WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

ANALYSTS' FORECASTS AS EARNINGS EXPECTATIONS

Patricia C. O'Brien

WP# 1743-86   revised March 1987

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

ANALYSTS' FORECASTS AS EARNINGS EXPECTATIONS

Patricia C. O'Brien

WP# 1743-86   revised March 1987

ANALYSTS' FORECASTS AS EARNINGS EXPECTATIONS

Patricia C. O'Brien

Massachusetts Institute of Technology

I examine three composite analyst forecasts of earnings per share as proxies for expected earnings. The most current forecast weakly dominates the mean and median forecasts in accuracy. This is evidence that forecast dates are more relevant for determining accuracy than individual error.

Consistent with previous research, I find analysts more accurate than time-series models. However, prior knowledge of forecast errors from a quarterly autoregressive model predicts excess stock returns better than prior knowledge of analysts' errors. This is inconsistent with previous research, and is anomalous given analysts' greater accuracy.

Sloan School of Management
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA   02139

1. Introduction

Analysts' forecasts of earnings are increasingly used in accounting and finance research as proxies for the unobservable "market" expectation of a future earnings realization. Since a diverse set of forecasts is available at any time for a given firm's earnings, composites are used to distill the diverse set into a single expectation. This paper considers the relative merits of several composite forecasts as expectations data. The primary result is that the most current forecast available is more accurate than either the mean or the median of all available forecasts. This suggests forecast timeliness as a characteristic for distinguishing better forecasts. A second and related result is that, conditional on only relatively recent forecasts being included, means or medians increase accuracy by aggregating across idiosyncratic individual error.

A second contribution of this paper is a comparison of analysts with time series models, a competing source of earnings expectations. Consistent with previous research, I find analysts generally more accurate than time-series models. However, I find that prior knowledge of the forecast errors from a simple autoregressive model on the univariate series of quarterly earnings provides better predictions of excess stock returns than prior knowledge of analysts' forecast errors. This result is inconsistent with prior research, and somewhat anomalous given analysts' greater accuracy.

A third contribution is a methodological refinement of techniques used to evaluate forecasts. I demonstrate the existence of significant time-period-specific effects in forecast errors. If time-series and cross-section data are pooled without taking these effects into account, statistical results may be overstated, and results are subject to an aggregation bias.

In section 2, I describe proxies for consensus in analysts' forecasts, as well as two quarterly time-series models used in the empirical tests. In section 3 I describe the data. The statistical tests and results are discussed in section 4, and section 5 is a summary with some concluding remarks.

2. Proxies for Expected Earnings

2.1 Defining Consensus for Analysts' Forecasts

The motivations for seeking a consensus expectation when many forecasts are available are primarily practical. In many contexts earnings expectations are not the central issue, but are necessary data. For example, to remove the effects of simultaneous earnings releases on non-earnings events, unanticipated earnings are necessary. Reducing measurement error in anticipated earnings, and therefore in unanticipated earnings, increases the power of tests of the non-earnings event. If individual forecasts contain idiosyncratic error which can be

diminished by aggregation. more accurate forecasts can be obtained by combining forecasts.[1]

Academic researchers have used a variety of methods to aggregate analysts' earnings forecasts into a single expectation.[2] Barefield and Comiskey (1975), and Fried and Givoly (1982) use the mean of a set of forecasts. Brown and Rozeff (1978) and Brown, Griffin, Hagerman, and Zmijewski (1986) use a single forecast from ValueLine. Givoly and Lakonishok (1979) select the "most active" forecaster for each firm from those available in Standard and Poors' Earnings Forecaster. Elton, Gruber, and Gultekin (1981) and Brown, Foster and Noreen (1984) consider both means and medians published in the I/B/E/S Summary database. Brown, Richardson, and Schwager (1986) use both the ValueLine forecasts and the I/B/E/S Summary forecasts. I compare three composites from a set of available forecasts: the mean, the median, and the most current forecast.

An implicit assumption behind the use of either the mean or the median forecast to represent consensus is that all forecasts are current, so that cross-sectional differences in forecasts are attributable to differential use of the same global set of information. Gains from combining forecasts arise either from the employment of more information in the aggregate than is used by any individual, or from diversification across individuals' idiosyncratic errors.

In fact, however, the analysts' forecasts available at any time have been produced at varying dates. If analysts

incorporate new information into their forecasts as the year progresses as previous research suggests (Crichfield, Dyckman and Lakonishok (1978), Collins and Hopwood (1980), Elton, Gruber, and Gultekin (1982)), then more current forecasts are expected to be more accurate than older ones. However, if diversifying across individual idiosyncracies is more important than eliminating out-dated forecasts, aggregations of many forecasts, regardless of date, may be more accurate than a single current one. I provide a comparison of the relative importance of forecast age versus diversification, by examining the most current forecast as an alternate to the mean and median definition of consensus.

My results indicate that the single most current forecast dominates aggregations which ignore forecast dates. When only relatively recent forecasts are included in the aggregations, it is possible to increase accuracy by aggregating forecasts.[3] Since the aggregate forecasts in published databases (e.g. the I/B/E/S Summary data and the Zacks Investment Research data) ignore forecast dates, my results are relevant for researchers using these sources.

2.2 Quarterly Time-series Models of Earnings

I use quarterly time-series models of earnings as benchmarks, against which analysts' forecasts are compared. Time-series models have been used frequently in previous research to provide earnings expectations. Analysts, however, have the

advantage of a broader information set, including industry and firm sales and production figures, general macroeconomic information, and other analysts' forecasts, in addition to the historical series of earnings. Analysts' forecasts, therefore, seem likely to be more accurate than forecasts from univariate models.

Several studies (Brown and Rozeff (1978), Collins and Hopwood (1980), Fried and Givoly (1982)) demonstrate that analysts are more accurate than univariate models, presumably because of the broader information set they can incorporate. Fried and Givoly (1982) also find that analysts' forecast errors are more closely associated with excess stock returns than are those of univariate models. Nonetheless, univariate models remain a common means of generating earnings expectations.

An advantage of univariate time-series models is the relative ease with which earnings data can be obtained for moderate samples of firms. This advantage is tempered by the caveat that the data requirements of the models impart a "survivorship" bias to samples. Another advantage to time-series models is the relative simplicity of the models used to generate expectations. Parsimonious models with a single, simple ARIMA structure applied to all firms have been shown to predict at least as well as univariate models with individually-specified structures, when one-step-ahead forecast errors are compared (Foster (1977), Watts and Leftwich (1977)).

I use the following two quarterly time-series models, from Foster (1977):

$$E[a_{jtq}] = a_{jtq-4} + \theta_{j0} + \theta_{j1}(a_{jtq-1} - a_{jtq-5}) \quad , \qquad (1)$$

and:

$$E[a_{jtq}] = a_{jtq-4} + \theta_{j2} \qquad , \qquad (2)$$

where a_{jtq} denotes quarterly reported earnings for firm j in quarter q of year t, and θ_{j0}, θ_{j1} and θ_{j2} are estimated parameters. The models are, respectively, a first order autoregressive process in fourth differences with a drift, and a random walk in fourth differences with a drift. I chose these models because of their relative simplicity, and because they have proven to be at least as good as other mechanical quarterly models. The data used and the estimation of these models are described in section 3.

3.   Data

3.1   Sources

The forecast data are from the Institutional Brokers Estimate System, or I/B/E/S, developed by the Lynch, Jones & Ryan brokerage house. The database consists of individual analysts' forecasts of earnings per share (EPS)[4] made between early 1975 and mid-1982, by analysts at between 50 and 130 brokerage houses. EPS are forecast for approximately 1000 to 2500 firms, depending on the month and year in question. The individual forecasts are

used by I/B/E/S to compute summary data, such as means, medians
and standard deviations. The summary statistics are sold to
clients, primarily institutional investors. The I/B/E/S Summary
data have been analyzed extensively by Brown, Foster and Noreen
(1984), among others.

Each brokerage house in the database employs many analysts,
but at most one forecast is reported from each brokerage house at
any time, for a given firm and year. Analysts and brokerage
houses are identified in the database by code numbers. The
I/B/E/S data are updated once per month with new forecasts. I
use primarily two pieces of information: individual analysts'
forecasts, and their associated forecast dates. My method of
selecting the sample of available forecasts for a given firm and
year, described more fully in section 3.2, differs from that used
by Lynch, Jones & Ryan to produce the Summary data. The most
important difference is that I use analysts' forecast dates, not
I/B/E/S publication dates, to define which forecasts are
available on a given day.

COMPUSTAT is the source of earnings data and most earnings
announcement dates. The remaining announcement dates are from
the Wall Street Journal (WSJ) and its Index. Stock return data
and the trading day calendar are from the CRSP Daily Returns
file. Data on stock splits and stock dividends are from the CRSP
Monthly Master file.

