# Analyst Disagreement, Forecast Bias and Stock Returns

Anna Scherbina

Finance Department
Harvard Business School

First version: June 2003

Final version: March 2004

Correspondence: Anna Scherbina, Finance Area, Harvard Business School, phone: (617) 495-1047, email: ascherbina@hbs.edu. This paper is based on the second part of my Ph.D. dissertation. I would like to thank the members of my dissertation committee: Kent Daniel, Ravi Jagannathan, Kathleen Hagerty, Todd Pulvino, and Beverly Walther for their consistent guidance and support. I am also grateful to Torben Andersen, Bob Korajczyk, Joel Lander, Tom Lys, Bob McDonald, Iwan Meier, Christopher Polk, Erik Stafford, and Volker Wieland for many valuable discussions and suggestions. Finally, I would like to thank I/B/E/S for making its dataset available for my research. I am responsible for any errors.

## Abstract

I present evidence on inefficient information processing in equity markets by documenting that biases in analysts' earnings forecasts result in biased stock prices. In particular, I show that investors fail to fully account for optimistic bias associated with analyst disagreement. This bias arises for two reasons. First, analysts issue more optimistic forecasts when earnings are uncertain. Second, analysts with sufficiently low earnings expectations who choose to keep quiet introduce an optimistic bias in the mean reported forecast that is increasing in the underlying disagreement. Indicators of the missing negative opinions predict earnings surprises and stock returns. By selling stocks with high analyst disagreement institutions exert correcting pressure on prices.

# I. Introduction

This paper examines whether the marginal investor correctly accounts for the biases in analyst earnings forecasts introduced by analysts' incentives. It has been empirically documented that analysts' opinions are reflected in stock prices.[1] Despite their influence, analysts often operate under incentives that are inconsistent with telling the truth. As less sophisticated consumers of analysts' opinions, retail investors might fail to make proper adjustments for the more nuanced of the resulting biases.[2] To the extent that arbitrage is limited, some of these biases might be reflected in market prices. Thus, a study of analysts' incentives, resulting forecast biases and their potential impact on market prices will shed light on the scope of market efficiency.

Sell-side analysts are pressured to issue optimistic forecasts and recommendations for several reasons. First, their compensation is tied to the amount of trade they generate for their brokerage firms. Given widespread unwillingness or inability to sell short, more trade will result from a "buy" than from a "sell" recommendation. Second, a positive outlook improves the chances of analysts' employers winning investment banking deals.[3] Third, being optimistic has historically helped analysts obtain inside information from the firms they cover. While all these pressures introduce an optimistic bias to analysts' views, the magnitude of the bias is held in check by reputational concerns. Ultimately, an analyst's livelihood—the ability to generate trades and attract investment banking business—depends on her credibility.

The tradeoff between career concerns and the pressure to be optimistic generates predictable patterns of forecast bias. In particular, optimistic forecast bias is increasing in the uncertainty of the underlying earnings. This bias has two components, one that analysts deliberately add to their private estimates, another that arises when sufficiently negative views are kept quiet.[4] Feeling less accountable in uncertain

[1] Evidenced by, among other things, stock price reactions to forecast and recommendation revisions and to corporate announcements of earnings that diverge from analysts' expectations (Abdel-khalik & Ajinkya (1982), Forbes & Skerratt (1992), Givoly & Lakonishok (1979), Gleason & Lee (2000), Gonedes et al. (1976), Hawkins et al. (1984), Imhoff & Lobo (1990), and Stickel (1990)).

[2] Odean (1998), Barber & Odean (1999) and Barber & Odean (2000) document the bounded rationality of small investors.

[3] See the September 6, 2002, *The Wall Street Journal* article "Analysts Pressured on Reports." Also consider a series of recent scandals involving influential analysts being paid for the investment banking deals they helped to generate. Finally, see Bradshaw et al. (2003) and Cowen et al. (2003) for documenting that the prospect of getting investment banking deal induces analysts to express more optimistic views.

[4] McNichols & O'Brien (1997) call this phenomenon "self-selection in analyst coverage" and provide evidence of its existence.

environments, analysts are inclined to issue more optimistic forecasts. High uncertainty will generally also have an effect of spreading out private estimates of future earnings. Those who choose to keep quiet will consequently be lower relative to the average forecast, biasing the mean of the reported forecasts further up. Thus, both components of optimistic bias—the one analysts add deliberately and the one due to self-selection—will be increasing in the underlying uncertainty and in the dispersion of the reported forecasts.

To empirically assess whether the marginal investor adjusts for this bias I construct predictors of the earnings surprises based on this theory and investigate whether they also forecast stock returns. I estimate the bias that arises due to self-selection in coverage based on the decrease in analyst following of a firm over the past three months. I further assume that the reported forecast distribution is a truncated normal and the absent analysts would have issued forecasts in the missing left tale of the distribution. I then compute the standard deviation of thus defined "true" forecast distribution and assume that it represents the underlying level of uncertainty. The estimate of the bias is a significant negative predictor of the future earnings surprise, with a coefficient close to one, implying that is correctly captures the magnitude of the bias. That the bias is also negatively related to abnormal returns around earnings announcement days suggests that the marginal investor does not adjust earnings forecasts for self-selection. Among the stocks in the highest quintile of forecast dispersion, those that have experienced a decrease in analyst following over the past three months earned on average a 4.8% lower annual return than the stocks with a nondecreased analyst following.

Another indication that the bias due to self-selection is factored into stock prices is that the right-skewed earnings forecast distribution is a negative predictor of earnings surprises and stock returns. Right-skewness of the forecast distribution indicates the absence of negative outliers from the reported forecasts.

A further confirmation that forecast biases are incorporated into stock prices is that the past-quarter revisions of the mean earnings forecast predict earnings surprises and stock returns. Both components of the bias will decrease gradually over the year as the earnings uncertainty goes down. Not surprisingly, firms with initially high levels of forecast dispersion experience the most pronounced future downward forecast revisions. But these downward revisions are not sufficient to fully eliminate the optimistic bias.

In fact, the optimistic error of the mean earnings forecast is increasing in the magnitude of the downward revision, implying that analysts are extremely reluctant to fully reflect bad news in their forecasts. That past-quarter forecast revisions also predict abnormal returns around earnings announcement days implies that the marginal investor does not fully adjust for analysts' tendency to "underrevise" their forecasts down.[5]

The empirical findings described above are evidence that the marginal investor does not fully adjust for the self-selection in analyst coverage. That this component of the bias is increasing in the forecast dispersion, implies that at least part of the predictive power of dispersion on future returns arises because the marginal investor does not properly adjust for its correlation with the bias. Therefore, this part of its predictive power on returns is an irrational phenomenon.[6] Sophisticated investors should then be trading against the mispricing. Among users of analysts' forecasts, institutional investors, in particular, mutual funds that engage in independent research are better positioned to understand inherent biases. But due to short-selling restrictions mutual funds are able to sell only shares they already own. Moreover, they will be able to correct mispricing only if they sell the stock in quantities sufficient to affect prices. I investigate whether institutional trades in response to dispersion affect dispersion's predictive ability on future returns. My finding that institutions on average reduce their holdings of a stock in response to increasing dispersion confirms the view that high-dispersion stocks tend to be overpriced. For the subset of stocks for which the reduction in holdings has been significant dispersion has no predictive power on future returns, indicating that institutions corrected potential mispricing through trades. This is evidence that sophisticated investors play a correcting role in the stock market.[7]

[5]The fact that earnings forecast revisions are gradual and insufficient might serve as an alternative explanation for the finding of Chan et al. (1996) that earnings revisions predict stock returns.

[6]Diether et al. (2002) document that dispersion is negatively related to future returns, although the authors use a slightly different explanation for this phenomenon. They interpret analyst disagreement as indicative of disagreement among investors and invoke the Miller (1977) argument that in the presence of short-sale constraints, stock prices will reflect the view of the more optimistic investors. When the disagreement is resolved, prices converge down to the fundamentals, earning low returns. In his intriguing paper, Johnson (2004) suggests that the phenomenon can be explained without invoking irrationality. He interprets dispersion in analysts' forecasts as a proxy for uncertainty in the underlying cash flows. Equity, which in the presence of debt has the payoff structure of a call option, increases in value with the uncertainty in the firm cash flows. To the extent that the cash flow risk is idiosynchratic it will be unpriced. Thus, the higher the uncertainty the lower will be the expected return on equity. As the uncertainty is resolved, equity falls in value and the required rate of return increases.

[7]Sadka & Scherbina (2004) argue that high-dispersion stocks persisted to be mispriced because arbitrage is costly. They show that high-dispersion stocks have significantly higher trading costs than otherwise similar stocks. One possible explanation is that the market marker perceives that some investors are better informed about how to aggregate analyst opinions.

## II. Data Description

Analysts' earnings forecasts are taken from the Institutional Brokers Estimate System (I/B/E/S) U.S. Detail History and Summary History datasets. The latter contains summary statistics on analyst forecasts, including mean, median, and standard deviation of the forecasts as well as information about the number of analysts making forecasts and the number of upward and downward revisions. These variables are calculated on (ordinarily) the third Thursday of each month. The Detail History file records individual analyst forecasts and dates of issue. Each record also contains a *revision date*, that is, the date on which the forecast was last confirmed to be accurate.

The standard-issue Summary and Detail files have a data problem that makes them unsuitable for the purposes of this paper.[8] In the standard-issue datasets, earnings per share forecasts are adjusted by I/B/E/S for stock splits and stock dividends since the date of the forecast in order to smooth the forecast time series. The adjusted number is then rounded to the nearest cent. For firms that had large numbers of stock splits or stock dividends, earnings per share forecasts (and the summary statistics associated with earnings) will be reported as zero. These will also tend to be the firms that did well ex-post. Observations with the standard deviation of zero (and/or mean forecast of zero) will thus include firms that have earned high future returns (which is what is actually observed in the data). To avoid inadvertently using this ex-post information I rely on forecasts not adjusted for stock splits produced by I/B/E/S at my request.

Data on stock returns, prices, and shares outstanding are from the daily and monthly stocks files of the Center for Research in Security Prices (CRSP). The accounting data are from the merged CRSP/Compustat database, extended through fiscal year 2002. If less than three months has elapsed since the latest fiscal-year-end date, accounting data for the preceding year is used.

Book value of equity is calculated using Compustat annual data (including the Research file). I use total common equity, if available, plus balance sheet deferred taxes and investment tax credit. If total common equity is not available I use shareholder's equity minus the value of preferred stock. For preferred stock I use redemption value, liquidating value, or carrying value in that order, as available. The book-to-market

[8] This problem was first reported in Diether et al. (2002).

ratio is defined as the ratio of book value to market value of equity. The latter is calculated as the product of month-end share price and the number of shares outstanding.

Stocks with high dispersion tend to be smaller, possibly because smaller stocks are more opaque. After controlling for size, stocks with high dispersion tend to have higher analyst coverage, possibly because there is more demand for expert opinion when it is difficult to interpret available information. High-dispersion stocks tend to be value stocks that have done poorly in the past and have higher systematic risk.[9]

To minimize the problem of bid-ask bounce I use stocks priced at no less than $5 per share. Because I am interested in dispersion in analysts' earnings per share forecasts, I consider only stocks in the I/B/E/S database that are followed by at least two analysts. As of January 1981 the number of stocks priced above $5 per share and followed by at least two analysts at the intersection of I/B/E/S and CRSP was 1,239. Of these, 858 were in the lower nine NYSE market-capitalization deciles. As of January 1983 the number of stocks at the intersection of I/B/E/S and CRSP, priced above $5 per share and followed by at least two analysts, had grown to 1,401. Of these, 962 were ranked in the lowest nine NYSE market-capitalization deciles. The numbers at the end of 1999 were 3,466, 2,525 in the lowest nine NYSE market capitalization deciles. At the intersection of the I/B/E/S, CRSP, and Compustat datasets the pattern is similar, although the total number of available observations is lower because Compustat contains only a subset of the stocks in CRSP. The number of stocks at this intersection priced above $5 per share and followed by at least two analysts grew from 1,049 in January 1981 to 1,178 in January 1983 and to 1,979 in December 1999. A more complete sample description is available in Table I of Diether et al. (2002). Although the data go back to 1976, the number of stocks in the I/B/E/S cross-section increases more than threefold between 1976 and 1983. I use data from January 1983 through December 2000 for tests that involve portfolio construction to allow for a larger cross-section of stocks.

