Investors' Use of Historical Forecast Bias to Adjust Current Expectations

Seung-Woog (Austin) Kwag*
Utah State University
Logan, UT 84322-3565
865-797-2361 voice
435-797-2701 fax
austin.kwag@usu.edu

Ronald E. Shrieves
The University of Tennessee
Knoxville, TN 37996-0540
865-974-3216 voice
865-974-1716 fax
rshrieve@utk.edu

We explore the extent to which investor response to earnings information differs in the presence of historical bias in earnings forecasts. Overall, the results are consistent with the notion that investors take historical forecast bias into account when interpreting information in earnings announcements and that the market's reaction to forecast errors is larger (less negative) when forecasts are historically more optimistic and suggests that the functional form commonly used in the earnings response literature does not appropriately capture the effect of real unexpected earnings information (i.e., investors' expectation errors as opposed to analysts' forecast errors) on stock returns.

Keywords: Earnings forecasts, investor expectation, historical bias, market reaction

JEL Classifications: G12, G14

We are indebted to Arnie Cowan (editor), and to an anonymous referee for their many useful suggestions.

Investors' Use of Historical Forecast Bias to Adjust Current Expectations

1. Introduction

This paper investigates whether investors factor the long-term error history of analysts' forecasts into their reaction to earnings announcements. We make no attempt to isolate the source of the forecast bias, which may be attributable to analysts' behavior, firm behavior, or a combination of these two information sources. Assessing how investors respond to historical bias requires a procedure for ex ante estimation of the bias in a given forecast, based on the information accessible to an investor. We estimate historical bias in forecasts based on our own research, Kwag and Shrieves (2006), and that of Lopez and Rees (2002). Kwag and Shrieves use historical forecast error data to classify a current forecast into one of five rank-ordered categories in an optimism-pessimism continuum. They find that a classification of forecast errors, based on historical accuracy, provides earnings prediction information incremental to that provided by the most recent forecast error. Further, Kwag and Shrieves find that the mean three-day announcement-period abnormal returns are larger for firms classified as having more optimistic forecast biases. They also report a positive (negative) post-announcement drift for relatively pessimistic (optimistic) forecasts over the sixty-day period following an announcement.

Lopez and Rees (2002) evaluate investors' potential ability to incorporate expectations about historical forecast bias into the analysis of earnings-announcement-period market reactions to forecast errors. They investigate how stock prices respond to forecast errors, and address, among other things, whether the market's reaction to a forecast error is conditioned on "the firm's history of consistently beating or missing analysts' forecasts," and conclude that (p. 181) "while the market adjusts for systematic forecast errors, it is not a full adjustment." Lopez and Rees emphasize using the earnings response coefficient (ERC) to evaluate the impact of systematic forecast errors. Though our conclusions are consistent with those of Lopez and Rees, this study uses a much longer history to assess the extent of bias in forecasts, and it provides a different method for examining market response.

We extend our earlier work (Kwag and Shrieves, 2006) by conducting more sophisticated multiple regression tests. These tests analyze the influence of historical bias on announcement-period responses to forecast errors, while conditioning on the lagged forecast error, whether earnings miss or beat the consensus forecast, firm size, and book-to-market ratios. We also perform several robustness checks to ensure the reliability of our results. We investigate whether investors react differently to a given forecast error depending on their expectations about forecast bias, and whether the adjustment is consistent with the application of an earnings expectation discount for optimistic forecasts relative to pessimistic forecasts. The study's empirical results are somewhat mixed. Nevertheless, given the historical evidence of bias in earnings forecasts, the results reinforce prior evidence that analysts' forecasts are a biased measure of investors' earnings expectations (Lopez and Rees, 2002; Kwag and Shrieves, 2006). The results further suggest that the functional form commonly used in the earnings response literature does not appropriately capture the effect real unexpected earnings information has on stock returns.