Some analysts occasionally forecast fully-diluted EPS,
rather than primary EPS. This is indicated in the I/B/E/S detail

data. I convert forecasts of fully-diluted EPS to primary EPS,
using the reported ratio of fully-diluted to primary EPS for that
firm and year from COMPUSTAT. I also adjust both analysts' and
time series model forecasts for any stock splits and stock
dividends announced between the forecast date and the annual
earnings announcement date.

3.2 Sample Selection

The sample selection criteria and effects on the sample size
are summarized in Table 1. The initial sample comprises the set
of firms in the I/B/E/S database with December yearends, and with
at least one forecast available in each year from 1975 through
1981. This set contains 508 firms, and 3556 firm-years. A firm-
year is excluded if the annual earnings number is not available
on COMPUSTAT, or if all four quarterly earnings announcement
dates are not available from COMPUSTAT or the WSJ. This
criterion reduces the sample to 497 firms, with 3440 firm-years.
The requirement that returns data be available on CRSP reduces
the sample to 410 firms. The estimation requirements of the
quarterly models, 30 continuous quarters of data prior to 1975-
IV, impose the most drastic reduction in the sample, to 184 firms
with 1260 firm-years.

Forecasts for each firm and year are selected at five fixed
horizons of less than one year. The horizons are: 240, 180,
120, 60 and 5 trading days prior to the announcement of annual

earnings. The first four horizons correspond roughly to dates following each of the year's quarterly earnings announcements; the fifth is immediately prior to the annual earnings announcement. For example, a horizon of 240 trading days usually corresponds to a date after the previous year's annual announcement, but before the current year's first quarter announcement. A horizon of 180 trading days typically corresponds to a date between the first quarter announcement and the second quarter announcement; and so on.[5]

I define the set of available analysts' forecasts for each horizon, firm and year as follows. The horizon date for horizon τ for firm j in year t is the date τ trading days prior to the announcement of firm j's year t EPS. Given a horizon date, I select the most recent forecast available from each brokerage house forecasting firm j's EPS for year t. The number of available forecasts increases as the horizon shrinks, as reported in Panel B of Table 1. Some brokerage houses issue forecasts before the start of the year, and update them periodically during the year. Many others add forecasts as the annual EPS announcement approaches.

To determine this set of available forecasts, I use the dates assigned to forecasts by the analysts, not the dates of I/B/E/S' first publication of the forecasts. The publication lag, or time between the analyst's forecast date and the date of the forecast's first appearance on I/B/E/S, averages 34 trading days, and has a standard deviation of 44.5 trading days. Thus,

some recently-updated forecasts are omitted from each monthly
listing by I/B/E/S.

From the set of available forecasts for each firm j, year t,
and horizon τ, I compute the mean and median, and find the most
current. I use the mean, median and most current as proxies for
the analysts' consensus. Since published means and medians are
computed from the monthly lists, the I/B/E/S Summary data fail to
reflect some recent updates. Eliminating the publication lag
probably results in my consensus analyst forecasts being somewhat
more accurate than those published in the Summary data.

Most previous studies of analysts' forecasts have used
publication dates, not analysts' dates, to select forecasts. For
example, Fried and Givoly (1982) and Givoly and Lakonishok (1979)
select their samples based on the publication date of the
Standard & Poors' Earnings Forecaster, their source of forecast
data. Fried & Givoly go on to use analysts' dates within that
sample to distinguish new and old forecasts. Brown and Rozeff
(1978), Brown, Foster and Noreen (1984), Brown, Griffin, Hagerman
and Zmijewski (1986) and Brown, Richardson and Schwager (1986)
use datasets for which individual analysts' forecast dates are
not available. Using publication dates instead of forecast dates
probably biases results against analysts, by failing to include
some recent updates of forecasts.

In spite of eliminating the publication lag, for any given
horizon date many of the forecasts available were made prior to
the last quarterly earnings announcement. This may indicate

analysts' failure to incorporate new information, but need not.
For example, announced quarterly EPS may be close to the
analyst's expectation, so little new information is conveyed by
the quarterly announcement and a revision of the annual EPS
forecast is unnecessary. I investigate a subsample consisting of
only those forecasts which have been updated since the most
recent announcement of quarterly EPS. The subsample is described
in section 4.4. where the results of this investigation are
reported.

3.3 Measuring Forecast Errors

Forecast errors are the elementary data I use to evaluate
forecasts. The forecast error $e_{ijt\tau}$ is defined as the difference
between A_{jt}. actual earnings per share of firm j in year t, and
$f_{ijt\tau}$. the forecast of EPS from source i, at a horizon τ prior to
the realization:[6]

$$e_{ijt\tau} = A_{jt} - f_{ijt\tau} \qquad . \qquad (3)$$

The source of the forecast, denoted by i, is one of the
following: the mean, the median, or the most current of available
analysts' forecasts; or one of the two benchmark quarterly models
described in section 2.2.

The fundamental difference between the most current analyst
forecast as a consensus definition and either the mean or the
median, is that the former is constructed using the forecast

date, while the latter two are not. In Table 2, I compare distributions of forecast ages in this sample. The age of a forecast is defined as the difference, in trading days, between the forecast date and the horizon date chosen for this study. More generally, this might correspond to a lag between the forecast date and an event date of interest to the researcher. For Table 2, I define the ages of the mean and median forecasts as, respectively, the mean and the median of the ages of the forecasts in the set of available forecasts for each firm, year and horizon. The distribution described in Table 2 is over all firms and years, for each horizon.

As expected from its definition, the most current forecast has a distribution of ages much closer to zero than either the mean or median. For the four longer horizons (240, 180, 120, and 60 trading days), over fifty percent of the most current forecasts are less than five trading days old. By contrast, over fifty percent of the mean or median ages at all horizons are more than forty trading days old. While some of these older forecasts may represent circumstances where little new information has arrived, so there was no need to update, the accuracy results which follow suggest that this is not always the case.

Quarterly models (1) and (2) are estimated for each firm, for each quarter from 1975-I through 1981-IV. Parameter estimates are updated each quarter, using the previous thirty quarters' observations. Observations are adjusted for changes in the number of outstanding shares. Annual forecasts are

constructed from quarterly forecasts by summing the appropriate

realizations and forecasts. For example, in quarter 3, there

have been two realizations of quarterly earnings for the year,

and two quarters remain to be forecast. The annual forecast from

a quarterly model during the third fiscal quarter is the sum of

actual earnings for quarters 1 and 2, and forecasts for quarters

3 and 4.

A small number of influential observations altered the

regression results reported below.[7] Since analysts and brokerage

houses are identified only by code numbers in the database, there

was no way to trace these observations to other sources. For

this reason, I imposed an arbitrary censoring rule on the data

for errors which could not be traced: forecast errors larger

than $10.00 per share in absolute value were deleted from the

sample. Since typical values for EPS numbers are in the range of

$1.00 to $5.00 per share, errors of sufficient magnitude to be

deleted are rare, and suggest a data entry or transcription

error.

3.3 Stock Returns

I measure the new information impounded in stock returns by

the cumulated prediction errors from a market model in

logarithmic form:

$$E[\ln(1 + R_{js})] = \alpha_j + \beta_j \ln(1 + R_{Ms}) \quad , \quad (4)$$

where R_{js} is the return to security j on day s, R_{Ms} is the return

on the CRSP equally-weighted market portfolio of securities on day s, and ln denotes the natural logarithm transformation.[8]

The parameters of (4) are estimated for each firm in the study using 200 trading days of data at a time, beginning in July 1974. Estimated parameters are used to predict ahead 100 trading days, and excess returns are the difference between the realization $\ln(1 + R_{js})$ and the prediction based on (4). After each iteration of estimation and prediction, the estimation period is rolled forward by 100 trading days, and new parameters are estimated.

The estimation procedure produces a stream of predicted daily excess returns, ε_{js}. The ε_{js} are cumulated over each forecast horizon, from the horizon date through the announcement date of annual EPS, to form $U_{jt\tau}$, the measure of new information arriving over horizon τ in year t for firm j.

4.   Results

4.1 Aggregating Forecast Errors

If a forecast incorporates all the information available on the forecast date in an unbiased manner, it is an expectation in the usual statistical sense of the word. Let f^* denote such a forecast:

$$f^*_{jt\tau} = E[A_{jt} | \phi_{t\tau}] \quad . \qquad (5)$$

where $\phi_{t\tau}$ represents the information available at a horizon τ prior to the realization, and $E[\cdot|\cdot]$ is the conditional expectation operator.

Between the forecast date and the realization date, new information may arrive. Even a forecast like (5), which may be ideal in the sense of employing all information available on the forecast date, omits unanticipated information which arrives later. Forecast errors consist, in part or entirely, of new information revealed over the forecast horizon, i.e. between forecast and realization.

Two closely-related implications of unanticipated information reflected in forecast errors are important for the specification of statistical tests. First, forecast errors within a year which are aggregated cross-sectionally may appear to be "biased" because of the common new information reflected in them. Second, if this common information is not accounted for, it will induce contemporaneous cross-sectional correlation in forecast errors.