I use Compustat's Quarterly Industrial and Annual Industrial datasets for firms' annual and quarterly earnings per share numbers. The I/B/E/S Summary file reports earnings-per-share forecasts on the diluted basis. That some of Compustat's reported quarterly diluted earnings per share observations do not add

[9] Ciccone (2003).

up to the annual number presents a slight problem. Explained a Standard and Poors representative who compiles the data:

"This inconsistency is due to the fact that companies were not required to report fully diluted EPS on a quarterly basis unless there was a significant (>3%) difference between it and the primary number. It was the same on an annual basis, but companies usually reported both numbers anyway. In 1983 it was our policy to carry the primary EPS number over to the fully diluted number when the fully diluted number was not reported. For IBM in 1983 this fully diluted number was not reported quarterly, but was annually, which would explain the inconsistencies."

Given this noise in the diluted reported earnings I convert analysts' forecasts to primary basis whenever I use comparisons with the realized earnings per share numbers.

Data on institutional holdings are from the Spectrum database. Spectrum collects quarterly data on stock holdings from the 13F reports institutions are required to file if their holdings exceed $100 million. I aggregate these holdings over all institutions to arrive at the institutional holdings number.

## III. Analysts' Incentives and Forecasts

Because sell-side analysts are not paid directly by investors and historically depended on firm managements for inside information, their incentives are not always consistent with telling the truth. Analysts build influence and reputation on the accuracy of their earnings per share forecasts and usefulness of their buy/sell recommendations, but their compensation is based on the profits they help generate for the brokerage firms that employ them. Consider the following quote from the article "Analyze This," published in the June 2001 issue of *Smart Money Magazine*. "At First Union and most other banks, research analysts' compensation is completely unrelated to their stock picking or their earnings estimates ... The real money—their bonus—is determined by how much trading they bring in for the sales force and, more important, how much business they generate for the firm's investment bankers." Analysts care about their reputation to the extent that it can be deployed to generate trades and attract investment banking business.

Due to widespread unwillingness or inability to sell short, analysts generate stock purchases more easily than sales. Incidentally, mutual funds, the client group with resources to generate large trades, are precluded by regulation from selling short. Analysts therefore prefer to focus on stocks for which they can issue a "strong buy" or "buy" recommendation. McNichols & O'Brien (1997) observe that analysts usually initiate coverage of stocks about which they feel optimistic and drop coverage of stocks about which they feel pessimistic.

Analysts' ties to investment banking recently received considerable public attention. Most notable is the April 2003 $1.4 settlement that grew out of charges that ten Wall Street firms misled investors by issuing overly optimistic stock research in bids to win stock-underwriting assignments and other investment banking business. The problem has been long in the making. Reported *Business Week* in the October 8, 1998 article "Wall Street's Spin Game": "Most Wall Street research is pitched to institutional investors who pay the firm about a nickel a share in commissions. But if an analyst spends his time trying to land an initial public offering, the firm can earn 15 to 20 times that amount per share. Investment banking deals are much more lucrative for the brokerage firm. Merger advisory fees can be sweet as well."

Even when their compensation is not directly tied to the investment banking deals they help generate, analysts are often pressured to issue positive reports about prospective investment banking clients. The September 6, 2002, *Wall Street Journal* article "Analysts Pressured on Reports" documents that analysts who issued downgrades were chided by their bosses and even denied coverage of the more lucrative sectors. One analyst was informed by an investment banker of the "unwritten rule number one: 'If you can't say something positive, don't say anything at all."' The "unwritten rule number two" is: "go with the flow of other analysts rather than try to be contrarian." Both rules discourage freely reporting negative opinions.

Managers can pressure an analyst who covers their firm to issue optimistic reports by denying access to inside information. Studies show that analysts' earnings forecasts contain private information in addition to a statistical model based only on public information. Not surprisingly, it is access to non-public information investors value most.[10] The Regulation "Fair Disclosure," adopted in October 2000 to make

[10]Institutional Investor, a firm that compiles annual analyst rankings, ranks access to management sixth out of thirteen valuable analyst attributes, ahead of accuracy of earnings estimates, written reports, stock selection, and financial modeling.

analysts more independent of management, forbids selective disclosure of information. But is has already failed in several instances. One is described in the October 2, 2001 *New York Times* article "In a Surprise Move, AOL Replaces Its Chief Financial Officer". "Late last month, after the company reported its third-quarter financial results, two prominent Merrill Lynch analysts, Jessica Reif Cohen and Henry Blodget downgraded the company's stock... That decision drew an angry reaction from Mr. Kelly [the AOL CFO] according to several analysts, who said that the company stopped returning telephone calls from Ms. Cohen and Mr. Blodget."[11]

These pressures notwithstanding, analysts are constrained from adding an arbitrarily high optimistic bias to their private estimates by fear of hurting their reputations with investors. Analysts will set the optimistic bias at an optimal point that balances the benefit of being upbeat against the cost to their reputations. If, in addition to being penalized for their forecast errors, analysts face an additional penalty (e.g. being cut off from the sources of inside information) for being too pessimistic relative to others, those with sufficiently low private estimates may decide to drop coverage altogether.

### A. A model of analysts' incentives

The historical earnings distribution is left-skewed.[12] Analysts forecast mean expected earnings. They base their forecasts on both historical data and signals about the future earnings. According to the Central Limit Theorem, an analyst's signals about a firm's expected earnings should be normally distributed. And so should be all analysts' earnings forecasts.[13]

Suppose, then, that all analysts and investors receive a public signal about next year's earnings per share that is normally distributed around the correct mean with the standard deviation of $\sigma_0$: $s_0 \sim N(EPS, \sigma_0)$. Each analyst, $i$, also receives a private signal, $s_i$, which is independent of the public signal. Suppose that all private signals are distributed around the correct mean with the same standard deviation $\sigma_s$, $s_0 \sim N(EPS, \sigma_s)$. An analyst combines the private and public signals to come up with an earnings estimate

[11] One must wonder how severe the conflict of interest might be for firms that pay analysts to cover them.

[12] One possible explanation is that firms engage in earnings smoothing to avoid negative earnings and when that does not work, the negative earnings they report tend to be very low.

[13] According to recent accounting papers, such as that of Gu & Wu (2003), analysts forecast median rather than mean earnings. But even then their earnings estimates should be normally distributed.

that would have the minimum variance: $E_i[EPS] = \frac{\sigma_s^2}{\sigma_0^2+\sigma_s^2} s_0 + \frac{\sigma_0^2}{\sigma_0^2+\sigma_s^2} s_i$. This forecast has a variance of $\sigma^2 = \frac{\sigma_s^2\sigma_0^2}{\sigma_s^2+\sigma_0^2}$. If the precision of private relative to public signal is the same across all analysts and firms, $\frac{\sigma_s}{\sigma_0} = \nu$, then there is a linear relationship between the standard deviations of analysts' earnings estimates and the public signal: $\sigma = \sigma_0\sqrt{\frac{\nu^2}{\nu^2+1}}$.

An analyst derives utility from issuing optimistic forecasts, but is penalized for forecast errors. Additionally, a high penalty, *C*, is exacted if an analyst forecast falls *k* standard deviations of the public signal below the mean of the public signal. If *f* is the reported earnings per share forecast and *EPS* the realized earnings per share, the utility function of the analyst is:

$$U(f) = f - EPS - \frac{\alpha}{\sigma_0}(f - EPS)^2 - C|f < s_0 - k\sigma_0 \tag{1}$$

Analysts set their forecasts to maximize their expected utility function:

$$EU(f) = f - E[EPS] - \frac{\alpha}{\sigma_0}\left(f^2 - 2fE[EPS] - E[EPS]^2 - \frac{\sigma^2}{2}\right) - C|f < s_0 - k\sigma_0 \tag{2}$$

The optimal forecast depends on how high an analyst's private earnings estimate, *E[EPS]*, is relative to the point at which the penalty C is imposed:

$$f^* = \begin{cases} E[EPS] + \frac{\sigma_0}{2\alpha} & \text{if} \quad E[EPS] \geq s_0 - k\sigma_0 - \frac{\sigma_0}{2\alpha} \equiv b \\ s_0 - k\sigma_0 & \text{if} \quad a \equiv s_0 - k\sigma_0 - \frac{\sigma_0}{2\alpha} - \sqrt{\frac{\sigma_0^2}{2\alpha^2} + \frac{\sigma^2}{2}} \leq E[EPS] < b \\ \text{do not report} & \text{if} \quad E[EPS] < a \end{cases} \tag{3}$$

The expected optimistic bias in an analyst's forecast is increasing in the level of uncertainty of the earnings estimate, $\sigma$ (and, hence, the level of uncertainty of the public signal, $\sigma_0$, since the two are proportional). There are two reasons for the positive correlation between the optimistic forecast error and level of uncertainty. First, the optimal bias analysts add to their earnings estimates is increasing in the level of the uncertainty of the public signal and, hence, the uncertainty of their estimates. Second, the component of the bias caused by self-selection in analyst coverage also increases in forecast uncertainty. Unreported low forecasts are likely to be lower relative to the true mean of the forecast distribution the more spread out the distribution, thus creating a positive relationship between the bias and dispersion of the reported distribution.

This intuition can be proved mathematically. The expected bias in a reported forecast can be calculated as:

$$E[f^* - EPS] = \frac{1}{\Phi(-a/\sigma)} \left( \int_b^\infty \left( \frac{\sigma_0}{2\alpha} + eps \right) f(eps) d(eps) + \int_a^b (s_0 - k\sigma_0) f(eps) d(eps) \right) \quad (4)$$

(The expression is divided by $\Phi(-a/\sigma)$ to arrive at the probability density function of the reported forecast distribution, which is truncated at point $a$ addable to 1.) It can be shown that the expected bias in analysts' forecasts is increasing in the level of underlying uncertainty about analysts' earnings estimates, $\sigma$.[14]

Given $k$ sufficiently large that the cut-off point $a$ below which analysts do not report their forecasts falls either below or slightly above the expected mean forecast, it can be shown that dispersion in the reported forecasts will be increasing in the level of forecast uncertainty, $\sigma$. The bias in the outstanding forecasts will thus also be increasing in the dispersion of the outstanding forecasts.

According to the model the bias in the outstanding forecasts has two components: one that analysts deliberately add to their private estimate, and one due to self-selection in analyst coverage. Investors might have an easier time correcting for the first than for the second component of the bias because to do so requires less data and skill. In this paper I try to determine the extent to which the second type of the bias is reflected in market prices. In the empirical analysis I use dispersion in analysts' forecasts as a proxy for the underlying level of uncertainty, and a measure based on the recent decrease in analyst coverage or skewness of the forecast distribution proxy for the component of the bias that is due to self-selection in coverage.

## B. Patterns of analyst forecasts

Figure 1 confirms the positive relationship between forecast bias and dispersion in outstanding forecasts. It plots the average dispersion-based portfolio statistics as a function of the number of months remaining until fiscal year end. The variables are first averaged over the firms in the portfolio and then over the years (the sample is 1983-2002). The variables plotted are the average forecast error in the mean, highest and lowest forecast, and average portfolio dispersion. Forecast error is defined as the forecast value minus the

[14] Proof is available upon request.

ex-post realized earnings value scaled by the book value of equity. Dispersion is defined as the standard deviation of outstanding forecasts scaled by book value of equity as of the previous fiscal year. Stocks with book values of less than $3 are excluded from the sample, as are observations for which the absolute value of the forecast error is greater than 30% of book equity. This is done to exclude extreme outliers that might be attributable in part to I/B/E/S data errors.

Portfolios are based on dispersion in analysts' earnings forecasts and include only stocks with December fiscal year end. Portfolio *D1* consists of stocks in the lowest dispersion-based quintile, *D5* of stocks in the highest dispersion-based quintile. In the graphs in the left column portfolios are re-formed monthly. In the graphs in the right column portfolios are formed in February of each year and held until the end of the year.