Early research on earnings expectations focuses on earnings prediction. The studies conclude that analysts' forecasts provide the best proxy for investors' earnings expectations, and that they generally outperform time-series models (Brown and Rozeff, 1978; Fried and Givoly, 1982; Givoly and Lakonishok, 1984; Conroy and Harris, 1987; Brown, Hagerman, Griffin, and Zmijewski, 1987; O'Brien, 1988; Kross, Ro, and Schroeder, 1990). In the spirit of the earlier research conducted by Kwag and Shrieves (2006) and Lopez and Rees (2002), this study contributes to the stream of research that investigates the relation between error patterns in earnings forecasts and investors' reactions to forecast errors (Abarbanell and Bernard, 1992; Dechow and Sloan, 1997; Ackert and Athanassakos, 1997; Shane and Brous, 2001; Bartov, Givoly, and Hayn, 2002). Our findings are broadly consistent with the idea that investors take historical forecast bias into account when interpreting earnings announcement information. This issue is central to the validity of model specifications in previous earnings response studies.

2. Hypotheses, empirical model, and bias determination

The voluminous literature that investigates announcement-period stock price reactions to earnings forecast errors or forecast revisions motivates this research. These studies typically model investors' reactions to earnings information as a function of the unexpected component of announced earnings, as measured by the difference between reported earnings and expected earnings. A general functional form of this relation is:

$$CAR_{it} = g(EE_{it}, X_{it}) = \alpha + \beta \ EE_{it} + \gamma \ X_{it} + \varepsilon_{it} \tag{1}$$

where CAR_{it} is the cumulative abnormal return for an event period around the earnings announcement date for firm i for quarter t; A_{it} is the actual earnings at the quarterly earnings announcement date for quarter t for firm i, scaled by the stock price of firm i prior to the announcement period; E_{it} is the investors' true earnings expectation for firm i in quarter t, scaled by the stock price of firm i prior to the announcement period; EE_{it} represents investors' true expectation errors ($= A_{it} - E_{it}$); and X_{it} is a vector of conditioning variables.

Many researchers use analysts' consensus earnings forecasts as a proxy for investors' earnings expectations. Because analysts' forecasts may not be the same as investors' true earnings expectations, researchers using functional forms similar to equation (1) could make invalid inferences about the relation between unexpected earnings information and investors' reactions to it. Most researchers acknowledge the issue, and add the conditioning variables (X_{it}) to equation (1) to account for potential misspecification. Examples of conditioning variables include: the year-on-year difference in actual fourth quarter earnings, lagged forecast errors, and forecast revisions (Hughes and Ricks, 1987; Cornell and Landsman, 1989; Mendenhall, 1991; Bartov, Givoly, and Hayn, 2002).

A number of recent studies find that the intercept (α) or the coefficient of the earnings expectation error variable (β) depend on whether the expectation error is positive or negative, i.e., whether the firm "beats" or "misses" analysts' forecasts (Kasznik and McNichols, 2002; Bartov, Givoly, and Hayn, 2002; Lopez and Rees, 2002). Collectively, their results suggest that one or both coefficients are higher for firms that meet or beat expectations. In this study the empirical specifications of the relation between returns and earnings expectation errors reflect this phenomenon.

Let F_{it} represent the most recent consensus analysts' forecast associated with A_{it}, scaled by stock price prior to the announcement date, and FE_{it} be the scaled earnings forecast error, for quarter t for firm i, based on the consensus forecast (i.e., $FE_{it} = A_{it} - F_{it}$). Then, if investors believe that analysts' forecasts are biased, the true earnings expectation error differs from the forecast error by an amount ϕ_{it},

$$E_{it} = F_{it} + \phi_{it} \text{ and } EE_{it} = A_{it} - (F_{it} + \phi_{it}) = FE_{it} - \phi_{it}. \tag{2}$$

Since the true (unobservable) measure of expected earnings is E_{it}, then $\phi_{it} = E_{it} - F_{it}$, so $\phi_{it} < 0 \; (> 0)$ for optimistic (pessimistic) forecasts. As a result, when investors take analysts' biased forecasts at face value, FE_{it} should be different from EE_{it}.