An example of information which may be reflected in forecast errors is an unanticipated macroeconomic shock affecting many firms in a similar manner. If the effect of the shock on firms has a non-zero mean, then a cross-sectional aggregation of forecast errors, even from unbiased forecasts, will also have a non-zero mean. This non-zero mean is not bias, but rather is time-period-specific new information. If time-period-specific

effects are ignored, they induce correlation in forecast errors across firms, for a given year and horizon, and across horizons within a year.

If forecast errors are positively correlated across firms within years, statistical comparisons based on pooled time-series and cross-section forecast error data which assume cross-sectional independence will overstate the statistical validity of the results. Several studies (Brown and Rozeff (1978), Elton, Gruber and Gultekin (1981), Malkiel and Cragg (1982), for example) have compared forecasts using criteria such as the number or proportion of times that one forecasting method outpredicts another. This criterion, or any other that assumes independent observations and is applied to a cross-section, could obtain the appearance of statistically significant superiority in forecasting ability from an anecdotal difference.

The tests developed in this paper adjust for time-period-specific shocks using a simple fixed effects model. This model, and its importance for the results, are described below.

4.2 Evaluating Forecasts – Bias

A simple model of time-period effects in forecast errors is:[9]

$$e_{jt\tau} = \mu_{t\tau} + \eta_{jt\tau} \qquad , \qquad (6)$$

where $\mu_{t\tau}$ is the average forecast error for year t and horizon τ,

and $\eta_{jt\tau}$ is a random error term, representing the deviation of firm j's forecast error from the common annual mean.

There may also be firm-specific information effects which persist through time, but the unanticipated information argument does not apply.[10] If a forecast fully impounds information from previous mistakes in an unbiased manner, systematically recurring events will not remain unanticipated year after year. Thus recurrent firm-specific forecast errors are not expected to arise on the basis of information that was unavailable at the time the forecast was made.

I estimate (6) using least squares with a dummy variable for each year. The test for bias is based on the grand mean of the estimated annual averages:

$$\bar{\mu}_\tau = \frac{1}{T} \sum_{t=1}^{T} \hat{\mu}_{t\tau} \qquad , \qquad (7)$$

where T is the number of years in the sample, and the $\hat{\mu}_{t\tau}$ are the year-specific average forecast errors, estimated separately for each horizon τ. The average $\bar{\mu}_\tau$ defined in (7) is a linear combination of least-squares coefficients, with estimated standard error:

$$s_{\bar{\mu}_\tau} = \frac{s_\eta}{T} \cdot [\underline{\iota}'(X'X)^{-1}\underline{\iota}]^{1/2} \qquad , \qquad (8)$$

where $\underline{\iota}$ is a vector of ones of length T, X is the matrix of dummy variables, and s_η is the regression residual standard error. The standard error (8) is a weighted version of the residual standard

error s_η, constructed from the $\eta_{jt\tau}$. The residuals $\eta_{jt\tau}$ are deviations from the annual averages, and so are purged of average time-period-specific information which induces cross-sectional correlation in the $e_{jt\tau}$.

In Table 3, the bias results are presented. The reported numbers are the forecast bias, estimated jointly for all forecast sources by stacking equation (7) across sources. The ratio of (7) to (8) is evaluated as a t-statistic, against a null hypothesis of no bias, for each analyst composite, mean, median and most current, and for the quarterly models.

Generally, forecast errors exhibit statistically significant negative bias. Of the three analyst consensus measures, the median uniformly exhibits the smallest bias, usually indistinguishable from zero.[11]

Negative bias corresponds to overestimates of EPS. Negative bias in analysts' forecasts is consistent with some conventional wisdom, that analysts prefer optimistic predictions and "buy" recommendations, to help maintain good relations with management.[12] The evidence supporting this story is weak, however, in two respects. First, the median analyst forecast appears to be unbiased. Second, and more importantly, when the analyst estimates are significantly negative, they are statistically indistinguishable from those of mechanical time-series models. The motive of maintaining good relations with management cannot be ascribed to these models. Thus, support for the contention that analysts preferentially issue optimistic forecasts is at best weak.

An alternative explanation which is also consistent with these results is that analysts issue unbiased forecasts, but this seven-year period, 1975 through 1981, is one with primarily negative unanticipated EPS. Unfortunately, the most obvious way to distinguish between the hypothesis of deliberate optimistic bias and this alternative is to collect data for a longer span of years. This is not possible using the I/B/E/S detail data.

4.3   Evaluating Forecasts - Accuracy

I use an approach similar to the bias evaluation described in the previous section for evaluating relative forecast accuracy. Accuracy is defined as absolute or squared forecast error. For the absolute error case, the model is:

$$|e_{jt\tau}| = \delta_{1j\tau} + \delta_{2t\tau} + \xi_{jt\tau} \qquad . \qquad (9)$$

For the squared error case, a similar model is estimated, with $(e_{jt\tau})^2$ as the left-hand-side variable. In equation (9) the $\delta_{1j\tau}$ measure average accuracy for each firm j, and the $\delta_{2t\tau}$ measure average accuracy for each year t. The $\xi_{jt\tau}$ are deviations from the average accuracy in this sample for firm j and for year t. Differences in accuracy across firms could arise, for example, if there are persistent differences in the amount of information available for different firms. Differences in accuracy across years could arise if there are more, or larger, unanticipated events in some years than in others.

Equation (9) is estimated using least squares on a set of dummy variables for firms and years. Average accuracy is computed as a linear combination of the estimated effects:[13]

$$\bar{\delta}_\tau = \frac{1}{J} \sum_{j=1}^{J} \hat{\delta}_{1j\tau} + \frac{1}{T} \sum_{t=1}^{T} \hat{\delta}_{2j\tau} \qquad . \qquad (10)$$

Equation (10) defines the average absolute error accuracy. Average squared error accuracy is defined similarly using the coefficients from the squared error version of (9). The estimated standard error of the average accuracy in equation (10) is:

$$s_{\bar{\delta}_\tau} = s_\xi \cdot [\underline{\omega}'(Z'Z)^{-1}\underline{\omega}]^{1/2} \qquad . \qquad (11)$$

In (11). Z represents the matrix of dummy variables used to estimate equation (9) or its squared error analogue, $\underline{\omega}$ is the vector of weights that transform the estimated parameters into the average defined in (10), and s_ξ is the residual standard error from the regression equation (9).

The estimates in equations (9) through (11) are computed jointly for all five forecast sources (mean, median and most current analyst; and two quarterly models), by stacking equations across sources. Pairwise differences in accuracy are compared across forecast sources using a t-statistic constructed from the average accuracies from (10) or its squared error analogue, and the standard error from (11).

Tables 4 through 6 summarize the results on forecast accuracy. The average absolute errors and average squared

21

errors, computed as described in equation (10) for each forecast source, appear in Table 4. Table 5 contains t-statistics testing pairwise differences in accuracy among analysts. Table 6 contains t-statistics testing pairwise differences in accuracy between analysts and the quarterly models.

Table 4 displays the expected pattern of increasing forecast accuracy as the earnings announcement date approaches, for all forecast sources. Both average absolute error and average squared error decline uniformly as the year progresses, for analysts and for quarterly models. For example, the average absolute error of the most current forecaster declines from $0.742 per share at a horizon of 240 days (almost a full year prior to the announcement) to $0.292 per share at a horizon of 5 days. This pattern of convergence toward the announced EPS number is consistent with forecasts incorporating some new information relevant to the prediction of EPS over the course of the year.

From Table 4 it appears that the most current analyst is no worse than the other sources, and that analysts dominate quarterly models in the longer horizons. An important caveat to this qualitative statement is that the relative performance results are highly correlated, both across horizons and across definitions of accuracy. The results of statistical tests for the differences in accuracy which are suggested by a perusal of Table 4 appear in Tables 5 and 6.

In Tables 5 and 6, a positive difference means that the

first of the pair is less accurate. Table 5 contains the results of pairwise comparisons among the three analyst consensus definitions. In terms of absolute error, which is reported in Panel A, when the differences are significant they favor the mean over the median and the most current forecaster over either the mean or the median. For example, at the 60-trading-day horizon, the t-statistic on the difference between the mean and median forecasts in average absolute accuracy is -2.14, which favors the mean, and is significant at the .05 level. The t-statistic on the difference between the median and the most current forecasts for the same horizon is 4.97, which favors the most current, and is significant at the .01 level.

In contrast, in Panel B, where differences in squared error accuracy are presented, there are no measurable differences in accuracy among the three analyst consensus definitions. The signs of differences in this panel generally are the same as those in Panel A, so the squared error evidence does not contradict the absolute error results. However, the very small t-statistics lend no additional support to the conclusions, either.

The results reported in Table 6 indicate that analyst forecasts generally dominate the time-series models at the longer horizons. For the 240, 180 and 120-trading-day horizons, wherever differences are statistically significant, the results favor analysts over the quarterly models. This evidence is consistent with the explanation that analysts use a broader information set than can be exploited by a univariate model.