As can be seen from the first panel, bias increases in portfolio dispersion as predicted by the model. As the quarterly earnings numbers come out during the fiscal year, the uncertainty surrounding annual earnings declines and with it the forecast dispersion (bottom panel). Dispersion declines faster for the portfolios formed in February and held for the entire year than for portfolios formed monthly. As dispersion declines, so does the optimistic forecast error. The average future forecast revision is more negative for stocks in the high-dispersion portfolio than for stocks in the low-dispersion portfolio, and the difference is statistically significant. (If the higher-than-usual magnitude of the future downward forecast revisions is unanticipated by the marginal investor, the stock price reaction will be negative.)

From the middle two graphs it can be seen that the highest and lowest forecasts are not distributed symmetrically around the ex-post realized earnings. In fact, even the lowest forecast tends to be slightly optimistic, and more so for the lower-dispersion portfolios. This indicates that analysts add a deliberate bias to their estimates in the beginning of the fiscal year, expecting to revise their forecasts down during the year, or that truly pessimistic analysts do not express their opinions, or a combination of both.[15]

[15] Richardson & Wysocki (2001) hypothesized the former to be the case.

### C. Self-selection in analyst coverage and bias

A direct implication of self-selection in analyst coverage is that the fewer analysts issuing forecasts, the higher the optimistic error in the mean outstanding forecast. The number of missing analysts can be calculated in two somewhat related ways. The first is to effect a cross-sectional comparison of analyst coverage at a given time. This method was first employed by Hong et al. (2000), who predicted analyst coverage for a given set of firm characteristics and took the deviations from the predicted number to be the missing (or excess) coverage. The second method is to use the time series variation in analyst coverage for a given firm to indicate missing coverage. Because the number of analysts issuing estimates is slightly lower at the beginning of the fiscal year, it is important to calculate changes in coverage *within* a fiscal year. The two methods should produce similar results. If the number of analysts has decreased in a time series, the cross-sectional results will likely indicate lower than normal coverage. When regressing the error of the mean forecast on the missing number of analysts, both methods produce a significant positive coefficient (meaning that the more analysts missing from coverage, the higher the optimistic error of the consensus forecast).[16]

In the remainder of this paper I use the time series variation in analyst coverage to estimate the number of missing analysts. If the number of analysts following a firm this month is lower than the number of analysts following the firm three months ago but in the same fiscal year, I set the difference to be the number of missing analysts. Otherwise, it is set to zero.

I then estimate the bias in outstanding earnings per share forecasts implied by the number of missing analysts. I estimate the bias using two different methods that yield similar results. The first assumes the distribution of reported forecasts to be a truncated normal and the amount of the bias caused by self-selection to be equal to $\frac{\phi(k)}{\Phi(k)} * \sigma^{True}$, where $k$ is the point below which analysts will not report their forecasts, estimated by inverting the standard normal CDF $\Phi(k) = \frac{missing}{reporting+missing}$, and $\sigma^{True}$ is the standard deviation of the true, not reported distribution that is assumed to be truncated. The dispersion of the true distribution is estimated under the assumption that the observed distribution is a truncated normal, so the true dispersion

[16]The results of these regressions are not reported.

will be somewhat higher: $\sigma^{True} = \sigma\left(\Phi(-k) + f\frac{\phi(k)}{\Phi(-k)}\right)$. The second method of estimating the bias assumes that the analysts who are no longer following the stock will have forecasted earnings lower by one cent than the currently lowest outstanding forecast. Estimated bias is scaled by the book value of equity. Additionally, the estimates are scaled by the square root of the number of quarters remaining until the end of the fiscal year in order to allow comparisons across firms with different fiscal year ends.

Table I provides a sample description of estimated bias as well as of other variables used in the paper. As can be seen from the table, the second method of estimating the bias produces a number roughly six times larger than the number produced by the first method. This is because the second method assumes that missing analysts will have issued forecasts one cent below the *lowest* forecast, whereas the first method simply places the missing forecasts at a certain number of standard deviations below the mean. If a forecast distribution is not exactly a truncated normal but has a negative outlier, the bias estimated using the second method will be much larger. It can also be seen from the table that relatively few observations have missing analysts. Even the 75th percentile observations have zero missing analysts.

In the remainder of the paper I report mostly results based on the second method of estimating the bias due to self-selection in analyst coverage. The second part of Table I shows a positive relationship between estimated bias and dispersion in outstanding forecasts. This is not surprising, as there needs to be a spread between the mean and lowest outstanding forecast in order to produce a substantial difference in the mean by placing the missing forecasts at one cent below the lowest. That estimated quarter-end bias is also likely to be preceded by a negative revision in the mean outstanding forecast over the previous quarter is consistent with the view that analysts drop coverage of firms that are trouble. Decreased analyst coverage over the course of the quarter leads to higher-than-average trading volume in the three-day window around the quarterly earnings announcement days. Finally, self-selection in analyst coverage tends to be more pronounced for smaller growth firms (which, presumably, are less transparent than larger, more established firms).

Table II reports the results of regressing the earnings surprise on a set of predictive variables. This regression is run using fiscal-year-end accounting variables. Earnings surprise is defined as the realized earnings per share number minus the latest available mean analyst forecast scaled by the book value of equity. The predictive variables used in the regression are dispersion in analyst forecasts, bias due to self-selection in analyst coverage (estimated using the second method), revision in the mean annual forecast over the past quarter, logarithm of the firm market capitalization level, and book to market ratio.

Analysts having been, on average, optimistic during the 1983-2002 time period, the average earnings surprise is negative. Consistent with the model of analysts' incentives, dispersion is a significant negative predictor of the earnings surprise. Estimated bias comes in significantly negative as well. That revision of the mean forecast over the past quarter is a positive predictor of the earnings surprise suggests that the magnitude of the past revision is informative of the forecast error because all analysts fail to completely incorporate new information into their forecasts. Finally, the size of the earnings surprise is less negative for large value firms.

## IV. Do Investors Adjust for the Biases in Analysts' Forecasts?

In this section I investigate whether biases in analysts' forecasts find their way into stock prices and whether the marginal investor is able to disentangle forecasts from these inherent biases. It is plausible that investors adjust for some obvious types of biases but not for the ones that require more sophisticated analysis. For example, analysts' forecasts are, on average, optimistic. If investors did not adjust for this the average price reaction around earnings announcement days would be negative. But, as seen in Table I, it is in fact slightly positive.[17] More nuanced biases are not as easily detected. For example, recognition that analyst disagreement or diminished analyst coverage might engender optimistic bias relies on investor access to an historical cross-section of forecasts across firms and analysts. Retail investors, who most likely observe only one analyst forecast at a time, will find it difficult to estimate the magnitude of the bias that is correctly conditioned on the fact that a forecast has been issued. Institutions with access to research resources are better positioned than retail investors to determine and trade against the biases in analysts' forecasts.

[17] Chari et al. (1988) suggest that the positive abnormal returns around earnings announcement days might be viewed as compensation for risk.

In the following sections I document that biases in analyst forecasts, both deliberate as related to uncertainty and unintentional due to self-selection, are reflected in stock prices.[18] I also show that institutions, when they sell shares in response to dispersion in analyst forecasts, ameliorate the mispricing.

## A. Evidence from returns around the earnings announcement days

Quarterly earnings announcements reveal a great deal about annual earnings. When investors' expectations diverge from reality the strongest price reaction should occur around the quarterly earnings announcement days. For firms with too high earnings expectations the price reactions should be negative.

I assume that all returns can be described by the market model with $\beta = 1$:

$$r_{it} = r_{mt} + \varepsilon_{it} \tag{5}$$

where $r_{it}$ and $r_{mt}$ are daily log-returns on an individual stock and the CRSP equally-weighted index, respectively.[19]

Assuming day 0 to be the announcement date, the cumulative abnormal return ($CAR$) is estimated as:

$$CAR_{it} = \frac{1}{3}\sum_{\tau=-1}^{\tau=1}(r_{i\tau} - r_{m\tau}) \tag{6}$$

Table III reports the results of regressing abnormal returns around quarterly earnings announcement days on predictive variables. Panel A presents regression results for the entire sample, 1983-2002. It can be seen from the table that both dispersion in analysts' forecasts and estimated bias due to self-selection in analyst coverage are significantly negative when the past forecast revision is not included in the regression. The past quarter forecast revision is a significantly positive predictor of the price reaction around earnings announcement days and subsumes the predictive power of both dispersion and estimated bias. Revision of the mean forecast, usually negative, can be interpreted as a measure of the initial forecast error. Both dispersion and estimated bias being strongly related to the positive error in the mean earnings forecast, it is not surprising that their significance declines when the past forecast revision is included in the regression.

[18] The predictive power of dispersion on returns is also open to the interpretations of Diether et al. (2002) and Johnson (2004).

[19] Brown & Warner (1985) find this model to perform adequately relative to more sophisticated models.

Returns around earnings announcements tend to be higher for smaller stocks. This result is consistent with Chari et al. (1988), who conjecture that for smaller firms that are less transparent most information becomes available at the time of the earnings announcements. This is when investors are exposed to and compensated for bearing the highest risk. Abnormal returns also tend to be higher for growth than for value firms, perhaps for the same reason. These results are driven by the outliers and go away when extreme price movements are excluded from the sample (Panel E).

Finally, high trading volume is, on average, associated with negative price reaction. This result might be due to the few observations with extreme negative movements in stock prices. As can be see from Panel E, when extreme price movements are eliminated abnormal volume turns out to be positively correlated with announcement-day returns. With the extreme observations removed, it appears that good news generates more trade and a larger immediate price reaction, and bad news leads to less immediate trade, but might still result in a more pronounced post-earnings-announcement drift. These questions should be addressed by future research.

Panel B considers only observations with at least 10 analysts. That these are relatively large stocks for which more information is available is perhaps the reason dispersion in analysts' forecasts loses its significance as a predictor of returns. The estimated forecast bias nevertheless remains a strongly negative, significant predictor of abnormal returns around earnings announcement days. This is perhaps due to the fact that the estimate of missing analysts is more reliable if analyst coverage is high to begin with.

Panels C and D present regression results for the later and earlier parts of the sample, respectively. The predictive power of dispersion is strong in the earlier part of the sample, but disappears in the later part. In contrast, bias due to self-selection is not significant in the earlier part of the sample, but gains significance later on. This must be due to the fact that it becomes a better predictor of forecast error later in the sample (the results are not reported here), perhaps due to the change in analysts' incentives. And self-selection in analyst coverage became more of an issue later in the sample, investors had not yet learned to adjust for it. Dispersion remained a significant negative predictor of forecast error throughout the years but lost its predictive power on future returns as investors learned to recognize it.

## B. Trading strategies

In this section I explore whether investors' failure to adjust forecasts for self-selection in analyst coverage can be profitably exploited. Table IV reproduces the finding first reported in Diether et al. (2002) that dispersion in analysts' forecasts predicts future stock returns. Every month stocks are subdivided into five groups based on the level of market capitalization as of the previous month.[20] Within each size group stocks are sorted into five portfolios based on dispersion in outstanding analyst earnings forecasts as of the previous month. Dispersion is defined as earlier in the paper (standard deviation scaled by the book value of equity and the square root of the number of quarters remaining until the fiscal year end). The table reports alphas of the Fama & French (1993) three-factor model.[21] High-dispersion portfolios consistently earn negative abnormal returns, thus producing positive returns on the strategy of buying low-dispersion and selling high-dispersion stocks.[22] The abnormal return of this strategy is highest for smaller stocks, 1.50% per month for the smallest size quintile, and insignificant for the largest size quintile. The return differential between low- and high-dispersion stocks in the entire I/B/E/S universe is 0.75% per month on average. Among several reasons for this is that analysts produce more upwardly biased forecasts for smaller stocks (see Table II) and investors assign more weight to analysts' forecasts when forming valuations of small stocks because less information about them is available independently. Additionally, smaller stocks usually incur higher costs of short-selling, thus making it difficult for rational investors to arbitrage away the mispricing.

Table V documents that a trading strategy based on self-selection in analyst coverage also produces abnormal returns. The bias in the mean outstanding forecast is estimated according to the first method (which assumes that the observed distribution of forecasts is a truncated normal and that analysts who stopped coverage within the past three months will have issued forecasts within the missing region of the

[20] I form my own size-based portfolios rather than sort based on the NYSE market capitalization deciles to obtain an equal number of stocks in each portfolio. Table III of Diether et al. (2002) reports portfolio returns based on the NYSE market capitalization breakdowns. Remarkably, there are almost no stocks in the lowest three deciles and only 135 stocks in the fourth decile.