Substituting for EE_{it} in (1), and expanding the regression to allow for differences in intercepts and ERCs based upon whether earnings missed (FE<0) or beat (FE>0) the consensus forecast gives:

$$CAR_{it} = \alpha^{-} NEGDUM_{it} + \alpha^{+} POSDUM_{it} + \beta^{-}(FE_{it} * NEGDUM_{it}) + \beta^{+}(FE_{it} * POSDUM_{it})$$

$$- \beta^{-}(\phi_{it} * NEGDUM_{it}) - \beta^{+}(\phi_{it} * POSDUM_{it}) + \gamma X_{it} + \varepsilon_{it} \tag{3}$$

where $NEGDUM_{it}$ ($POSDUM_{it}$) is coded as one if $FE_{it}<0$ ($FE_{it}>0$) and zero otherwise.

We examine the hypothesis that investors consider historical forecast bias and that the market's response to earnings announcements is dependent on the long-term forecast history of the announcing firm, implying that ϕ_{it} in (3) is different from zero. In other words, investors who are aware of analysts' historical optimism expect earnings to be less than the forecast ($\phi_{it} < 0$). They respond to forecast optimism by applying a negative earnings adjustment to the analysts' consensus forecast. Similarly, for forecast pessimism ($\phi_{it} > 0$), investors apply a positive earnings adjustment to the analysts' consensus forecast. Under this hypothesis, with perceived historical optimism, and for a given magnitude of analysts' forecast error, $EE_{it} < FE_{it}$, the announcement-period market response is larger than it would be absent the perceived bias. Conversely, the market's response to historically pessimistic analysts' forecasts involves a smaller price response than would occur if the investors took the analysts' forecast at face value, since $EE_{it} > FE_{it}$. Equation (3) captures this effect via the $-\beta^{-}(\phi_{it} * NEGDUM_{it})$ and $-\beta^{+}(\phi_{it} * POSDUM_{it})$ terms, which are positive for forecasts that investors perceive to have an optimistic bias ($\phi_{it} < 0$), and negative for forecasts that investors perceive to have a pessimistic bias ($\phi_{it} > 0$). We refer to this hypothesis as the bias adjustment hypothesis.

We test the bias adjustment hypothesis in two stages: first, we find an instrument for ϕ'_{it}, the perceived forecast optimism or pessimism, and second, we investigate differences in the stock market's responses to earnings announcements for forecasts that are classified as optimistic and those that are considered pessimistic. The test must take into consideration the joint hypotheses that the instrument for historical bias is reasonable and that investors take the bias into account when they react to earnings announcements. Evidence consistent with our hypothesis suggests that investors may rationally evaluate systematic bias in analysts' forecasting behavior when they form their own earnings expectations (i.e., that awareness of the systematic bias significantly influences investors' expectations). If investors take forecast bias into account, use of analysts' forecasts as a proxy for investor expectations involves misspecification of the presence of varying degrees of perceived optimism and pessimism in earnings forecasts. Under these circumstances, parameter estimates are biased and inefficient. We estimate equation (3) to investigate whether bias, as revealed in the history of forecast errors, influences the market's response to earnings announcements.

3. Data

The study's sample includes data on quarterly individual analysts' earnings forecasts and actual earnings per share (EPS) from 1990 to 2001, as well as price and stock return data. We construct quarterly analysts' consensus earnings forecasts by extracting individual forecasts from the Thomson IBES detail file. They include only forecasts that fall within the 60-day window prior to the quarterly earnings announcement date. For example, if a company schedules its earnings announcement on May 15, we use an interval from March 15 to May 14 as the window within which we collect individual forecasts.[1] If an analyst issues two or more forecasts during

[1] Since several prior studies identify an errors-in-variables problem associated with a regime shift in the way IBES reported actual earnings numbers during the early 1990s, we perform a robustness test of our main results using data over the relatively error-free period from 1996 to 2001 (Abarbanell and Lehavy, 2002; Abarbanell and Lehavy, 2003; Abarbanell and Lehavy, 2007; Cohen, Hann, and Ogneva, 2007).

the 60-day period, we use the latest estimate. Median analysts' forecast errors are calculated using reported earnings from IBES' Actual file. We extract stock price and return information from the Center for Research in Security Prices (CRSP) daily database and collect quarterly Compustat data for market and book values of common equity (*MVE* and *BVE*). Firms in the final sample must be listed on IBES for at least 21 consecutive quarters (due to the bias classification process discussed in the following section), and on CRSP for at least 250 days prior to the earnings announcement.