At the 60-trading-day horizon, however, the quarterly time-series models dominate the mean and the median analyst forecast in all comparisons where significant differences exist. The most current forecast is never dominated to a statistically significant extent by the quarterly models, but generally is indistinguishable from them.

In summary, the accuracy results reported in Tables 4 through 6 generally support the conjectures that a current analyst forecast, presumably incorporating a broader information set, is at least as accurate as a forecast from a time-series model, and is at least as accurate as aggregations which ignore the forecast date. Although the statistical significance of results varies across forecasting horizons and accuracy criteria, wherever differences in accuracy between forecast sources are statistically significant, the results conform with expectations.

The results reported here probably understate the difference between the most current forecast and the mean and median definitions which appear in most other published work. My sample is selected to eliminate the publication lag, and so the mean and median forecasts in my sample are probably more accurate than, e.g., those in the I/B/E/S Summary Data.

The comparisons between analysts and quarterly models may understate differences because of the sample selection process. Since the sample of firms is weighted toward stabler and longer-lived firms by the data requirements of the time-series models, the selection process may exclude firms where analysts'

information advantage is largest: newer, or less stable firms, where time-series models that assume stationarity are less suitable. However, if firm size is a proxy for longevity or stability, this may be mitigated. Brown, Richardson and Schwager (1986) find that the superiority of ValueLine forecasts to a random walk model increases with firm size.

4.4   Accuracy Within a Subsample of Timely Forecasts

The results reported in section 4.3 suggest an advantage to timeliness in forecast selection. A single current forecast is no worse than, and sometimes dominates, aggregations which include both recently-updated and out-of-date forecasts. Given this, the question arises whether, conditional on timeliness, there is an advantage to diversifying idiosyncratic error. In this section, I address this question using a subsample which includes only forecasts made after the firm's most recent announcement of quarterly EPS.[14]

The selection of the subsample, described in Table 7, is as follows. For each horizon τ, firm j, and year t, I select the set of available forecasts which were made after the most recent previous quarterly earnings announcement. That is, all forecasts in this subsample have forecast dates indicating they have been revised since the last quarterly earnings release. This criterion reduces the number of forecasts available by about one-third at the 5-trading-day horizon, and by 70 to 80 percent or more at the other horizons.

The dramatic difference between the 5-trading-day and the other horizons in sample reduction, evident in Table 7, is largely due to the definition of forecast horizons. The 5-trading-day horizon date is, by construction, approximately three months after the third quarter announcement. All other horizon dates, also by construction, tend to be closer to the previous quarterly announcement.

Both the timeliness and the number of forecasts included in this subsample depend, among other things, on the number of days between the quarterly EPS announcement date and the horizon date. This fact is evident in Table 8, where summary information about the distribution of forecast ages in this subsample is reported. For horizons between 60 and 240 trading days, differences between the age of the most current forecast and the mean or the median of forecast ages are much less pronounced than those reported in Table 2 for the full sample.

Table 7 illustrates that up to 27% of the firm-years in the sample are eliminated entirely by this timeliness criterion. That is, for some firms in some years, none of the analysts' forecasts available on the fixed horizon date had been updated since the last quarterly EPS announcement. This is true even at the 5-trading-day horizon, where one percent of firm-years are eliminated. A related feature of the subsample is an increase in the number of firm-years with only forecast available. These observations, where by definition the mean, median and most current forecast are the same, reduce the power of statistical

tests to distinguish between the three consensus definitions.
However, in the full sample they are a trivial proportion of
observations.   In this subsample, they are 4 percent of
observations at the 5-trading-day horizon, but 20 to 30 percent
of observations at other horizons.

Tests of pairwise differences in accuracy among the three
consensus definitions in this subsample are reported in Table 9.
In contrast to the results in Table 5, the most current forecast
no longer dominates when all forecasts are reasonably current.
There are no statistically significant differences in accuracy in
the longer horizons, in part due to the sample selection issues
raised in the previous paragraph.   At the 5-trading-day horizon,
where sufficient forecasts are available to distinguish the
different consensus measures, the most current forecast is
significantly worse than either the mean or the median.   That is,
conditional on timeliness and on the availability of sufficient
numbers of forecasts, there are gains in accuracy from
aggregating to reduce idiosyncratic error.

4.5   Evaluating Forecasts - Market Association

The criteria developed in the previous sections, bias and
relative accuracy, are common in the literature on forecast
evaluation.   They do not, however, address the context in which
forecasts are used.   Both researchers' and investors' use of
forecast data in contexts related to securities markets suggests

that association with stock returns may provide a relevant empirical comparison.

If forecast errors reflect information relevant to the firm's prospects arriving after the forecast date, then, subject to two important qualifications discussed below, forecast errors will be positively correlated with new information impounded in stock returns over the forecast horizon. The first qualification to this implied association is that the information relevant to valuing the firm's common stock is not precisely the same as the information relevant to current-year earnings. There are errors in both variables with respect to the measured association between them. Non-recurring events, whether they are treated as extraordinary items or not, may affect earnings in a particular year, but may be inconsequential to the long-term value of the firm. Conversely, events that influence longer-term prospects, such as changes in investment opportunities, may affect the value of the firm without altering current earnings.

The second qualification is that excess returns are constructed to exclude one source of unanticipated information. The excess return is purged of its systematic relation with market returns, which includes both anticipated and unanticipated market returns. It is desirable to purge the stock returns of the anticipated component of the market return, since informationally inefficient forecasts will be correlated with anticipated information. Eliminating the unanticipated market return, however, may reduce the measurable association between

excess returns and forecast errors. On the other hand, it is important to note that while the power of tests for positive association is reduced, the reduction in power does not vary across forecast sources, since they are all evaluated relative to the same excess returns. In other words, the relative degree of association across sources will be unaffected.

Both qualifications noted above will have the effect of reducing the measurable association between forecast errors and excess returns. Nevertheless, previous studies facing the same inherent difficulties have found statistically significant positive associations between unexpected earnings and excess stock returns (Ball and Brown (1968), Beaver, Clarke and Wright (1979), and Fried and Givoly (1982), for example).

The regression model used to estimate the association between cumulated excess returns, represented by $U_{jt\tau}$, and forecast errors, $e_{ijt\tau}$, is:

$$e_{ijt\tau} = \alpha_{1ijt} + \alpha_{2it\tau} + \beta_{i\tau} U_{jt\tau} + v_{ijt\tau} \qquad (12)$$

In (12), $\beta_{i\tau}U_{jt\tau}$ is the portion of the forecast error from source i at horizon τ which is systematically related to excess returns. The slope coefficient $\beta_{i\tau}$ is the covariance between excess returns and forecast errors, adjusted for firm and year effects, in units of the variance of excess returns. Using excess returns as the independent variable and forecast errors as dependent has the desirable feature that $\beta_{i\tau}$ and its associated t-statistic have the same scale for all sources i. If the roles of these two variables were reversed in the regression equation,

the estimated regression slope coefficient would depend explicitly on the forecast error variance from source i.

The constants $\alpha_{1ij\tau}$ and $\alpha_{2it\tau}$ measure, respectively, firm- and year-specific average forecast errors, conditional on the systematic relation with excess returns. The $\alpha_{2it\tau}$, the year effects, play an important role in equation (12), since they capture the time-period-specific information in forecast errors which is not captured by excess returns.[15] Among other things, they include the average effect of omitting the unanticipated component of the market return. If the $\alpha_{2it\tau}$ are not included in the model, they are impounded in the regression residuals as an omitted variable. This induces cross-sectional correlation in the residuals, which if ignored leads to incorrect statistical inferences, as was discussed above for the bias computation.[16]

Equation (12) is estimated by stacking the regressions for the five forecast sources, and estimating them jointly. The forecast sources are the mean, median and most current analyst forecast and the two quarterly models. Estimations are performed jointly for the five forecast sources, and separately for each forecast horizon.

Since the matrix of independent variables, which consists of cumulated excess returns over the forecast horizon and dummy variables indicating firms and years, is the same for each of the forecast sources, there is no efficiency gain over equation-by-equation least squares (see Zellner (1962)). The advantage of stacking the equations is for joint estimation of the firm and

year effects, so that observations from each of the five forecast sources are adjusted for the same firm and year effects.

Tables 10 through 12 contain the regression results from estimation of model (12). In Table 10, I report the estimated slope coefficients and their associated t-statistics, testing the statistical significance of the relation between forecast errors and excess returns over the forecasting horizon. Table 11 contains t-statistics which test for differences in the slope coefficients across forecast sources. Table 12 contains regression summary statistics, including adjusted R^2, F-statistics, and numbers of parameters, and sample sizes.

According to results reported in Table 12, equation (12) explains between 9% and 16% of the variation in forecast errors, with slight variations across horizons. The largest adjusted R^2 appears at the 5-trading-day horizon, though differences in explanatory power are not large. The model has statistically significant explanatory power, according to the regression F-statistics, which reject the null of no explanatory power at the .001 level.