[21] Returns here are equally-weighted, but value-weighting produces quantitatively similar results because stocks are pre-sorted by size.

[22] Scherbina (2001) shows that a dispropotionally large portion of future underperformance of high-dispersion stocks falls in the 3-day windows around quarterly earnings announcement days, when most of the uncertainty about annual earnings is resolved.

distribution).[23] Stocks are sorted into five size-based groups and then into three bias-based portfolios. As most observations do not exhibit decreased analyst coverage within the past three months, most firm-date observations fall into the zero-bias portfolio. Returns for smaller firms are biased slightly up since I require that a firm be present in the I/B/E/S database one month before returns are calculated. Were all analysts to drop coverage, the firm might still experience low future returns, but will not show up in the portfolio. Consequently, for smaller firms there are fewer observations with estimated bias (28 on average over the time period). The number of stocks in the positive-bias portfolio is increasing with size, up to 63 stocks per portfolio, on average, for the highest size portfolio. As can be seen from the table portfolios of stocks with the highest estimated bias earn negative abnormal returns. Returns on high-bias stocks increase with size but, unlike dispersion-based sorting, remain significant even for stocks in the largest size quintile.

To show that bias due to self-selection in analyst coverage has predictive power on returns after controlling for forecast dispersion, I sort stocks independently into five dispersion-based groups and groups with zero and positive estimated bias. Table VI presents portfolio returns. As expected, self-selection in analyst coverage leads to high bias and, hence, low future returns only when the level of analyst disagreement is high as only then will the missing opinions deviate significantly from the reported mean. Positive-bias portfolios earn significantly low returns for the two highest, but not the three lowest, dispersion-based groups. Estimated bias therefore has predictive power on future returns independently of the level of forecast dispersion. A portfolio of stocks with high forecast dispersion and for which analyst coverage has declined over the past few months earns the lowest returns.

## C. Evidence based on skewness in forecast distribution

This section provides further corroboration for the hypothesis that missing negative opinions are responsible for low future stock returns. When negative opinions are missing because relatively pessimistic analysts have dropped coverage or further biased up their forecasts, the reported forecast distribution will be right-skewed. A measure based on skewness in analysts' forecasts will therefore predict the forecast error of the mean outstanding forecast and future low stock returns. This is, indeed, the case in the data.

[23] Estimating bias according to the second method produces a similar result.

I measure skewness as the difference between the mean and median forecast scaled by the book value of equity. At the beginning of each month I sort stocks at the intersection of the I/B/E/S Summary File and the CRSP dataset, with outstanding forecasts by at least three analysts, into five size groups and further into five skewness groups with both size and skewness measured as of the end of the previous month. Table VII shows average monthly returns on the equally-weighted portfolios formed this way. It can be seen from the table that portfolios with positive skewness in the forecasts earn negative abnormal returns in the smallest and third smallest size quintiles. For the entire universe the quintile of stocks with the highest (positive) skewness underperforms the quintile of stocks with the lowest (negative) skewness by, on average, 0.23% per month.

## V. Does the Presence of Sophisticated Investors Ameliorate Mispricing? Evidence from Institutional Holdings Data

As conjectured by Miller (1977) and Viswanathan (2001), the optimistic bias in stock prices will be lower the less constrained pessimistic investors are from selling shares. Institutions, being on average a more sophisticated class of investors, are more likely to understand when high-dispersion stocks are overpriced. Because some institutions, in particular, mutual funds, are prohibited by regulators from selling short they will be able to influence prices only if they own sufficient shares to begin with and are free to substantially reduce their holdings in response to mispricing (which will not be the case for index funds). If institutions are able to sell sufficient shares, mispricing will be corrected and dispersion will not predict future returns. To see if this is the case I check the predictive power of dispersion on future returns dependent on whether institutions as a group have historically traded against dispersion in analysts' forecasts.

I calculate total institutional ownership, $Instit_{it}$, of stock $i$ at time $t$ by summing individual holdings recorded quarterly on SEC 13F reports and divide the result by the total number of shares outstanding to arrive at the fraction held by the institutional sector. I quantify how this fraction responds in a time series to predictors of institutional ownership such as dispersion and price momentum. For each stock $i$ I run the following quarterly time-series regression:

$$Instit_{it} = \beta_{0i} + \beta_{1i}\sigma_{it} + \beta_{2i}ret3mo_{it} + \varepsilon_{it} \quad (7)$$

where $\sigma_{it}$ is the standard deviation in analysts' earnings per share forecasts scaled by the book value of equity and the square root of the number of months remaining until the fiscal year end, and *ret3mo* is the average stock return over the past three months, included to control for price momentum.

Coefficient $\beta_{1i}$ captures the degree to which institutional ownership responds to changes in analyst forecast dispersion for firm $i$. Theory suggests that if institutional ownership responds negatively to dispersion (i.e., if $\beta_{1i}$ is significantly negative) mispricing should be ameliorated. Such stocks should not earn as low risk-adjusted returns as stocks with nonnegative sensitivity of ownership to dispersion ($\beta_{1i} \geq 0$).

To check whether this is the case in the data I sort all stocks at the intersection of Spectrum (the dataset of SEC 13F report filers), the I/B/E/S Summary File, CRSP, and Compustat into three groups based on the p-value of coefficient $\beta_{1i}$ of regression (7). I then sort the stocks independently into five groups based on dispersion in analyst forecasts as of the previous month, defined in the usual way. Resulting portfolio returns are equally-weighted.

Table VIII presents the average Fama-French-adjusted return differential between the low- and high-dispersion portfolios. Consistent with the view that high-dispersion stocks tend to be irrationally over-priced, institutions on average reduced their holdings of a stock in response to dispersion (see the table on the right). For the majority of the firms, however, the negative response is not statistically significant and so is classified as a "hold" response. There are relatively few stocks that institutions bought in response to increased dispersion. Consistent with the hypothesis, the high-dispersion stocks that institutions have historically sold in response to dispersion have not exhibited negative risk-adjusted returns. The stocks institutions held irrespective of dispersion experienced negative abnormal returns over the time period (-0.31% per month with a *t*-statistic of -2.58). The small number of stocks that institutions, on average, bought in response to dispersion, perhaps even further aggravating the mispricing created by retail investors, experienced even more significant negative returns (-0.81% per month with a *t*-statistic of -4.30).[24]

[24]Viswanathan (2001) sorts stocks based on responsiveness of short interest to dispersion and finds that the returns differential between the high- and low-dispersion stocks is lower for the group of stocks with a higher responsiveness coefficient. I repeat the same exercise by estimating the responsiveness of total institutional holdings and the breadth of institutional ownership to dispersion and obtain a similar result: when institutions trade in response to increased dispersion, the return differential is less pronounced.

## VI. Discussion

As discussed in the introduction, there are several plausible explanations for why dispersion forecasts returns. Johnson (2004) argues that it is a rational phenomenon. When a firm is levered, equity has the payoff structure of a call option. The value of a call option will be increasing in the uncertainty of the firm's cash flows, which is captured by dispersion in analyst earnings forecasts. If this uncertainty is idiosyncratic it will not be priced in, lowering the expected returns on equity. When the uncertainty is resolved equity will earn low returns.[25]

Diether et al. (2002) conjecture that analyst disagreement predicts future underperformance because it proxies for differences of opinion among investors. When pessimistic investors are short-sale constrained, optimists end up driving up stock prices to reflect their own valuations and in the future prices fall as the uncertainty is resolved. In this paper I argue that analyst disagreement leads to low future returns because it is correlated with the upward bias in analysts' earnings forecasts that is mistakenly incorporated into stock prices. These arguments are similar in that they rely on the marginal investor being optimistic for whatever reason and rational investors being unable to sell the stock and correct the mispricing.[26]

The results here are consistent with the notion that at least part of the predictive power of dispersion derives from its correlation with forecast bias. The estimate of the bias that is due to self-selection in analyst coverage is positively correlated with dispersion. As evidenced by Table III, it predicts returns even in the presence of dispersion as a control variable. Additionally, the prior-quarter forecast revision subsumes the predictive power of dispersion. One possible explanation is that it is a better predictor of the earnings surprise.

[25] The interaction between leverage and dispersion is also consistent with investors' irrational overpricing of firms with high dispersion for whatever reason. According to the second proposition of Modigliani and Miller, the required return on equity increases with leverage. When investors irrationally overprice equity, they underestimate leverage and thus the required rate of return. When the equity value converges to fundamentals, investors learn about the actual leverage and demand a higher rate of return. Thus, the underperformance should be more pronounced for the more levered firms.

[26] Another plausible conjecture is that analysts are slow to revise their forecasts down in response to bad news, thus causing dispersion in analysts' forecasts, a downward drift in forecast revisions, and a positive earnings surprise in the future. This concern has been addressed in Scherbina (2001) who shows that even when only most recent forecasts, issued in the past 30 days, are considered dispersion still forecasts future returns.

That the negative relationship between forecast dispersion and returns is not a fully rational phenomenon is evidenced by institutional investors trading against it. For stocks for which institutional response to dispersion has been significant, the mispricing appears to have been eliminated. A natural follow-up question is why do institutions sell some overpriced stocks but not others? Several explanations seem plausible. Institutions might be required to hold certain stocks that are part of an index. Other stocks might be highly illiquid such that the price concession institutions will pay for selling them will exceed the expected price decline. It also also plausible that the institutions that hold on to overpriced shares are less sophisticated, i.e. insurance companies rather than mutual funds. Further investigation is required to evaluate these possibilities.

The question of why the mispricing has not been fully eliminated is addressed by Sadka & Scherbina (2004), who claim that it is due to the endogenously high costs of arbitrage. When analyst disagreement is high, the market maker is informationally disadvantaged with respect to some investors on how to aggregate analysts' views. She will consider trades to be informative of the true value of the stock and adjust the price in the direction of the trades.[27] The authors document that high-dispersion stocks have significantly higher costs of trade than low-dispersion stocks, even after controlling for size and momentum. Thus, analyst disagreement leads to an upward bias in forecasts misleading some investors into forming high valuations, but also resulting in high trading costs that render arbitrage unprofitable.

## VII. Conclusion

When the level of analyst disagreement about future earnings is high the average forecast tends to be overly optimistic. I hypothesize that this happens for two reasons. First, high analyst disagreement usually coincides with high uncertainty, which makes analysts less accountable for errors in their forecasts and encourages a adding a higher bias to their private estimates. Second, if some analysts with low private estimates cease coverage the upward bias in the reported forecast will be higher the lower the unreported forecasts are relative to the truth, which would be the case when the underlying forecasts are highly dispersed.

[27] See the Kyle (1985) model for the determinants of trading costs.

I show dispersion in analysts' earnings forecasts to be a significant predictor of the earnings surprise and the price reaction around earnings announcement days. I further compute a proxy for the bias due to self-selection in analyst coverage by estimating the number of missing analysts and placing their collective forecast somewhere below the mean. Two different methods of estimating this bias produce similar results. Both are reliable predictors of the future earnings surprise, price reaction around earnings announcement days, and returns throughout the fiscal year. Right-skewness in the reported forecast distribution, which is another indicator that pessimistic opinions are missing from the forecast distribution, turns out to be a significant predictor of future returns as well.

These results indicate that the marginal investor, on average, fails to adjust analysts' earnings forecasts for inherent biases. When I check whether the resulting mispricing is ameliorated when sophisticated investors trade against it, I find this to be the case. When institutions, a more sophisticated class of investors with access to independent earnings forecasts, trade against the dispersion in analysts' forecasts dispersion loses its predictive power on future returns. This finding suggests a beneficial effect of sophisticated investors on market efficiency.

## References

Abdel-khalik, A. & Ajinkya, B. (1982), 'Returns to informational advantages: The case of analysts' forecast revisions', *Journal of Accounting Research* **October**, 661–680.

Barber, B. & Odean, T. (1999), 'The courage of misguided convictions: The trading behavior of individual investors', *Financial Analyst Journal* pp. 41–55.