We calculate market-model adjusted daily abnormal returns in the usual manner, using the least squares estimates of firm *i*'s market model parameters calculated over 240 days to 11 days prior to the earnings announcement. We estimate the market's response to an announcement (CAR_{it}) as the cumulative three-day (–2, 0) earnings announcement-period for abnormal stock returns.

4. Ex ante classification of earnings forecasts according to bias

Since actual investor expectations are unobservable, we use the method of Kwag and Shrieves (2006) to characterize a consensus forecast ex ante objectively according to the degree of analysts' optimism or pessimism. We classify observations based on historical accuracy, i.e., the expected directions and magnitude of the forecast error, and the likelihood that it will be negative. Kwag and Shrieves call the classification system the Mean-Frequency Forecast Error (MFFE) method.[2] The method classifies forecasts according to historical bias (i.e., optimism and pessimism). The results are predictive with respect to subsequent forecast errors.

Using the MFFE classification method, for each firm-quarter median forecast error, we

[2] Following common practice, we truncate forecast errors at the one and 99 percentile levels to exclude the effects of influential extreme observations. For a detailed discussion of the category formation procedure, refer to Kwag and Shrieves (2006).

calculate the mean of the 20 previous quarterly forecast errors (*MQFE*) and a weighted average of the number of negative forecast errors (*WNNFE*) to construct portfolios based on the likely direction and magnitude of bias in the current forecast error. The logic behind using the dual classification method is that combining the two measures (*MQFE* and *WNNFE*) forms subsets of forecasts that are less likely to involve classification error relative to subsets formed using either *MQFE* or *WNNFE* alone.[3]

We form portfolios by ranking the observations into quintiles on the basis of the historical *MQFE* and *WNNFE* values. The two-way rankings result in the 25 groups of firm-quarter observations shown in Table 1. The entries in parentheses in Table 1 indicate the number of quarterly analysts' forecasts in each group. We apply dominance criteria to the 25 groups to align the biases into five categories along a continuum from the most optimistic (category 1, or P1) to the most pessimistic (category 5, or P5). The rationale for using category formation is that observations in each category should dominate observations in the next higher numbered category on at least one of the two metrics, and be at least equivalent on the other.

Observations falling in the first or second quintile with both MQFE and WNNFE, with at least one of the two measures in the first quintile, are classified in the "optimistic" category, P1. Observations falling in the second or third quintile with both MQFE and WNNFE, but in the second quintile on at least one of the two measures are in P2. Observations falling in the third quintile on both MQFE and WNNFE are in P3. Observations falling in the third or fourth quintile with both measures, but in the fourth quintile on at least one measure, are in P4. Finally, we classify observations falling in either the fourth or fifth quintile with MQFE and WNNFE,

[3] The weights decline geometrically from the most recent to the most distant prior errors. The argument for using frequency, in addition to average forecast error, is illustrated in the following hypothetical example from Kwag and Shrieves (2006, p.83): "… suppose that a company at quarter *t* has 1 large negative forecast error at t-20 followed by 19 small positive forecast errors but that the magnitude of the 1 negative forecast error outweighs the sum of 19 small positive forecast errors. In this case, the company for that quarter would be assigned to the optimistic portfolio if the *MQFE* were the only formation method, …"

but in the fifth quintile on at least one measure, as P5, the "pessimistic" category. By design, mean category MQFE is increasing (going from negative to positive) and WNNFE is decreasing as the category index number increases.