The incremental F-statistics in Table 12 confirm that the year-specific effects are important in equation (12).[17] The F-statistic on the year-specific effects tests the null hypothesis:

$$H_0 : \alpha_{2i1\tau} = \alpha_{2i2\tau} = \cdots = \alpha_{2iT\tau} = 0$$

That is, the F-statistic tests the null hypothesis that estimation of year-specific intercepts adds no explanatory power to the model. This hypothesis is rejected at the .05 level at

all horizons, and at the .001 level or better at the horizons longer than 5 days. The importance of year effects in the model increases with the length of the horizon. This is consistent with the information-based explanation for their inclusion in the model, namely that forecast errors impound time-period-specific unanticipated information. Over longer horizons, loosely speaking, the "quantity" of unanticipated information is greater. The strength of this result also confirms the assertions made earlier that a cross-section of forecast errors for a single time period is not a set of independent observations.

The F-statistics on firm-specific effects reported in Table 12 also reject the null hypothesis, which is:

$$H_0 : \alpha_{1i1\tau} = \alpha_{1i2\tau} = \cdots = \alpha_{1ij\tau} = 0$$

There are measurable firm-specific differences in average forecast error at all horizons, even after the adjustment for firm-specific information impounded in excess returns. The strength of this result varies little across forecast horizons and across transformations of the dependent variable.

The importance of the slope coefficients in model (12), indicated by the F-statisic reported in Table 12, varies across forecast horizons and across transformations of the dependent variable. The individual slope coefficients reported in Table 10, however, are of greater relevance. Generally, the results in Table 10 show a pattern of positive association between forecast errors and excess returns. A positive association is expected if, first, there is some overlap between information relevant to

firm value and information relevant to current-year earnings, and second, some of this overlapping information is unanticipated both by investors and by the predictor of EPS.

The statistical significance of the positive association varies somewhat across forecast horizons, and more importantly across forecast sources. The strongest results are for the 120-day horizon. Among the analyst consensus forecasts, the strongest results are generally for the most current forecaster, which is consistent with the most current forecaster acting as a reasonable composite, or expectation. The strongest results, and the only ones which are consistently positive and statistically significant, however, are for the quarterly autoregressive model, equation (1). This pattern of relative performance does not vary substantially across horizons or across transformations of the dependent variable. This result is anomalous, especially in light of the quarterly model's relative inaccuracy. It indicates that prior knowledge of the forecast error from a quarterly autoregressive model is a better predictor of excess returns than prior knowledge of the forecast error from analysts' forecasts.[18]

The greater association of excess returns with forecast errors from a time-series model than with those from analysts is not consistent with the results of Fried and Givoly (1982). I use quarterly data in the time series models of annual earnings, while Fried and Givoly (1982) use annual data. Since models using quarterly data produce more accurate forecasts of annual earnings than models using only annual data (see Hopwood, McKeown

and Newbold (1982)), presumably my tests are more demanding of
analysts than those used by Fried and Givoly. However, the
result remains anomalous since analysts are more accurate and can
employ more information than quarterly time series models.

This anomaly is further investigated in Table 11, by testing
for the statistical significance of differences in the slope
coefficients. A further advantage of stacking equations across
forecast sources to estimate them is that statistical testing is
simplified, since a set of linear constraints on the estimated
slope coefficients $\beta_{i\tau}$ generates direct tests of differences in
slope across forecast sources. For example, if $\underline{\beta}_\tau$ is the vector
of five slope coefficients, one for each forecast source, and if
$\underline{c}_{12}'$ is the vector $(1,-1,0,0,0)$, then $\underline{c}_{12}'\underline{\beta}_\tau$ estimates the
difference in slopes between the first and second sources, with
estimated standard error:

$$s_{\underline{c}'\underline{\beta}} = s_v \cdot [\underline{c}'(X'X)_\beta^{-1}\underline{c}]^{1/2} \qquad (13)$$

In (13), s_v is the residual standard error from the joint
estimation of equation (12). $(X'X)^{-1}_\beta$ is the lower-right
submatrix of five rows and five columns from the $(X'X)^{-1}$ matrix
of equation (12). This submatrix determines the variance-
covariance relations among the five slope coefficients.

Linear constraints of the form of $\underline{c}_{12}$ are used to evaluate
differences in the slope coefficients that appear in Table 10,
i.e. differences across forecast sources in the association
between excess returns and forecast errors. Table 11 contains

the results of these statistical tests. Results are shown for
tests of pairwise differences between the quarterly
autoregressive model and all other forecast sources, and for
differences between the most current analyst and the mean and
median analyst forecasts. The statistical tests confirm the
anomalous result, that errors from a mechanical quarterly model
often are significantly more closely related to excess returns
than errors from analysts. In addition, the tests indicate that
while the most current forecast typically shows the strongest
result among analyst consensus forecasts, the difference is not
statistically significant, in general.

5.   Summary and Conclusions

Analysts' predictions of EPS are a potential source of
"market expectations" information. I have examined properties of
different composite forecasts, on arbitrarily-chosen dates which
span approximately a year. Results reported here indicate that
the most current forecast available from an analyst dominates the
mean or median of the available forecasts.

Five alternate sources of earnings expectations are
examined: the mean, median, and most current of available
analysts' forecasts; an autoregressive model in fourth
differences of the univariate series of quarterly EPS, and a
fourth-differenced random walk using quarterly EPS. The two

quarterly time-series models are included primarily as benchmarks.

The most current forecast is at least as accurate as either the mean or median forecast, and generally dominates them in absolute error terms, when all available forecasts are considered. This result indicates that the date of the forecast is relevant for determining its accuracy, and dominates the diversification obtained by aggregating forecasts from different sources as a "first cut" criterion. Since most published aggregations of forecasts and much previous research treat all forecasts for a given firm as if they are equally current, they ignore this relevant piece of information.

When forecasts are censored in a way that eliminates the most out-of-date forecasts from the sample, it is possible to improve accuracy by aggregating forecasts. When only forecasts from the fourth fiscal quarter are included, the mean or median forecast at a horizon of 5 trading days before the annual EPS announcement is more accurate than the single most current forecast.

In this sample the forecast error from the most current forecast is more closely associated with excess returns over the forecast horizon than the error from the mean or the median forecast, but the difference in association is not, in general, statistically significant.

Analysts generally are significantly more accurate than time-series models. Errors from the quarterly autoregressive

model, however, appear to be more closely related with excess returns over the forecasting horizon than those of analysts. Because of this anomalous result, it is unclear that analysts provide a better model of the "market expectation" than mechanical models.

It should be noted, though, that the sample of firms was reduced sharply by the data requirements of the time-series models. This sample, with its selection bias toward longer-lived firms with continuous data available, does not clearly isolate cases where analysts might be expected to have the most advantage over mechanical models, and perhaps eliminated many of these cases. These firms, where there is a substantial amount of non-earnings information expected to have an impact on earnings, may be a fruitful area for future investigation.

ACKNOWLEDGEMENT

FOOTNOTES

[1]See Bates and Granger (1969), Granger and Newbold (1977) and Figlewski and Urich (1982).

[2]I use words like "aggregate", "composite" and "consensus" in the general sense, to include weighting schemes which put all weight on a single forecast, and none on others.

[3]Dimson and Marsh (1984) find similar results for forecasts of stock returns. Their data are from a designed experiment in which forecasts were collected at regular intervals, so they were assured a sample of nearly simultaneous forecasts. In analysts' EPS forecast data such as those used here, the availability of recent forecasts depends, in part, on analysts' private decisions about when to update. This point is discussed further in section 4.4.

[4]"Earnings" and "earnings per share" (or "EPS") are used interchangeably in this paper. The data are forecasts of EPS.

[5]Since firms' annual earnings announcement dates are remarkably consistent year after year, choosing fixed lengths of time prior to the announcement date is a fairly accurate means of finding dates that differ by one quarter. For the 6218 horizons in this paper, 13 horizon dates did not fall between the quarterly announcements as intended. The results are not affected by deletion of these observations.

[6]Results reported in this paper are for unscaled forecast errors as defined in equation (3). Other forecast error metrics, or scales, may be appropriate, for example to control for heteroscedasticity. However, the qualitative conclusions of the paper, other than the bias results, are not affected by the choice of scale. The scales I have investigated are: (1) standardized forecast errors, where the denominator is (a) the average, over the previous five years, of absolute changes in EPS, or (b) the standard deviation of EPS changes; and (2) percent forecast errors, where the denominator is (a) the absolute value of the prior year's EPS, or (b) where the sample is censored to exclude negative denominators, or (c) to exclude denominators less than $0.20. These results are available upon request from the author.

[7]The numbers of observations deleted are: 1, 2, 2, 5, and 5, at the 5-, 60-, 120-, 180-, and 240-trading-day horizons, respectively.

[8]My results do not differ if the value-weighted market portfolio of securities is used as a proxy for the market.

[9]The subscript i, which indexes the source of the forecast, is

suppressed in the following discussion for readability.