Barber, B. & Odean, T. (2000), 'Trading is hazardous to your wealth: The common stock investment performance of individual investors', *The Journal of Finance* pp. 773–806.

Bradshaw, M. T., Richardson, S. A. & Sloan, R. G. (2003), Pump and dump: An empirical analysis of the relation between corporate financing activities and sell-side analyst research. Working Paper.

Brown, S. J. & Warner, J. B. (1985), 'Using daily stock returns: The case of even studies', *Journal of Financial Economics* **14**, 3–31.

Chan, L., Jegadeesh, N. & Lakonishok, J. (1996), 'Momentum strategies', *The Journal of Finance* **51**, 1681–1713.

Chari, V., Jagannathan, R. & Offer, A. (1988), 'Seasonalities in security returns. the case of earnings announcements', *Journal of Financial Economics* **21**, 101–121.

Ciccone, S. J. (2003), 'Does analyst optimism about future earnings distort stock prices?', *Journal of Behavioral Finance* **4**, 59–64.

Cowen, A., Groysberg, B. & Healy, P. (2003), Which types of analyst firms make more optimistic forecasts? Harvard Business School Working Paper.

Diether, K., Malloy, C. & Scherbina, A. (2002), 'Differences of opinion and the cross-section of stock returns', *The Journal of Finance* **52**(1).

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns of stocks and bonds', *Journal of Financial Economics* **33**, 3 – 56.

Forbes, W. & Skerratt, L. (1992), 'Analysts' forecast revisions and stock price movements', *Journal of Business Finance and Accounting* **June**, 555–564.

Givoly, D. & Lakonishok, J. (1979), 'The information content of financial analysts' earnings forecasts', *Journal of Accounting and Economics* **March**, 165–185.

Gleason, C. & Lee, C. (2000), Characteristics of price informative analyst forecasts. Working Paper.

Gonedes, N., Dopuch, N. & Penman, S. (1976), 'Disclosure rules, information production, and capital market equilibrium: The case of forecast disclosure rules', *Journal of Accounting Research* **Spring**, 89–137.

Gu, Z. & Wu, J. (2003), 'Earning skewness and analyst forecast bias', *Journal of Accounting and Economics* **35**, 5–29.

Hawkins, E., Chamberlin, S. & Daniel, W. (1984), ‘Earnings expectations and security prices’, *Financial Analysts Journal* **September/October**, 24–37.

Hong, H., Lim, T. & Stein, J. (2000), ‘Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies’, *The Journal of Finance* **55**, 265 – 295.

Imhoff, E. & Lobo, G. (1990), ‘Information content of analysts’ composite forecast revisions’, *Journal of Accounting Research* **Autumn**, 541–554.

Johnson, T. C. (2004), ‘Forecast dispersion and the cross-section of expected returns’, *The Journal of Finance* . forthcoming.

Kyle, A. S. (1985), ‘Continuous auctions and insider trading’, **53**, 1315–1335.

McNichols, M. & O’Brien, P. (1997), ‘Self-selection and analyst coverage’, *Journal of Accounting Research* **353**, 167–199.

Miller, E. (1977), ‘Risk, uncertainty, and divergence of opinion’, *The Journal of Finance* **32**(4), 1151–1168.

Odean, T. (1998), ‘Are investors reluctant to realize their losses?’, *The Journal of Finance* pp. 1775–1798.

Richardson, S, S. T. & Wysocki, P. (2001), The walkdown to beatable analyst forecasts: the roles of equity issuance and insider trading incentives. Working paper.

Sadka, R. & Scherbina, A. (2004), Mispricing, liquidity and convergence to fundamentals: Empirical evidence. Working Paper.

Scherbina, A. (2001), Stock prices and differences of opinion: empirical evidence that prices reflect optimism. Working paper.

Stickel, S. (1990), ‘Predicting individual analyst earnings forecasts’, *Journal of Accounting Research* **Autumn**, 409–417.

Viswanathan, S. (2001), Strategic trading, heterogeneous beliefs and short constraints. Working paper.

**Table I**

**Sample Description**

This table reports summary statistics for the variables used in the subsequent two regressions. To be included in the regressions a firm-date observation must be present at the intersection of I/B/E/S, CRSP, and Quarterly Compustat. The time period is 1983-2002. Annual earnings surprise is defined as the realized annual earnings per share minus the mean forecasted number, scaled by the book value of equity as of the end of the previous fiscal year (only fourth-quarter numbers are reported). Cumulative abnormal return around the earnings announcement date is calculated as the mean difference between the return on the stock and the return on the CRSP equally-weighted portfolio from one day before to one day after the earnings announcement day. $\widehat{bias}$ is bias due to self-selection in analyst coverage, estimated under the assumption that if the number of analysts following the stock decreased since the previous quarter, the distribution of the reported forecasts is a truncated normal and the bias is equal to $\frac{\phi(k)}{\Phi(k)} * \sigma$, where $k$ is the point below which analysts will not report their forecast, estimated as $\Phi(k) = \frac{missing}{reporting+missing}$, and $\sigma$ is the standard deviation of the true, not reported, distribution. Dispersion is then defined as $\sigma$; both $\widehat{bias}$ and dispersion are estimated based on the latest available forecasts before the announcement date and are scaled by book equity and the square root of the number of quarters remaining until fiscal year end, in order to facilitate comparison across quarters of a fiscal year. Past-quarter forecast revision in the difference between the latest mean forecast preceding the earnings announcement day and the mean forecast three months ago, scaled by book equity. Skew of the forecast distribution is defined as the difference between the mean and the median forecast scaled by book equity. Market capitalization is the dollar value of all common equity as of the last annual report. Book-to-market ratio is the ratio of book value of equity to the market value of equity as of the last annual report. Book equity is defined as the Compustat book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, I use redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. The observations used in the regressions are required to have at least three outstanding analysts' forecasts and book value of equity of at least $3 per share.

**Descriptive Statistics**

| | Mean | Median | 25th percentile | 75th percentile |
|---|---|---|---|---|
| Average 3-day abnormal return (x$10^{-3}$) (*CAR*) | 0.30 | -0.14 | -9.41 | 9.82 |
| Annual earnings surprise (x$10^{-2}$) (*surprise*) | -1.87 | -0.15 | -1.81 | 0.34 |
| Annual forecast dispersion of the true distribution (x$10^{-2}$) (*dispersion*) | 1.03 | 0.53 | 0.26 | 1.10 |
| Estimated bias due to self-selection (x$10^{-2}$) ($\widehat{bias}$) | 0.05 | 0.00 | 0.00 | 0.00 |
| Past-quarter forecast revision (x$10^{-2}$) (*revision*) | -0.99 | -0.11 | -1.14 | 0.16 |
| Skew of the forecast distribution (x$10^{-3}$) (*skew*) | 0.17 | 0 | -0.69 | 0.82 |
| Market capitalization (in millions of $) (*ME*) | 2,317 | 322 | 102 | 1,156 |
| Book-to-market ratio (*BE/ME*) | 0.75 | 0.64 | 0.40 | 0.96 |

**Pearson Correlation Coefficients**

| | *dispersion* | $\widehat{bias}$ | *revision* | *skew* | $\ln(ME)$ | *BE/ME* |
|---|---|---|---|---|---|---|
| *dispersion* | 1 | 0.41 | -0.23 | 0.05 | -0.04 | 0.05 |
| | | $(< .0001)$ | $(< .0001)$ | $(< .0001)$ | $(< .0001)$ | $(< .0001)$ |
| $\widehat{bias}$ | | 1 | -0.14 | -0.05 | -0.01 | 0.03 |
| | | | $(< .0001)$ | $(< .0001)$ | (0.02) | $(< .0001)$ |
| *revision* | | | 1 | -0.05 | 0.00 | 0.01 |
| | | | | $(< .0001)$ | (0.12) | $(< .0001)$ |
| *skew* | | | | 1 | -0.00 | -0.00 |
| | | | | | (0.29) | (0.05) |
| $\ln(ME)$ | | | | | 1 | -0.11 |
| | | | | | | $(< .0001)$ |

## Table II
## Predictors of the Earnings Surprise

This table is based on analysts' annual earnings per share forecasts immediately preceding the annual earnings announcement dates. The time period is 1983-2002. Earnings surprise is defined as the realized annual earnings per share minus the average forecasted number scaled by the book value of equity. The table reports coefficients of regressing earnings surprise on various predictive variables. These variables are: $\widehat{bias}$ is bias due to self-selection in analyst coverage, estimated under the assumption that if the number of analysts following the stock decreased since the previous quarter, the distribution of the reported forecasts is a truncated normal and the bias is equal to $\frac{\phi(k)}{\Phi(k)} * \sigma$, where $k$ is the point below which analysts will not report their forecast, estimated as $\Phi(k) = \frac{missing}{reporting+missing}$, and $\sigma$ is the standard deviation of the true, not reported, distribution; dispersion is defined as $\sigma$; past-quarter revision, is the change in the mean outstanding annual earnings forecast over the past three months; revision$^+$/revision$^-$ are variables that are equal to revision if it is positive/negative and zero otherwise; skew is the skew of the forecast distribution, defined only if there are more than two forecasts outstanding (dispersion, $\widehat{bias}$, revision and skew are scaled by book equity as of the last annual report); $\ln(ME)$, the logarithm of the firm's market value of equity, and $\ln(BE/ME)$, the logarithm of the ratio of the firm book value of equity to the market value of equity, both as of the latest available annual report. Book equity is defined as the Compustat book value of stockholders' equity plus balance sheet deferred taxes and investment tax credit (if available) minus the book value of preferred stock. Depending on availability, I use redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. The observations used in the regressions are required to have at least three outstanding analysts' forecasts and book value of equity of at least $3 per share. Adjusted $R^2$ is reported in per cent.

| *intercept* (x$10^{-2}$) | $\widehat{bias}$ | *dispersion* | *skew* | past-quarter *rev* | past-quarter *rev*$^-$ | past-quarter *rev*$^+$ | $\ln(ME)$ (x$10^{-3}$) | $BE/ME$ (x$10^{-3}$) | $R^2$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| $-0.56^a$ | $-1.11^a$ | $-0.48^a$ | $-0.43^a$ | $0.18^a$ | - | - | $-0.83^a$ | 0.48 | 4.53 |
| (-3.27) | (-11.42) | (-19.18) | (-8.62) | (15.06) | - | - | (-3.71) | (0.66) | |
| $-0.72^a$ | $-1.01^a$ | $-0.51^a$ | - | - | $0.16^a$ | $0.39^a$ | $-0.71^a$ | 0.76 | 4.39 |
| (-4.22) | (-10.55) | (-19.80) | - | - | (11.52) | (10.18) | (-3.17) | (1.04) | |
| $-1.53^a$ | $-2.07^a$ | - | - | - | - | - | - | - | 1.84 |
| (-42.88) | (-23.80) | - | - | - | - | - | - | - | |
| $-1.15^a$ | - | $-0.67^a$ | - | - | - | - | - | - | 3.03 |
| (-29.65) | - | (-30.75) | - | - | - | - | - | - | |
| $-1.65^a$ | - | - | $-0.34^a$ | - | - | - | - | - | 0.15 |
| (-46.45) | - | - | (-6.70) | - | - | - | - | - | |
| $-1.39^a$ | - | - | - | $0.29^a$ | $0.23^a$ | - | - | - | 1.91 |
| (-35.35) | - | - | - | (22.81) | (5.94) | - | - | - | |
| Past-quarter revision $\leq 0$ (20,068 observations) | | | | | | | | | |
| $-0.64^a$ | $-1.03^a$ | $-0.60^a$ | $-0.44^a$ | $0.09^a$ | - | - | $-1.13^a$ | -0.86 | 5.28 |
| (-2.89) | (-9.56) | (-20.32) | (-7.70) | (6.04) | - | - | (-3.86) | (-0.94) | |
| $-2.02^a$ | $-2.09^a$ | - | - | - | - | - | - | - | 2.27 |
| (-43.51) | (-21.64) | - | - | - | - | - | - | - | |
| $-1.54^a$ | - | $-0.76^a$ | - | - | - | - | - | - | 4.35 |
| (-30.81) | - | (-30.22) | - | - | - | - | - | - | |
| $-2.16^a$ | - | - | $-0.27^a$ | - | - | - | - | - | 0.11 |
| (-46.59) | - | - | (-4.72) | - | - | - | - | - | |
| $-1.77^a$ | - | - | - | $-0.24^a$ | - | - | - | - | 1.48 |
| (-34.25) | - | - | - | (17.38) | - | - | - | - | |
| Past-quarter revision $> 0$ (10,126 observations) | | | | | | | | | |
| $-0.78^a$ | $-1.30^a$ | 0.05 | $-0.45^a$ | 0.05 | - | - | $-0.27^a$ | $0.42^a$ | 0.46 |
| (-3.03) | (-4.64) | (0.92) | (-3.84) | (1.31) | - | - | (-0.81) | (3.81) | |
| $-0.60^a$ | $-1.04^a$ | - | - | - | - | - | - | - | 0.15 |
| (-11.39) | (-4.01) | - | - | - | - | - | - | - | |
| $-0.63^a$ | - | -0.01 | - | - | - | - | - | - | 0.01 |
| (-10.74) | - | (-0.12) | - | - | - | - | - | - | |
| $-0.65^a$ | - | - | $-0.41^a$ | - | - | - | - | - | 0.11 |
| (-46.59) | - | - | (-3.51) | - | - | - | - | - | |
| $-0.66^a$ | - | - | - | 0.03 | - | - | - | - | -0.00 |
| (-11.53) | - | - | - | (0.93) | - | - | - | - | |