<Table 1 here>

The intersection of the sample of firms from IBES and CRSP contains 34,810 firm-quarters. Application of the MFFE method yields 26,335 observations dispersed over categories P1 to P5. Applying the requirement for data from the quarterly Compustat file and excluding sample observations with extreme values eliminate an additional 6,590 observations, leaving 19,745 usable observations with complete data.[4]

Table 2 summarizes descriptive statistics for the entire sample and across the five bias classification categories. For the full sample, the mean CAR is significantly positive while the mean FE is significantly negative. The fact that the abnormal three-day return is positive is not surprising in light of prior research that supports the “uncertainty resolution hypothesis.” This hypothesis states that a firm’s return variances, betas, and returns will, on average, be higher for firms announcing earnings results (Chari, Jagannathan, and Ofer, 1988; Ball and Kothari, 1991). These results reflect the impact of impending news on the risk and return to investors, and imply that earnings announcement period returns for representative samples of earnings announcements tend to be positive, irrespective of time trends in forecast bias. Unreported results investigate these data for time trends in average MQFE and average FE. Both exhibit significant increasing patterns (i.e., they become less negative) over time (Kendall's taus are 0.87 and 0.73 respectively and significant at the 5% level). We do not find a significant trend for average CAR (Kendall’s

[4] The final sample differs slightly from that reported in Kwag and Shrieves (2006) for the following reasons: they use beta-adjusted CARs, while this study uses underlying data to estimate the market-model adjusted returns, thus requiring additional CRSP data; removal of the top and bottom one percentiles on the CARs in addition to the removal of the top and bottom one percentiles on the *MQFE* or *FE*.

tau is 0.2, not significant at the conventional significance levels). The juxtaposition of an increasing trend in average forecast errors (i.e., the mean and median errors became less negative), and the absence of a trend in CARs is consistent with the uncertainty resolution hypothesis. Alternatively, the positive mean announcement-period return may well result from survivorship bias induced by sample selection procedures, since firms with the most adverse earnings performance are more likely to have disappeared over time.[5] Therefore, we turn our attention to whether the returns are systematically different across the bias categories P1 to P5.

<Table 2 here>

We calculate the mean and median forecast error for the quarter subsequent to the computation of each MFFE category. As shown in Table 2, the MFFE classification method produces categories that have mean forecast errors ranging from -0.00108 (significant at the 0.01 level for both the t-test and non-parametric binomial test) for P1 to +0.00034 (significant at the 0.01 level for both t-test and non-parametric binomial test) for P5. These mean (median) forecast errors are consistent with the presence of optimism in earnings forecasts for categories P1and P2, pessimism in forecasts in categories P4 and P5, and neither optimism nor pessimism in P3.[6] Table 2 reports that the categories possess distinct characteristics in terms of firm size (equity market capitalization) and book-to-market ratio. The average CAR and average book-to-market ratio tend to be larger for more optimistic categories; the average firm sizes for P1 and P2 are smaller than those for P3, P4, and P5. These results suggest that CAR may be negatively associated with logSize and positively associated with the book-to-market ratio. The former

[5] Any bias classification procedure employing a long series of earnings and forecasts, in addition to requiring data from IBES, CRSP, and Compustat may introduce survivorship bias (Bernard and Thomas, 1989; Ball and Kothari, 1989).

[6] Kwag and Shrieves (2006) perform extensive parametric and nonparametric tests of the reliability of the classifications from the MFFE algorithm. They examine each of the categories to see whether it exhibits clustering in time, in industry, or even stock exchange listing, concluding that these factors do not determine the “bias persistence” features of the categories.

relation indicates the size effect, while the latter implies profitability of value-firm investments (Fama and French, 1993, 1995). Although a nonparametric correlation analysis (Kendall's tau) does not indicate significant size and book-to-market effects, we nevertheless take the precaution of controlling for those effects in the regressions.

In summary, the bias classification method rank-orders consensus forecasts into categories based upon historical forecast errors. Kwag and Shrieves (2006) find that their rank-ordering method has the ability to provide significant predictive power with respect to subsequent forecast errors.