[10]The estimation was also done with firm-specific effects in the model. The bias results do not differ qualitatively from those reported here.

[11]Fried and Givoly (1982) report negative statistically significant bias in analysts' April forecasts for December yearends for the period 1969 through 1979, but none in time-series forecasts. I replicate the bias test for percent forecast errors (see footnote 6), which is approximately their definition. I find no bias in 120-day through 240-day analyst forecasts, and significant positive bias in the time-series model forecasts. Thus, I conclude that their bias result is not replicated in this later (1975-1981) sample.

[12]For example, see Dirks and Gross (1974), especially pp. 252-257. Also see "Bank Analysts Try to Balance their Ratings", WSJ, May 29, 1984, p. 33; and "Picking a Loser", WSJ, September 28, 1983, p. 1.

[13]The estimated effects are not the same as the estimated coefficients, because model (9) has two sets of effects, firms and years. A simple example will illustrate this. If there were two years and three firms in the sample, model (9) could be estimated with no intercept, using two year dummy variables (DY1 and DY2) and two firm dummy variables (DF1 and DF2):

$$|e_{jt\tau}| = d_{11\tau}DF1 + d_{12\tau}DF2 + d_{21\tau}DY1 + d_{22\tau}DY2 + \xi_{jt\tau}.$$

The "year 1 effect" is the average $|e_{jt\tau}|$ for t=1. This effect is not estimated by $d_{21\tau}$. Rather, $d_{21\tau}$ is the average $|e_{jt\tau}|$ for year 1 for the omitted (third) firm. The "year 1 effect" is estimated in this formulation by:

$$\delta_{21\tau} = d_{21\tau} + (1/3)\, d_{11\tau} + (1/3)\, d_{12\tau} .$$

The "firm 1 effect" is estimated by:

$$\delta_{11\tau} = d_{11\tau} + (1/2)\, d_{21\tau} + (1/2)\, d_{22\tau} .$$

The "firm 3 effect" is estimated by:

$$\delta_{13\tau} = (1/2)\, d_{21\tau} + (1/2)\, d_{22\tau} .$$

Since any non-redundant spanning set of dummy variables can be used, the particular linear combinations of coefficients to estimate the firm and year effects depend on the model used.

[14]This analysis was suggested to me by the referee.

[15]Dropping the $\alpha_{1ij\tau}$, the firm-specific effects, does not alter the estimates of the slope coefficients $\beta_{i\tau}$ or their statistical significance by a substantial amount. Omitting the $\alpha_{2it\tau}$, however, alters both the estimates and their statistical significance.

[16]An alternate way to model the problem is to include the year or firm effects as "random effects", contributing off-diagonal elements to the covariance matrix of the $v_{ijt\tau}$. Mundlak (1978) shows that the estimate of the slope coefficient $\beta_{i\tau}$ obtained from a model like (12), is identical to the GLS estimate which would be obtained if the firm- and year-specific effects, $\alpha_{1ij\tau}$ and $\alpha_{2it\tau}$, were modeled as random effects and included in the covariance matrix.

[17]The statistical significance of year-specific effects as a determinant of forecast errors, although not reported in the previous tables, is similar in the estimates of bias and accuracy.

[18]An alternate method of measuring the association between forecast errors and excess returns, similar to that of Ball and Brown (1968), is to construct portfolios based on foreknowledge of the sign of EPS forecast error. That is, a long position is taken in each of the securities for which the forecast error is positive, and a short position is taken in those with negative errors. This procedure, applied to these data, produces results qualitatively identical to those reported here.

REFERENCES

Ball, R. and P. Brown (1968). An Empirical Evaluation of Accounting Income Numbers. Journal of Accounting Research 6: 159-178.

Barefield, R. and E. Comiskey (1975). The Accuracy of Analysts' Forecasts of Earnings Per Share. Journal of Business Research 3: 241-252.

Bates, J. and C.W.J. Granger (1969). The Combination of Forecasts. Operations Research Quarterly 20: 451-468.

Beaver, W.H., R. Clarke and W. Wright (1979). The Association Between Unsystematic Security Returns and the Magnitude of Earnings Forecast Errors. Journal of Accounting Research 17:316-340.

Brown, L., P. Griffin, R. Hagerman, and M. Zmijewski (1986). Security Analyst Superiority Relative to Univariate Time-Series Models in Forecasting Quarterly Earnings. working paper.

Brown, L., G. Richardson, and S. Schwager (1986). Financial Analyst Forecast Superiority: Is There an Information Interpretation?. working paper.

Brown, L. and M. Rozeff (1978). The Superiority of Analysts' Forecasts as Measures of Expectations: Evidence from Earnings. Journal of Finance 33: 1-6.

Brown, P., G. Foster and E. Noreen (1984). Security Analyst Multi-Year Earnings Forecasts and the Capital Market. American Accounting Association Studies in Accounting Research #21.

Collins, W.A. and W.S. 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 18: 390-406.

Crichfield, T., T. Dyckman, and J. Lakonishok (1978). An Evaluation of Security Analysts' Forecasts. Accounting Review 53: 651-668.

Darlin, D. (1983). Picking a Loser. The Wall Street Journal. September 28. p. 1.

Dimson, E., and P. Marsh (1984). An Analysis of Brokers' and Analysts' Unpublished Forecasts of UK Stock Returns. Journal of Finance 39: 1257-1292.

Dirks, R. and L. Gross (1974). The Great Wall Street Scandal.
    McGraw-Hill Book Company.

Elton, E., M. Gruber, and M. Gultekin (1981). Professionnal
    Expectations: Accuracy and Diagnosis of Errors. working
    paper, New York University.

Figlewski, S. and T. Urich (1982). Optimal Aggregation of Money
    Supply Forecasts: Accuracy, Profitability, and Market
    Efficiency. working paper, New York University.

Foster, G. (1977). Quarterly Earnings Data: Time Series
    Properties and Predictive Ability Results. Accounting
    Review 52: 1-21.

Fried. D. and D. Givoly (1982). Financial Analysts' Forecasts of
    Earnings: A Better Surrogate for Market Expectations.
    Journal of Accounting and Economics 4: 85-107.

Gibson, R. and W.L. Wall (1984). Bank Analysts Try to Balance
    Their Ratings. The Wall Street Journal. May 29. p. 33.

Givoly, D. and J. Lakonishok (1979). The Information Content of
    Financial Analysts' Forecasts of Earnings. Journal of
    Accounting and Economics 1: 165-185.

Granger, C.W.J., and P. Newbold (1977). Forecasting Economic
    Time Series. Academic Press.

Hopwood, W.S., J.C. McKeown, and P. Newbold (1982). The
    Additional Information Content of Quarterly Earnings
    Reports: Intertemporal Disaggregation. Journal of
    Accounting Research 20: 343-349.

Malkiel, B. and J. Cragg (1982). Expectations and the Structure
    of Share Prices. University of Chicago Press.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross
    Section Data. Econometrica 46: 69-85.

Watts, R., and R. Leftwich (1977). The Time Series of Annual
    Earnings. Journal of Accounting Research 15: 253-271.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly
    Unrelated Regressions and Tests for Aggregation Bias.
    Journal of the American Statistical Association 57: 348-
    368.

Table 1

Sample Selection Criteria and Their Effects on the Sample

Panel A: Elimination of Firms and Firm-Years

Criteria	# firms	# firm-years
1. December yearend, and $\geq$ 1 forecast in each year, '75-81	508	3556
2. Annual EPS on COMPUSTAT, and quarterly announcement dates on COMPUSTAT or in WSJ	497	3440
3. Return data on CRSP	410	not tabulated
4. Quarterly EPS on COMPUSTAT for all of 69-III through 75-IV	184	1260

Panel B: Number of Analyst Forecasts Available, by Forecast Horizon

Criteria (from Panel A)	# firms	Forecast Horizon, in Trading Days				
		240	180	120	60	5
2.	497	31425	37450	40805	43155	45016
3.	410	26308	31193	33676	35644	37135
4.	184	16134	19294	20969	22078	22864

Notes:

Forecast horizons are measured in trading days prior to the annual
earnings announcement.

Table 2

Selected Characteristics of the Distribution of Forecast Ages,
for the Mean, Median and Most Current Analyst Forecasts,
Measured Across Firms and Years

[forecast ages are measured in trading days]

| | Horizon | Fractiles | | | | | Sample Moments | | | |
		.1	.25	.5	.75	.9	mean	stdev	skew	N
mean	240	34	46	60	74	89	61.5	24.7	0.8	1198
	180	33	43	58	73	91	59.9	23.9	0.8	1235
	120	34	47	62	82	102	65.9	28.1	1.0	1254
	60	38	51	67	88	110	71.7	31.8	1.6	1258
	5	40	52	67	87	109	72.3	31.2	2.5	1260
median	240	7	20	51	83	127	59.4	47.9	1.1	1198
	180	5	16	41	78	131	56.6	54.5	1.5	1234
	120	7	19	47	80	135	61.1	58.0	1.8	1254
	60	7	19	49	88	155	66.6	67.7	2.0	1258
	5	15	26	43	71	131	61.7	60.8	2.6	1260
current	240	1	2	4	14	43	14.4	24.7	3.3	1198
	180	1	1	3	9	24	9.0	16.9	4.8	1233
	120	1	2	4	11	25	10.0	18.8	6.3	1254
	60	1	1	4	10	26	10.6	22.5	6.8	1259
	5	1	5	11	20	34	15.8	22.0	8.9	1259

Notes:

A forecast's age is the number of trading days between the analyst's
forecast date and the horizon date selected in this study. Forecast
horizons are measured in trading days prior to the annual earnings
announcement.