**Table III**

**Predictors of Cumulative Abnormal Returns Around Quarterly Earnings Announcement Days**

Cumulative abnormal return is defined as the average stock return net of the equal-weighted market return in a three-day window around the earnings announcement day (in per cent). The table reports coefficients of regressing cumulative returns on various predictive variables; the intercept coefficient is not reported. The predictive variables are: $stdev$, the standard deviation of the latest outstanding analyst forecasts scaled by the book value of equity as of the latest available annual report (previous year's book equity is used for the first quarter of the fiscal year); $\widehat{bias}_2$, the predicted bias, calculated under the assumption that if the number of analysts following the stock decreased since the previous quarter, the analysts no longer following the stock will have forecasted earnings lower by 1 cent than the currently lowest outstanding forecast and scaled by the book value of equity; $rev_3$, the change in the mean outstanding annual earnings forecast over the past three months scaled by book equity as of the last annual report; *turnover*, defined as the average daily volume during the 3-day period around the earnings announcement date scaled by the total number of common shares outstanding as of the last annual report; *volume ratio*, the ratio of the average daily volume within the 3-day period around the earnings announcement date to the average daily volume within the period from 61 days to one day prior the earnings announcement date; $\ln(ME)$, the logarithm of the firm's market value of equity, and $\ln(BE/ME)$, the logarithm of the ratio of the firms' book value of equity to the market value of equity, both as of the latest available annual report. Book equity is defined as the Compustat book value of stockholders' equity plus balance sheet deferred taxes and investment tax credit (if available) minus the book value of preferred stock. Depending on availability, we use redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. The observations used in the regressions are required to have at least three outstanding analysts' forecasts and book value of equity of at least $3 per share. Adjusted $R^2$ is reported in per cent.

Panel A: All observations (142,123 observations)

| *dispersion* $(*10^{-3})$ | $\widehat{bias}_2$ $(*10^{-3})$ | *past-qtr. revision* $(*10^{-3})$ | *turnover* (%) | *volume ratio* $(*10^{-6})$ | $\ln(ME)$ $(*10^{-3})$ | $BE/ME$ $(*10^{-3})$ | $R^2$ |
|---|---|---|---|---|---|---|---|
| -4.31 | -6.76 | $16.88^a$ | $-0.09^a$ | $7.47^a$ | $-0.15^a$ | $-4.99^a$ | 0.49 |
| (-0.97) | (-1.22) | (8.93) | (-24.03) | (4.75) | (-4.01) | (-4.06) | |
| $-14.38^a$ | - | - | - | - | - | - | 0.01 |
| (-3.39) | - | - | - | - | - | - | |
| - | $-16.21^a$ | - | - | - | - | - | 0.01 |
| - | (-3.00) | - | - | - | - | - | |
| - | - | $18.50^a$ | - | - | - | - | 0.07 |
| - | - | (10.12) | - | - | - | - | |
| $-12.31^a$ | $-13.03^b$ | - | - | - | - | - | 0.01 |
| (-2.84) | (-2.36) | - | - | - | - | - | |
| -4.55 | - | $18.04^a$ | - | - | - | - | 0.07 |
| (-1.04) | - | (9.59) | - | - | - | - | |
| - | -8.35 | $18.50^a$ | - | - | - | - | 0.07 |
| - | (-1.53) | (9.78) | - | - | - | - | |
| $-13.41^a$ | - | - | $-0.08^a$ | - | - | - | 0.40 |
| (-3.17) | - | - | (-23.77) | - | - | - | |
| $-14.38^a$ | - | - | - | $-3.62^b$ | - | - | 0.01 |
| (-3.39) | - | - | - | (-2.41) | - | - | |
| - | - | - | $-0.08^a$ | - | - | - | 0.40 |
| - | - | - | (-23.81) | - | - | - | |
| - | - | - | - | $-3.62^b$ | - | - | 0.00 |
| - | - | - | - | (-2.41) | - | - | |

[a,b,c] Statistically significant at the 1, 5, and 10 percent levels, respectively

Panel B: Firms covered by at least 10 analysts (23,223 observations)

| *dispersion* ($*10^{-3}$) | $\widehat{bias}_2$ ($*10^{-3}$) | *past-qtr. revision* ($*10^{-3}$) | *turnover* | *volume ratio* ($*10^{-6}$) | ln($ME$) ($*10^{-3}$) | $BE/ME$ ($*10^{-3}$) | $R^2$ |
|---|---|---|---|---|---|---|---|
| 10.03 | $-82.63^a$ | -3.02 | $-0.03^a$ | $-137.91^a$ | $-0.43^a$ | $-1.74^a$ | 2.04 |
| (0.39 ) | (-2.85) | (-0.60) | (-3.72) | (-18.47) | (-3.48) | (-3.69) | |
| 0.05 | - | - | - | - | - | - | 0.00 |
| (0.27) | - | - | - | - | - | - | |
| - | $-71.98^b$ | - | - | - | - | - | 0.02 |
| - | (-2.50) | - | - | - | - | - | |
| - | - | 0.10 | - | - | - | - | 0.00 |
| - | - | (0.62) | - | - | - | - | |
| - | - | - | $-0.07^a$ | - | - | - | 0.52 |
| - | - | - | (-11.05) | - | - | - | |
| - | - | - | - | $-145.83^a$ | - | - | 1.92 |
| - | - | - | - | (-21.33) | - | - | |

Panel C: 1992-2002 subperiod (89,476 observations)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1.85 | $-15.02^b$ | $14.33^a$ | $-0.09^a$ | $-6.59^a$ | $-0.24^a$ | $-0.81^a$ | 0.69 |
| (0.23) | (-2.06) | (5.54) | (-21.67) | (-3.07) | (-4.59) | (-4.42) | |
| 10.36 | - | - | - | - | - | - | 0.00 |
| (-1.43) | - | - | - | - | - | - | |
| - | $-20.23^a$ | - | - | - | - | - | 0.01 |
| - | (-2.83) | - | - | - | - | - | |
| - | - | $16.06^a$ | - | - | - | - | 0.04 |
| - | - | (6.35) | - | - | - | - | |
| - | - | - | $-0.09^a$ | - | - | - | 0.62 |
| - | - | - | (-23.56) | - | - | - | |
| - | - | - | - | $-19.98^a$ | - | - | 0.10 |
| - | - | - | - | (-9.75) | - | - | |

Panel D: 1983-1991 subperiod (52,649 observations)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -5.88 | 6.39 | $22.26^a$ | $-0.07^a$ | $28.90^a$ | -0.07 | 0.23 | 0.56 |
| (-1.29) | (0.78) | (8.91) | (-0.67) | (13.10) | (-1.32) | (1.57) | |
| $-15.50^a$ | - | - | - | - | - | - | 0.02 |
| (-3.71) | - | - | - | - | - | - | |
| - | -8.36 | - | - | - | - | - | 0.00 |
| - | (-1.08) | - | - | - | - | - | |
| - | - | $22.76^a$ | - | - | - | - | 0.17 |
| - | - | (9.58) | - | - | - | - | |
| - | - | - | $0.05^a$ | - | - | - | 0.05 |
| - | - | - | (4.98) | - | - | - | |
| - | - | - | - | $28.34^a$ | - | - | 0.38 |
| - | - | - | - | (14.26) | - | - | |

[a,b,c] Statistically significant at the 1, 5, and 10 percent levels, respectively

Panel E: Subsamples based on abnormal return

| *dispersion* ($*10^{-3}$) | $\widehat{bias}_2$ ($*10^{-3}$) | *past-qtr. revision* ($*10^{-3}$) | *turnover* | *volume ratio* ($*10^{-6}$) | $\ln(ME)$ ($*10^{-3}$) | $BE/ME$ ($*10^{-3}$) | $R^2$ |
|---|---|---|---|---|---|---|---|
| | | | $\lvert CAR\rvert > 5\%$ (6,071 observations) | | | | |
| 0.29 | -17.48 | $35.73^c$ | $-0.19^a$ | $-57.71^a$ | $-3.05^a$ | -1.87 | 2.77 |
| (0.01) | (-0.28) | (1.74) | (-9.46) | (-4.33) | (-4.50) | (-1.36) | |
| 30.24 | - | - | - | - | - | - | -0.01 |
| (0.64) | - | - | - | - | - | - | |
| - | 18.61 | - | - | - | - | - | -0.01 |
| - | (0.30) | - | - | - | - | - | |
| - | - | 22.40 | - | - | - | - | 0.00 |
| - | - | (1.12) | - | - | - | - | |
| - | - | - | $-0.22^a$ | - | - | - | 2.25 |
| - | - | - | (-11.87) | - | - | - | |
| - | - | - | - | $-92.17^a$ | - | - | 0.84 |
| - | - | - | - | (-7.26) | - | - | |
| | | | $\lvert CAR\rvert \leq 5\%$ (136,052 observations) | | | | |
| $-8.38^b$ | $-10.96^a$ | $16.67^a$ | $0.01^a$ | $14.47^a$ | -0.01 | -0.15 | 0.27 |
| (-2.53) | (-2.66) | (11.85) | (2.85) | (11.72) | (-0.48) | (-1.60) | |
| $-19.19^a$ | - | - | - | - | - | - | 0.03 |
| (-6.10) | - | - | - | - | - | - | |
| - | $-19.76^a$ | - | - | - | - | - | 0.02 |
| - | (-4.93) | - | - | - | - | - | |
| - | - | $18.06^a$ | - | - | - | - | 0.13 |
| - | - | (13.27) | - | - | - | - | |
| - | - | - | $0.02^a$ | - | - | - | 0.02 |
| - | - | - | (5.66) | - | - | - | |
| - | - | - | - | $15.65^a$ | - | - | 0.12 |
| - | - | - | - | (13.07) | - | - | |

[a,b,c] Statistically significant at the 1, 5, and 10 percent levels, respectively

**Table IV**
**Forecast Dispersion and Portfolio Returns**

The left table reports alphas of the Fama-French three-factor model for the time period of July 1983-December 2002. Stocks are sorted first by size and then by dispersion in analyst earnings per share forecasts. Dispersion is defined as the standard deviation of all outstanding analysts' earnings per share forecasts for the current fiscal year, scaled by the book value of equity as of the latest available annual report (previous year's book equity is used for the first quarter of the fiscal year). To facilitate comparison across firms with different months fiscal year end, both estimated bias and dispersion are divided by the square root of the number of months remaining until the announcement month of annual earnings (this method assumes that signals about annual earnings are evenly spread and independent across months). Portfolios are formed every month and stocks in the portfolios held for one month. The observations used are required to have at least three outstanding analysts' forecasts, book value of equity of at least $3 per share, and stock price of at least $5 per share, *t*-statistics are adjusted for autocorrelation.

