol. 9, pp. 295–320.

________ (1997), 'Analyst Forecasting Errors: Additional Evidence', Working Paper (University at Buffalo).

________ and M. Rozeff (1978), 'The Superiority of Analysts' Forecasts as Measures of Expectations: Evidence from Earnings', *Journal of Finance*, Vol. 33, pp. 1–16.

________ R. Hagerman, P. Griffin and M. Zmijewski (1987a), 'Security Analyst Superiority Relative to Univariate Time-Series Models in Forecasting Quarterly Earnings', *Journal of Accounting and Economics*, Vol. 9, pp. 61–87.

________ ________ ________ ________ (1987b), 'An Evaluation of Alternative Proxies for the Market's Assessment of Unexpected Earnings', *Journal of Accounting and Economics*, Vol. 9, pp. 159–93.

Brown, P., G. Foster and E. Noreen (1985), 'Security Analyst Multi-Year Earnings Forecasts and the Capital Market', *American Accounting Association Studies in Accounting Research*, Vol. 21.

Bulkley, G. and R. Harris (1997), 'Irrational Analysts' Expectations as a Cause of Excess Volatility in Stock Prices', *Economic Journal*, Vol. 107, pp. 359–71.

Capstaff, J., K. Paudyal and W. Rees (1995), 'The Accuracy and Rationality of Earnings Forecasts by UK Analysts', *Journal of Business, Finance & Accounting*, Vol. 22, pp. 67–85.

________ ________ ________ (1998), 'A Comparative Analysis of Earnings Forecasting in Europe', manuscript (University of Glasgow).

Christie, A. (1987), 'On Cross-Sectional Analysis in Accounting Research', *Journal of Accounting and Economics*, Vol. 9, pp. 231–58.

Collins, W. and W. Hopwood (1980), 'A Multivariate Analysis of Annual Earnings Forecasts Generated from Quarterly Forecasts of Financial Analysts and Univariate Time-Series Models', *Journal of Accounting Research* Vol. 18, pp. 390–406.

DeBondt, W. (1992), *Earnings Forecasts and Share Price Reversals* (Foundation of Chartered Financial Analysts, Charlottsville, Virginia).

________ and R. Thaler (1990), 'Do Security Analysts Overreact?' *American Economic Review*, Vol. 80, pp. 52–57.

Dechow, P. and R. Sloan (1997), 'Returns to Contrarian Investment Strategies: Tests of Naive Expectations Hypotheses', *Journal of Financial Economics*, Vol. 43, pp. 3–27.

Elgers, P., and D. Murray (1992), 'The Relative and Complementary Performance of Analyst and Security-Price Based Measures of Expected Earnings', *Journal of Accounting and Economics*, Vol. 15, No. 303–16.

Elton, E., M. Gruber and M. Gultekin (1984), 'Professional Expectations: Accuracy and Diagnosis of Errors', *Journal of Financial and Quantitative Analysis*, Vol. 19, pp. 351–62.

Foster, G. (1977), 'Quarterly Accounting Data: Time Series Properties and Predictive-Ability Results', *Accounting Review*, Vol. 52, pp. 1–21.

Francis, J. and D. Philbrick (1993), 'Analysts' Decisions as Products of a Multi-Task Environment', *Journal of Accounting Research*, Vol. 31, pp. 216–30.

Frankel, R. and C. Lee (1996), 'Accounting Valuation, Market Expectation, and the Book-to-Market Effect', *Journal of Finance*, Vol. 51, pp. 1040–41.

Fried, D. and D. Givoly (1982), 'Financial Analysts' Forecasts of Earnings', *Journal of Accounting and Economics*, Vol. 4, pp. 85–107.

Froot, K. (1989), 'Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data', *Journal of Financial and Quantitative Analysis*, Vol. 24, pp. 333–55.

Givoly, D. (1985), 'The Formation of Earnings Expectations', *Accounting Review*, Vol. 60, pp. 372–86.

Lakonishok, J., A. Schleifer and R. Vishny, (1994), 'Contrarian Investment, Extrapolation, and Risk', *Journal of Finance*, Vol. 49, pp. 1541–78.

La Porta, R. (1996), 'Expectations and the Cross-Section of Stock Returns', *Journal of Finance*, Vol. 51, pp. 1715–42.

Mincer, J. and V. Zarnovitz (1969), 'The Valuation of Economic Forecasts', in J. Mincer (ed.), *Economic Forecasts and Expectations* (National Bureau of Economic Research).

Newey, W. and K. West (1987), 'A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica*, Vol. 55, pp. 703–8.

O'Brien, P. (1988), 'Analysts' Forecasts as Earnings Expectations', *Journal of Accounting and Economics*, Vol. 10, pp. 53–83.

Ou, J. and S. Penman (1987), 'Financial Statement Analysis and the Prediction of Stock Returns', *Journal of Accounting and Economics*, Vol. 11, pp. 295–330.

Ramakrishnan, R. and J. Thomas (1992), 'What Matters from the Past: Market Value, Book Value or Earnings? Earnings Valuation and Sufficient Statistics for Prior Information', *Journal of Accounting, Auditing and Finance*, Vol. 7, pp. 423–64.

Schipper, K. (1991), 'Commentary on Analysts' Forecasts', *Accounting Horizons*, Vol. 5, pp. 105–21.

Theil, H. (1964), *Optimal Decision Rules for Government and Industry* (North Holland).

________ (1966), *Applied Economic Forecasting* (North Holland).