The distributions described above are over firms and years for ages
defined as follows:
                mean - the mean of the ages of available analyst forecasts,
                        for each firm and year.
              median - the median of the ages of available analyst
                        forecasts, for each firm and year.
            current - the age of the most recent analyst's forecast, for
                        each firm and year.
Note that the median age is not necessarily the age of the forecast
which is median, nor is the mean age necessarily the age of a forecast
equal to the mean forecast.

Table 3

Forecast Bias, for Five Forecast Sources, at Five Forecast Horizons

[bias numbers are denominated in $ per share]

Source	Forecast Horizon, in Trading Days				
	240	180	120	60	5
q.a.r.	-0.15	-0.11	-0.14	-0.08	
	(-3.82)	(-3.15)	(-5.18)	(-4.11)	
q.r.w.	-0.09	-0.05	-0.11	-0.06	
	(-2.17)	(-1.49)	(-4.24)	(-3.25)	
mean	-0.08	-0.12	-0.12	-0.09	-0.05
	(-2.09)	(-3.57)	(-4.70)	(-4.96)	(-2.74)
median	-0.03	-0.05	-0.06	-0.02	0.02
	(-0.72)	(-1.52)	(-2.25)	(-1.19)	(1.00)
current	-0.08	-0.10	-0.09	-0.04	-0.02
	(-1.91)	(-2.93)	(-3.27)	(-2.34)	(-1.42)

Notes:

The computation of forecast bias and its associated t-statistic (in parentheses) are described in equations (6) through (8) in the text.

The forecast sources are:
    q.a.r. - a quarterly autoregressive model in fourth differences; equation (1) in the text.
    q.r.w. - a random walk model with drift in fourth differences; equation (2) in the text.
    mean - the mean of the available analysts' forecasts.
    median - the median of the available analysts' forecasts.
    current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual earnings announcement.

The degrees of freedom for all reported t-statistics are over 1,000, so they are approximately normal. For a two-sided test, the .05 and .01 critical points of the N(0,1) distribution are 1.96 and 2.58, respectively.

Table 4

Forecast Accuracy: Average Absolute or Squared Forecast Error
for Five Forecast Sources, and Five Forecast Horizons

[forecast errors are denominated in $ per share]

		Forecast Horizon, in Trading Days				
	Source	240	180	120	60	5
	q.a.r.	0.975	0.780	0.592	0.350	
Average	q.r.w.	0.963	0.781	0.620	0.363	
Absolute	mean	0.747	0.645	0.516	0.395	0.291
Forecast	median	0.788	0.677	0.546	0.435	0.321
Error	current	0.742	0.610	0.468	0.342	0.292
	q.a.r.	2.771	1.877	1.203	0.460	
Average	q.r.w.	2.626	1.897	1.255	0.447	
Squared	mean	1.531	1.239	0.818	0.478	0.351
Forecast	median	1.595	1.246	0.790	0.546	0.409
Error	current	1.638	1.178	0.669	0.389	0.361

Notes:

The computation of average absolute error is described in equations
(9) and (10) in the text. The computation of average squared error is
analogous to the computation of average absolute error.

The forecast sources are:
     q.a.r. - a quarterly autoregressive model in fourth differences;
              equation (1) in the text.
     q.r.w. - a random walk model with drift in fourth differences;
              equation (2) in the text.
       mean - the mean of the available analysts' forecasts.
     median - the median of the available analysts' forecasts.
    current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

Table 5

Pairwise Differences in Forecast Accuracy
Among the Mean, Median and Most Current Analyst Forecasts,
for Five Forecast Horizons

Panel A: t-Statistics on Differences in Average Absolute Error

| | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
mean - median	-1.06	-0.98	-1.11	-2.14	-1.62
mean - current	0.11	1.07	1.82	2.83	-0.02
median - current	1.18	2.05	2.93	4.97	1.60

Panel B: t-Statistics on Differences in Average Squared Error

| | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
mean - median	-0.26	-0.04	0.20	-0.83	-0.59
mean - current	-0.44	0.30	1.06	1.09	-0.10
median - current	-0.18	0.34	0.86	1.91	0.50

Notes:

The reported numbers are t-statistics on pairwise differences in
average absolute or squared forecast error. See Table 3 for the
average absolute and squared errors. The computations are described
in equations (9) through (11) in the text.

The degrees of freedom for all reported t-statistics are over 2,000,
so they are approximately normal. For a two-sided test, the .05 and
.01 critical points from the $N(0,1)$ distribution are 1.96 and 2.58,
respectively.

The forecast sources are:
        mean - the mean of the available analysts' forecasts.
        median - the median of the available analysts' forecasts.
        current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

Table 6

Pairwise Differences in Forecast Accuracy Between Analysts and
Quarterly Time-Series Models, for Five Forecast Horizons

Panel A: t-Statistics on Differences in Average Absolute Error

| | Quarter: 1 | 2 | 3 | 4 |
	Horizon: 240	180	120	60
q.a.r. - mean	5.85	4.15	2.85	-2.36
q.a.r. - median	4.79	3.17	1.74	-4.51
q.a.r. - current	5.96	5.22	4.67	0.46
q.r.w. - mean	5.56	4.20	3.92	-1.70
q.r.w. - median	4.50	3.21	2.81	-3.84
q.r.w. - current	5.67	5.27	5.74	1.13

Panel B: t-Statistics on Differences in Average Squared Error

| | Quarter: 1 | 2 | 3 | 4 |
	Horizon: 240	180	120	60
q.a.r. - mean	5.08	3.16	2.73	-0.22
q.a.r. - median	4.81	3.13	2.93	-1.04
q.a.r. - current	4.64	3.47	3.79	0.87
q.r.w. - mean	4.49	3.26	3.09	-0.38
q.r.w. - median	4.22	3.23	3.29	-1.21
q.r.w. - current	4.05	3.56	4.15	0.70

Notes:

The reported numbers are t-statistics on pairwise differences in
average absolute or squared forecast error. See Table 3 for the
average absolute and squared errors.

The forecast sources are:
    q.a.r. - a quarterly autoregressive model in fourth differences;
           equation (1) in the text.
    q.r.w. - a quarterly random walk model in fourth differences;
           equation (2) in the text.
      mean - the mean of the available analysts' forecasts.
    median - the median of the available analysts' forecasts.
   current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

The degrees of freedom for all reported t-statistics are over 2,000.
For a two-sided test, the .05 and .01 critical points from the N(0,1)
distribution are 1.96 and 2.58, respectively.

Table 7

Selection of the Subsample, for Each Forecast Horizon,
of Forecasts Made Between the Last Announcement
of Quarterly Earnings and the Horizon Date

	Forecast Horizon, in Trading Days				
Full Sample:	240	180	120	60	5
# forecasts	16134	19294	20969	22078	22864
# firms	184	184	184	184	184
# firm-years	1198	1235	1254	1259	1260
# firm-years with 1 forecast	68	32	31	22	25
Subsample:					
# forecasts	3017	5601	4425	3859	15352
# firms	181	184	183	183	184
# firm-years	875	1104	1039	961	1246
# firm-years with 1 forecast	250	209	246	259	53

Notes:

The full sample, whose selection is described in Table 1, includes,
for each firm and year, all available forecasts that have not been
revised or withdrawn prior to the horizon date.

The subsample includes, for each firm and year, only those forecasts
which have been revised or initiated between the firm's last
announcement of quarterly EPS and the horizon date.

Table 8

Selected Characteristics of the Distribution of Forecast Ages,
Measured Across Firms and Years, for the Subsample of Forecasts
Made Between the Last Announcement of Quarterly Earnings
and the Horizon Date

| | Horizon | Fractiles | | | | | Sample Moments | | | |
		.1	.25	.5	.75	.9	mean	stddv	skew	N
mean	240	2	4	6	8	11	6.1	4.6	7.86	875
	180	3	6	9	12	17	9.4	5.5	1.12	1103
	120	2	4	7	11	15	8.1	5.2	1.07	1039
	60	2	3	6	9	12	6.5	4.4	1.06	961
	5	25	29	34	41	47	35.2	9.4	.53	1246
median	240	1	3	6	8	11	6.2	4.9	6.61	875
	180	3	5	9	13	17	9.4	6.0	1.01	1104
	120	2	4	7	11	16	8.1	5.6	1.04	1039
	60	1	3	6	9	12	6.4	4.6	1.10	961
	5	22	29	35	43	49	35.8	11.1	.28	1246
current	240	0	1	3	5	9	3.8	3.7	1.49	875
	180	0	1	3	6	12	4.6	5.6	2.10	1102
	120	0	1	3	6	11	4.6	5.1	2.02	1039
	60	0	1	3	5	9	3.7	4.1	1.96	961
	5	1	5	11	20	33	14.3	12.8	1.50	1245

Notes:

A forecast's age is the number of trading days between the analyst's
forecast date and the horizon date selected in this study. Forecast
horizons are measured in trading days prior to the annual earnings
announcement.