**Returns**

| Dispersion | small *S1* | *S2* | *S3* | *S4* | large *S5* | All Stocks |
|---|---|---|---|---|---|---|
| *D1 (low)* | $0.46^b$ | $0.28^c$ | $0.28^c$ | 0.11 | 0.14 | $0.23^c$ |
| | (1.99) | (1.70) | (1.90) | (0.71) | (1.26) | (1.73) |
| *D2* | 0.19 | $0.31^b$ | 0.12 | -0.08 | 0.12 | 0.13 |
| | (0.82) | (2.01) | (0.91) | (-0.61) | (1.27) | (1.30) |
| *D3* | -0.02 | 0.00 | -0.10 | -0.10 | 0.01 | -0.02 |
| | (-0.08) | (0.00) | (-0.77) | (-0.86) | (0.18) | (-0.15) |
| *D4* | $-0.42^c$ | -0.04 | -0.12 | -0.04 | -0.12 | -0.07 |
| | (-1.80) | (-0.22) | (-0.92) | (0.32) | (-1.44) | (-0.63) |
| *D5 (high)* | $-1.03^a$ | $-0.36^c$ | $-0.45^a$ | $-0.24^c$ | -0.17 | $-0.53^a$ |
| | (-4.34) | (-1.82) | (-3.05) | (-1.81) | (-1.36) | (-3.79) |
| *D1-D5* | $1.50^a$ | $0.64^a$ | $0.74^a$ | $0.35^c$ | 0.31 | $0.75^a$ |
| | (5.53) | (3.17) | (3.41) | (1.71) | (1.51) | (4.07) |

**Portfolio statistics**

| Dispersion | | small *S1* | *S2* | *S3* | *S4* | large *S5* |
|---|---|---|---|---|---|---|
| *D1* | size ($*10^3$ $) | 97.19 | 253.35 | 556.94 | 1,135.29 | 10,992.04 |
| | dispersion $*10^{-3}$ | 1.67 | 1.39 | 1.30 | 1.36 | 1.26 |
| | coverage | 4.24 | 5.30 | 7.29 | 11.64 | 21.30 |
| *D2* | size ($*10^3$ $) | 94.52 | 251.87 | 556.58 | 1,137.20 | 11,063.93 |
| | dispersion $*10^{-3}$ | 3.87 | 3.05 | 2.68 | 2.56 | 2.30 |
| | coverage | 4.61 | 6.02 | 8.02 | 12.19 | 21.26 |
| *D3* | size ($*10^3$ $) | 90.79 | 250.98 | 553.40 | 1,138.11 | 11,565.71 |
| | dispersion $*10^{-3}$ | 6.64 | 5.02 | 4.35 | 3.97 | 3.52 |
| | coverage | 4.69 | 6.33 | 8.42 | 12.32 | 21.26 |
| *D4* | size ($*10^3$ $) | 86.76 | 250.23 | 548.73 | 1,137.54 | 11,185.55 |
| | dispersion $*10^{-3}$ | 11.71 | 58.55 | 7.32 | 6.38 | 5.52 |
| | coverage | 4.74 | 6.54 | 8.70 | 12.81 | 21.07 |
| *D5* | size ($*10^3$ $) | 79.17 | 245.92 | 544.48 | 1,135.76 | 9.245.01 |
| | dispersion $*10^{-3}$ | 39.35 | 26.36 | 22.09 | 18.50 | 14.24 |
| | coverage | 4.89 | 6.75 | 9.17 | 13.03 | 20.43 |

**Table V**

**Estimated Truncation-Caused Bias and Portfolio Returns**

The left table reports alphas of the Fama-French three-factor model for the time period of July 1983-December 2002. The stocks are sorted first by size and then by estimated bias in current year's analyst earnings forecasts. Estimated bias is calculated under the assumption that if the number of analysts forecasting earnings for the current fiscal year decreased from three months ago, the analysts who dropped coverage would have forecasted a number somewhere below the truncation point, according to the normal distribution. Truncation point, $k$, is estimated as $\Phi(k) = \frac{missing}{reporting+missing}$. The bias is equal to $\frac{\phi(k)}{\Phi(k)} * \sigma$, where $\sigma$ is the standard deviation of the true, not the reported, distribution. This number is then scaled by the firm's book value of equity and, to facilitate comparison across firms with different months' fiscal year end, divided by the square root of the number of months remaining until the announcement month of annual earnings (this method assumes that signals about annual earnings are evenly spread and independent across months). Portfolios are formed every month and held for one month. Forecasts for the second, third, and fourth month of the fiscal year are not considered because it is not possible to calculate drop in coverage in the beginning of the fiscal year. Only observations with book value of equity of at least $3 per share and stock price of at least $5 per share are considered, *t*-statistics are adjusted for autocorrelation.

**Returns**

| Bias | Size: small *S1* | *S2* | *S3* | *S4* | large *S5* | All Stocks |
|---|---|---|---|---|---|---|
| $\widehat{bias} = 0$ | -0.16 | -0.03 | -0.01 | -0.11 | -0.12 | -0.09 |
| | (-0.92) | (-0.24) | (-0.07) | (-0.98) | (-1.46) | (-0.85) |
| $\widehat{bias}$ *low* | -0.07 | -0.11 | 0.09 | 0.09 | 0.10 | -0.01 |
| | (-0.21) | (-0.42) | (0.45) | (0.47) | (0.88) | (-0.01) |
| $\widehat{bias}$ *high* | $-1.10^a$ | $-0.94^a$ | -0.35 | $-0.37^b$ | $-0.25^c$ | $-0.51^a$ |
| | (-3.15) | (-3.89) | (-1.55) | (-2.18) | (-1.85) | (-2.95) |
| *zero-*$\widehat{bias}$*–* | $0.94^a$ | $0.91^a$ | $0.34^c$ | $0.26^c$ | 0.13 | $0.42^a$ |
| *high-*$\widehat{bias}$ *stocks* | (3.31) | (4.20) | (1.91) | (1.83) | (1.28) | (3.66) |

**Portfolio statistics**

| Bias | | Size: small *S1* | *S2* | *S3* | *S4* | large *S5* |
|---|---|---|---|---|---|---|
| $\widehat{bias} = 0$ | size ($*10^3$ $) | 41.79 | 130.24 | 314.75 | 842.09 | 7,636.61 |
| | $\widehat{bias} * 10^{-3}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | *missing* | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| $\widehat{bias}$ *low* | size ($*10^3$ $) | 48.73 | 136.15 | 326.21 | 894.53 | 10,871.44 |
| | $\widehat{bias} * 10^{-3}$ | 2.50 | 1.27 | 0.80 | 0.50 | 0.33 |
| | *missing* | 1.16 | 1.18 | 1.21 | 1.26 | 1.43 |
| $\widehat{bias}$ *high* | size ($*10^3$ $) | 40.45 | 129.33 | 312.04 | 38.50 | 7,214.10 |
| | $\widehat{bias} * 10^{-3}$ | 24.53 | 12.00 | 6.50 | 4.15 | 1.84 |
| | *missing* | 1.33 | 1.46 | 1.56 | 1.71 | 2.12 |

**Table VI**

**Portfolio Returns as a Function of Truncation Bias and Dispersion**

The left table reports alpha's of the Fama-French three-factor model for the time period of July 1983-December 2002. The stocks are sorted first by dispersion and then within each dispersion group by estimated bias in the current year's analyst earnings forecasts. Estimated bias is calculated under the assumption that if the number of analysts forecasting earnings for the current fiscal year decreased from three months ago, the analysts who dropped coverage would have forecasted a number somewhere below the truncation point, according to the normal distribution. Truncation point, $k$ is estimated as $\Phi(k) = \frac{missing}{reporting+missing}$. The bias is equal to $\frac{\phi(k)}{\Phi(k)} * \sigma$, where $\sigma$ is the standard deviation of the true, not the reported, distribution. Dispersion is calculated as the standard deviation in analysts' current year's earnings-per-share forecasts scaled by book value of equity (only observations with at least three outstanding analysts' forecasts are considered). To facilitate comparison across firms with different fiscal year ends, both estimated bias and dispersion are scaled by the square root of the number of months remaining until the announcement month of the annual earnings announcement day (this method assumes that signals about annual earnings are evenly spread and independent across months). Portfolios are formed every month and held for one month. Forecasts for the second, third, and fourth month of the fiscal year are not considered because it is not possible to calculate the drop in coverage at the beginning of the fiscal year. Only observations with the book value of equity of at least \$3 per share and stock price of at least \$5 per share are considered, $t$-statistics are adjusted for autocorrelation.

**Returns**

| Bias | Dispersion: fewer than 3 forecasts | 3 forecasts or more: low | | | | high |
|---|---|---|---|---|---|---|
| | N/A | *D1* | *D2* | *D3* | *D4* | *D5* |
| $\widehat{bias} = 0$ | -0.02 | 0.15 | 0.11 | -0.13 | $-0.20^c$ | $-0.54^a$ |
| | (-0.18) | (0.95) | (0.98) | (-1.00) | (-1.82) | (-3.67) |
| $\widehat{bias} > 0$ | $-0.28^c$ | 0.16 | -0.00 | 0.01 | $-0.42^b$ | $-0.94^a$ |
| | (-1.70) | (0.92) | (-0.24) | (0.09) | (-2.10) | (-4.06) |
| *return* | $0.26^c$ | -0.01 | 0.11 | -0.15 | 0.22 | $0.40^a$ |
| *difference* | (1.73) | (-0.06) | (0.92) | (-1.27) | (1.56) | (2.63) |

**Portfolio statistics**

| Bias | | Dispersion: fewer than 3 forecasts | 3 forecasts or more: low | | | | high |
|---|---|---|---|---|---|---|---|
| | | *N/A* | *D1* | *D2* | *D3* | *D4* | *D5* |
| $\widehat{bias} = 0$ | dispersion $*10^{-3}$ | n/a | 1.21 | 2.83 | 4.60 | 7.83 | 25.30 |
| | $\widehat{bias} * 10^{-3}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | missing | 0.02 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\widehat{bias} > 0$ | dispersion $*10^{-3}$ | n/a | 1.54 | 2.84 | 4.62 | 7.84 | 27.38 |
| | $\widehat{bias} * 10^{-3}$ | 9.41 | 0.21 | 0.41 | 0.69 | 1.21 | 5.11 |
| | missing | 1.23 | 1.50 | 1.49 | 1.54 | 1.56 | 1.61 |

**Table VII**
**Skewness in Analyst Forecasts and Portfolio Returns**

The left table reports alphas of the Fama-French three-factor model for the time period of July 1983-December 2002. The stocks are sorted first by size and then by skewness of outstanding analyst earnings per share forecasts. Skewness is defined as the difference between the mean and median of the outstanding earnings per share forecasts scaled by the book value of equity as of the latest available annual report (previous year's book equity is used for the first quarter of the fiscal year). To facilitate comparison across firms with different months' fiscal year ends, the skewness measure is divided by the square root of the number of months remaining until the announcement month of annual earnings (this method assumes that signals about annual earnings are evenly spread and independent across months). Portfolios are formed every month and stocks in the portfolios held for one month. The observations used are required to have at least three outstanding analysts' forecasts, book value of equity of at least \$3 per share, and stock price of at least \$5 per share; *t*-statistics are adjusted for autocorrelation.