The distributions described above are over firms and years for ages
defined as follows:

        mean – the mean of the ages of available analyst forecasts,
             for each firm and year.

      median – the median of the ages of available analyst
             forecasts, for each firm and year.

    current – the age of the most recent analyst's forecast, for
             each firm and year.

Note that the median age is not necessarily the age of the forecast
which is median, nor is the mean age necessarily the age of a forecast
equal to the mean forecast.

The subsample includes, for each firm and year, only those forecasts
which have been revised or initiated between the firm's last
announcement of quarterly EPS and the horizon date.

Table 9

Pairwise Differences in Forecast Accuracy Among the Mean,
Median and Most Current Analyst Forecasts, in the Subsample of
Forecasts Made Between the Last Announcement of Quarterly Earnings
and the Horizon Date, for Five Forecast Horizons

Panel A: t-Statistics on Differences in Average Absolute Error

| | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
mean - median	0.034	0.048	0.059	-0.067	0.388
mean - current	-0.240	-0.210	-1.169	-1.106	-3.508
median - current	-0.275	-0.258	-1.228	-1.039	-3.896

Panel B: t-Statistics on Differences in Average Squared Error

| | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
mean - median	-0.078	-0.318	-0.101	-0.099	0.168
mean - current	-0.489	-0.008	-0.704	-0.370	-2.146
median - current	-0.411	0.310	-0.603	-0.270	-2.314

Notes:

The subsample includes, for each firm and year, only those forecasts
which have been revised or initiated between the firm's last
announcement of quarterly EPS and the horizon date.

The computation of average absolute error is described in equations
(9) and (10) in the text.   The computation of average squared error is
analogous to the computation of average absolute error.

The forecast sources are:
    q.a.r. - a quarterly autoregressive model in fourth differences;
              equation (1) in the text.
    q.r.w. - a random walk model with drift in fourth differences;
              equation (2) in the text.
      mean - the mean of the available analysts' forecasts.
    median - the median of the available analysts' forecasts.
   current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

Table 10

Slope Coefficients from the Regression of EPS Forecast Error
on Excess Return, for Five Forecast Horizons:

$$e_{ijt\tau} = \alpha_{1ij\tau} + \alpha_{2it\tau} + \beta_{i\tau} U_{jt\tau} + v_{ijt\tau} \qquad (12)$$

| Source | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
q.a.r.	0.98	0.96	0.99	0.48	
	(5.38)	(5.72)	(6.46)	(3.26)	
q.r.w.	0.29	0.24	0.63	0.18	
	(1.61)	(1.44)	(4.15)	(1.25)	
mean	0.14	0.01	0.34	0.14	0.66
	(0.78)	(0.04)	(2.20)	(0.96)	(1.68)
median	0.07	0.07	0.41	0.20	0.77
	(0.40)	(0.40)	(2.68)	(1.38)	(1.97)
current	0.32	0.29	0.53	0.37	0.25
	(1.75)	(1.75)	(3.49)	(2.49)	(0.64)

Notes:

The forecast sources are:
    q.a.r. - a quarterly autoregressive model in fourth differences;
        equation (1) in the text.
    q.r.w. - a random walk model with drift in fourth differences;
        equation (2) in the text.
     mean - the mean of the available analysts' forecasts.
   median - the median of the available analysts' forecasts.
  current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

In equation (12), $e_{ijt\tau}$ is the forecast error, in \$ per share, from
source i for firm j in year t at horizon τ, and $U_{jt\tau}$ is cumulated
excess returns for firm j in year t over horizon τ. Equation (12) is
estimated jointly for five forecast sources and separately for each
horizon. The $\alpha_{1ij\tau}$ are firm-effects, and the $\alpha_{2it\tau}$ are year-effects.

A slope $\beta_{i\tau}$ is estimated for each forecast source i and horizon τ.
The t-statistic on each slope coefficient is reported in parentheses.
The t-statistics have degrees of freedom greater than 1,000. For a
one-sided test, the .05 and .01 critical points from the N(0,1)
distribution are 1.65 and 2.33, respectively.

Table 11

Pairwise Differences in Slope Coefficients from the Regression
of EPS Forecast Error on Excess Return, for Five Forecast Horizons:

$$e_{ijt\tau} = \alpha_{1ij\tau} + \alpha_{2it\tau} + \beta_{i\tau} U_{jt\tau} + v_{ijt\tau} \tag{12}$$

Panel A: t-Statistics on Differences between Quarterly Auto-
regressive Model [eqn. (1)] and Other Forecast Sources

	Quarter:	1	2	3	4	
	Horizon:	240	180	120	60	5
q.a.r. - q.r.w.		2.69	3.08	1.68	1.47	
q.a.r. - mean		3.28	4.08	3.09	1.68	
q.a.r. - median		3.55	3.82	2.74	1.37	
q.a.r. - current		2.59	2.85	2.16	0.56	

Panel B: t-Statistics on Differences between the Most Current
and the Mean or Median Analyst Forecasts

	Horizon:	240	180	120	60	5
mean - current		-0.69	-1.23	-0.94	-1.12	0.76
median - current		-0.96	-0.97	-0.59	-0.82	0.97

Notes:

The forecast sources are:
        q.a.r. - a quarterly autoregressive model in fourth
                differences; equation (1) in the text.
        q.r.w. - a random walk model with drift in fourth
                differences; equation (2) in the text.
          mean - the mean of the available analysts' forecasts.
        median - the median of the available analysts' forecasts.
       current - the most recent forecast from an analyst.

The forecast horizons are measured in trading days prior to the annual
earnings announcement.

In equation (12), $e_{ijt\tau}$ is the forecast error, in \$ per share, from
source i for firm j in year t at horizon τ, and $U_{jt\tau}$ is cumulated
excess returns for firm j in year t over horizon τ. Equation (12) is
estimated jointly for five forecast sources and separately for each
horizon. The $\alpha_{1ij\tau}$ are firm-effects, and the $\alpha_{2it\tau}$ are year-effects.

A slope $\beta_{i\tau}$ is estimated for each forecast source i and horizon τ.
The t-statistic on each slope coefficient is reported in parentheses.
The t-statistics have degrees of freedom greater than 1,000. For a
one-sided test, the .05 and .01 critical points from the N(0,1)
distribution are 1.65 and 2.33, respectively.

Table 12

Regression Summary Statistics for the Regression
of EPS Forecast Error on Excess Return, for Five Forecast Horizons:

$$e_{ijt\tau} = \alpha_{1ij\tau} + \alpha_{2it\tau} + \beta_{i\tau} U_{jt\tau} + v_{ijt\tau} \qquad (12)$$

| | Forecast Horizon, in Trading Days | | | | |
	240	180	120	60	5
adjusted R^2	.116	.095	.107	.114	.156
full model $F(k-1,N-k)$	5.02	4.27	4.79	5.09	4.60
year-effect $F(k1,N-k)$	54.13	36.79	25.17	19.25	2.94
firm-effect $F(k2,N-k)$	3.34	3.08	3.87	4.82	4.65
excess returns $F(k3,N-k)$	6.89	7.33	14.37	3.57	2.18
d.f. for F-statistics:					
sample size (N)	5986	6171	6267	6293	3779
# of parameters (k)	199	199	199	199	195
# of years (k1+1)	7	7	7	7	7
# of firms (k2+1)	184	184	184	184	184
# of slopes (k3)	5	5	5	5	3

Notes:

In equation (12), $e_{ijt\tau}$ is the forecast error, in $ per share, from
source i for firm j in year t at horizon τ, and $U_{jt\tau}$ is cumulated
excess returns for firm j in year t over horizon τ. Equation (12) is
estimated jointly for five forecast sources and separately for each
horizon. The $\alpha_{1ij\tau}$ estimate firm-specific effects, and the $\alpha_{2it\tau}$
estimate year-specific effects.

The full model F-statistic tests the null hypothesis that the
regression model (12) has explanatory power. The year, firm, and
excess return F-statistics test the incremental explanatory power of
including groups of parameters in the model. Selected critical points
for the F distribution are:

	$\alpha = .05$		$\alpha = .001$	
	Denominator d.f.		Denominator d.f.	
	120	∞	120	∞
Numerator d.f.				
3	2.68	2.60	5.78	5.42
5	2.29	2.21	4.42	4.10
6	2.18	2.10	4.04	3.74
120	1.35	1.22	1.77	1.45