**Returns**

| Dispersion | Size: small *S1* | *S2* | *S3* | *S4* | large *S5* | All Stocks |
|---|---|---|---|---|---|---|
| *Skew1 (low)* | $-0.48^b$ | -0.02 | -0.06 | 0.02 | 0.09 | -0.06 |
| | (-2.31) | (-0.07) | (-0.39) | (0.12) | (1.05) | (-0.45) |
| *Skew2* | 0.12 | $0.27^c$ | 0.07 | -0.08 | 0.02 | 0.04 |
| | (0.47) | (1.69) | (0.51) | (-0.63) | (0.15) | (0.32) |
| *Skew3* | 0.08 | 0.12 | 0.04 | 0.05 | 0.07 | 0.06 |
| | (0.38) | (0.74) | (0.33) | (0.40) | (0.88) | (0.60) |
| *Skew4* | 0.05 | -0.05 | -0.02 | -0.18 | -0.02 | -0.05 |
| | (0.20) | (-0.29) | (-0.17) | (-1.30) | (-0.22) | (-0.36) |
| *Skew5 (high)* | $-0.55^b$ | -0.18 | $-0.33^b$ | $-0.17^c$ | $-0.21^b$ | $-0.29^b$ |
| | (-2.36) | (-1.26) | (-2.48) | (-1.80) | (-2.26) | (-2.54) |
| *Skew1-Skew5* | 0.08 | 0.16 | $0.28^b$ | 0.19 | $0.30^b$ | $0.23^a$ |
| | (0.46) | (0.98) | (2.02) | (1.54) | (2.54) | (3.24) |

**Portfolio statistics**

| Skewness | | Size: small *S1* | *S2* | *S3* | *S4* | large *S5* |
|---|---|---|---|---|---|---|
| *Skew1* | skewness $*10^{-3}$ | -9.44 | -6.34 | -5.09 | -4.10 | -3.03 |
| | dispersion $*10^{-3}$ | 21.35 | 14.57 | 12.19 | 10.46 | 8.40 |
| | coverage | 4.53 | 6.12 | 8.21 | 12.00 | 20.16 |
| *Skew2* | skewness $*10^{-3}$ | -1.52 | -1.13 | -0.94 | -0.78 | -0.62 |
| | dispersion $*10^{-3}$ | 5.80 | 4.45 | 3.86 | 3.59 | 3.25 |
| | coverage | 4.64 | 6.06 | 8.13 | 12.55 | 21.32 |
| *Skew3* | skewness $*10^{-3}$ | 0.95 | 0.05 | 0.03 | 0.03 | 0.02 |
| | dispersion $*10^{-3}$ | 5.11 | 4.10 | 3.74 | 3.59 | 3.23 |
| | coverage | 4.60 | 6.10 | 8.26 | 12.58 | 21.73 |
| *Skew4* | skewness $*10^{-3}$ | 1.92 | 1.34 | 1.07 | 0.88 | 0.66 |
| | dispersion $*10^{-3}$ | 6.85 | 5.02 | 4.20 | 3.94 | 3.41 |
| | coverage | 4.74 | 6.19 | 8.31 | 12.37 | 21.38 |
| *Skew5* | skewness $*10^{-3}$ | 10.80 | 7.27 | 5.82 | 4.51 | 3.14 |
| | dispersion $*10^{-3}$ | 24.41 | 16.33 | 13.80 | 11.18 | 8.55 |
| | coverage | 4.68 | 6.40 | 8.60 | 12.39 | 20.44 |

**Table VIII**

**Institutional Trades and Portfolio Returns**

The left table reports alphas of the Fama-French three-factor model for the time period of July 1983-December 2002. The stocks are sorted independently based on sensitivity of intitutional ownership to dispersion in analysts' forecasts and dispersion. Dispersion is calculated as the standard deviation in all outstanding analyst forecasts scaled by the book value of equity and the square root of the number of months remaining to the annual earnings announcement (the last scaling is done to facilitate cross-sectional comparisons across stocks with different months' fiscal year end and assumes that an equal amount of uncertainty about annual earnings is resolved every month). Institutional ownership is calculated by summing the quarterly holdings of all SEC 13F report filers. Sensitivity of institutional ownership of stock $i$ to dispersion is defined as the coefficient $\beta_{1i}$ of the time series regression: $Instit_{it} = \beta_{0i} + \beta_{1i}\sigma_{it} + \beta_{2i}Ret3_{it} + \varepsilon_{it}$, where $\sigma_{it}$ is the standard deviation in analysts' earnings per share forecasts available before the 13F report, scaled by the book value of equity and $Ret3_{it}$ is the average stock return for the past three months. At least 5 quarterly observations are required to be included in the sample. Given $\beta_{1i}$ statistically significant with a probability of 5%, if it has a negative sign it is classified as a "sell," if a positive sign, a "buy," and if it is insignificant at the 5% level it is classified as a "hold." Stocks are sorted independently into five groups based on dispersion in analysts' earnings per share forecasts. There are 15 portfolios total. Portfolio returns are equally-weighted. Stocks are reassigned into dispersion portfolios every month, Stocks with share prices of less than $5 and book equity values of less than $3 per share are excluded from the portfolios; *t*-statistics in parentheses are adjusted for autocorrelation.

**Returns**

| | Institutional response to analyst disagreement | | |
|---|---|---|---|
| Dispersion | Sell | Hold | Buy |
| *D1 (low)* | 0.02 | 0.29[b] | 0.23 |
| | (0.74) | (2.10) | (1.22) |
| *D2* | 0.16 | 0.16 | 0.16 |
| | (1.16) | (1.42) | (0.80) |
| *D3* | -0.08 | -0.01 | 0.01 |
| | (-0.57) | (-0.05) | (0.03) |
| *D4* | -0.00 | 0.02 | -0.21 |
| | (-0.01) | (0.13) | (-1.18) |
| *D5 (high)* | -0.15 | −0.31[b] | −0.81[a] |
| | (0.87) | (-2.58) | (-4.30) |
| *D1-D5* | 0.17 | 0.59[a] | 1.01[a] |
| | (0.74) | (3.11) | (4.07) |

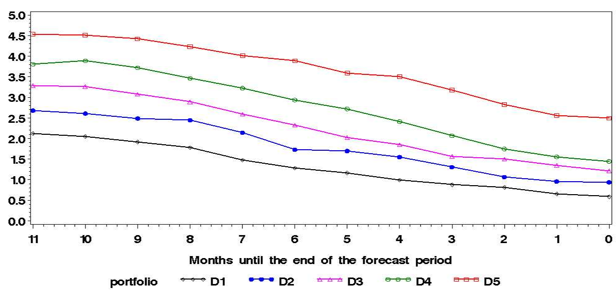
**Portfolio statistics**

| | | Institutional response to analyst disagreement | | |
|---|---|---|---|---|
| Dispersion | | Sell | Hold | Buy |
| *D1* | dispersion $*10^{-3}$ | 0.90 | 0.89 | 0.88 |
| | size ($*10^3$ $) | 2,742.41 | 3,649.52 | 5,865.07 |
| | $\beta_1$ | -23.38 | -1.38 | 19.91 |
| | t-statistic of $\beta_1$ | (-3.32) | (-0.20) | (2.88) |
| | number of firms | 88 | 116 | 20 |
| *D2* | dispersion $*10^{-3}$ | 1.81 | 1.81 | 1.81 |
| | size ($*10^3$ $) | 3,031.81 | 3,044.94 | 5,990.80 |
| | $\beta_1$ | -19.66 | -1.54 | 16.74 |
| | t-statistic of $\beta_1$ | (-3.35) | (-0.23) | (2.90) |
| | number of firms | 92 | 114 | 18 |
| *D3* | dispersion $*10^{-3}$ | 2.88 | 2.88 | 2.88 |
| | size ($*10^3$ $) | 2,840.57 | 2,952.14 | 5,371.12 |
| | $\beta_1$ | -16.61 | -1.31 | 15.03 |
| | t-statistic of $\beta_1$ | (-3.36) | (-0.24) | (2.94) |
| | number of firms | 93 | 114 | 17 |
| *D4* | dispersion $*10^{-3}$ | 4.71 | 4.73 | 4.74 |
| | size ($*10^3$ $) | 2,743.84 | 2,594.93 | 4,976.89 |
| | $\beta_1$ | -13.97 | -1.11 | 11.77 |
| | t-statistic of $\beta_1$ | (-3.33) | (-0.23) | (2.86) |
| | number of firms | 92 | 113 | 19 |
| *D5* | dispersion $*10^{-3}$ | 13.58 | 14.00 | 14.22 |
| | size ($*10^3$ $) | 1,702.08 | 1,445.89 | 1,849.86 |
| | $\beta_1$ | -9.76 | -0.88 | 8.51 |
| | t-statistic of $\beta_1$ | (-3.42) | (-0.26) | (2.91) |
| | number of firms | 87 | 115 | 22 |

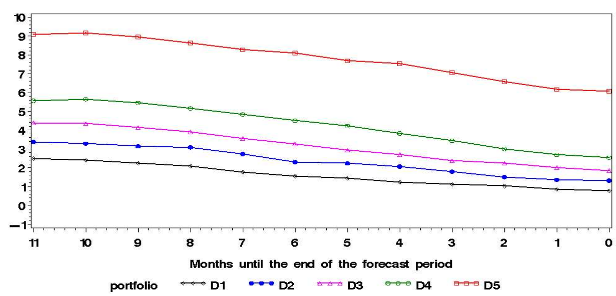
**Figure 1.** In February of each year stocks with the December fiscal year end are sorted into five portfolios based on dispersion in analysts' earnings per share forecasts for the current fiscal year. Forecast error is defined as the mean reported forecast minus the realized earnings, scaled by the book value of equity as of the end of the previous fiscal year. Forecast error is averaged over portfolios by the number of months remaining until the end of the current fiscal year.

## Portfolios are formed every month

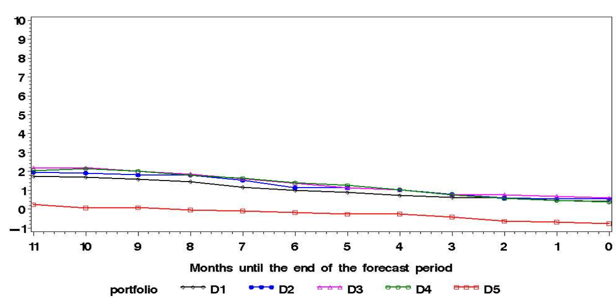
Bias in the mean outstanding forecast



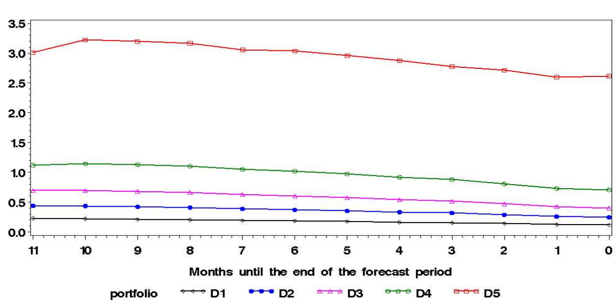

Bias in the highest outstanding forecast




Bias in the lowest outstanding forecast

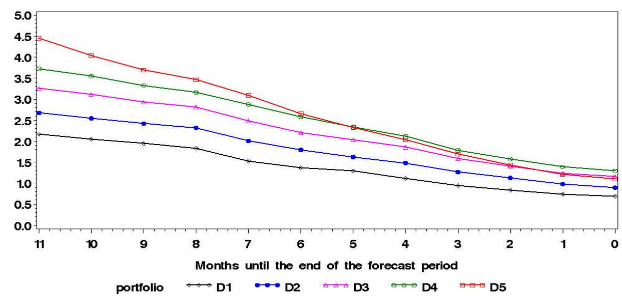



Forecast dispersion

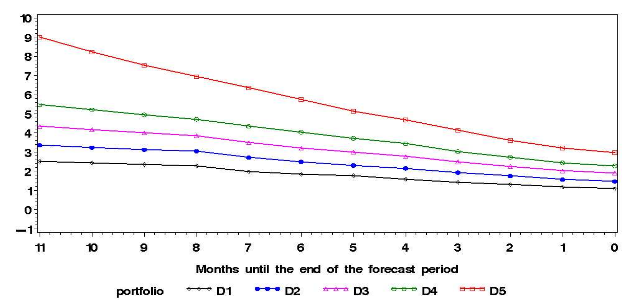



## Portfolios are formed in February of each year

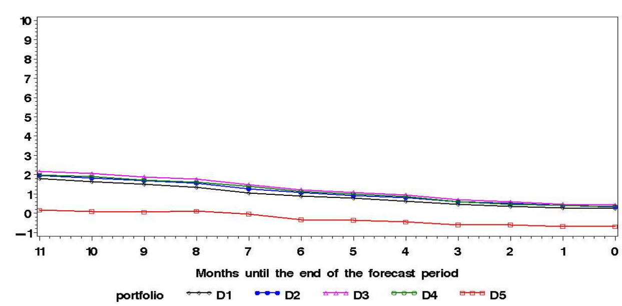
Bias in the mean outstanding forecast




Bias in the highest outstanding forecast




Bias in the lowest outstanding forecast




Forecast dispersion