5. Regression tests of the bias adjustment hypothesis

This section outlines the method for using the ex ante rankings to test whether the market response to earnings announcements reflects the resulting bias classifications.[7] We use the bias category indicator variable, Pn_{it}, as an instrument for the unobservable measure of analysts' forecast bias, ϕ_{it}. The regression specification (3) controls for the contemporaneous forecast error, FE_{it}. Since many studies conclude that investors' response to earnings announcements is influenced by the serial correlation in forecast errors (e.g., Abarbanell and Bernard, 1992; Shane and Brous, 2001), we include the lagged forecast error to ensure that the estimated impact of the category classifications is incremental to any influence captured by the most recent forecast error. We also estimate separate intercept and earnings response coefficients based on whether current earnings miss (FE<0) or beat (FE>0) the most recent consensus forecast. For these regressions, we also eliminate observations with zero FE, reducing the full sample to 19,503. We consider the

[7] This approach for evaluating bias contrasts sharply with that of Lopez and Rees (2002), who use five quarters of forecasts and earnings data (including the current period), classify forecasts as "consistently" beating (missing) analysts' forecasts if all five forecast errors are positive (negative), and use the mean forecast error over the five quarters as the "systematic" component of forecast error.

impact of systematic forecast errors on both the equation intercept (α), as well as the earnings response coefficient (β), and we also test for the combined impact of the effects of historical bias on both parameters. Finally, we condition the market response on the logarithm of firm capitalization and the firm's book-to-market ratio (Fama and French, 1992, 1993, 1995). The following model summarizes our estimation:

$$\begin{aligned} CAR_{it} = {} & \sum_{n=1}^{5}\alpha_n^-(Pn_{it} * NEGDUM_{it}) + \sum_{n=1}^{5}\alpha_n^+(Pn_{it} * POSDUM_{it}) + \\ & \sum_{n=1}^{5}\beta_n^-(FE_{it} * Pn_{it} * NEGDUM_{it}) + \sum_{n=1}^{5}\beta_n^+(FE_{it} * Pn_{it} * POSDUM_{it}) + \\ & \gamma_1 FE_{i,t-1} + \gamma_2 diffSize_{it} + \gamma_3 diffBtoM_{it} + u_{it} \end{aligned} \tag{4}$$

where CAR_{it} is the three-day (–2, 0) earnings announcement-period abnormal stock returns; Pn_{it} is a binary variable that equals one if an observation belongs to category n, n = 1, …, 5; FE_{it} is analysts' earnings forecast error for firm i in quarter t, scaled by stock price ten days prior to the announcement date; $diffSize_{it}$ is the difference between $\log(P_{it} \times Shr_{it})$ and the grand mean of $\log(P_{it} \times Shr_{it})$ in quarter t (but prior to the earnings announcement date), where P_{it} is the closing stock price at the third month of quarter t and Shr_{it} is the number of common shares used to calculate earnings per share (EPS) in quarter t; $diffBtoM_{it}$ is the difference between $BtoM_{it}$ and the grand mean of $BtoM_{it}$, where $BtoM_{it}$ is the ratio of book value of equity to market value of equity in quarter t (but prior to the earnings announcement date); $NEGDUM_{it}$ is an binary variable coded one if $FE_{it}<0$ and zero otherwise; $POSDUM_{it}$ is an binary variable coded one if $FE_{it}>0$ and zero otherwise; and u_{it} is an identically and independently distributed random error term.

Since the size and book-to-market ratio effects are measured in terms of differences from a grand mean, the category intercept and slope effects indicate average fixed and marginal market impacts for each category for a firm of average equity capitalization and average book-to-market ratio. This allows direct comparison of one category's intercept coefficients with those of another for inference regarding the impact of historical forecast bias.

Related literature suggests there are three reasons to restrict the analysis by eliminating the tails of the distribution of forecast errors. First, Freeman and Tse (1992) find that stock prices respond to earnings forecast errors in a nonlinear manner. A second rationale for restricting the range of FEs over which we test the bias adjustment hypothesis is to mitigate the potential confounding effect of changes in the manner in which IBES computes its "actual" earnings data. Abarbanell and Lehavy (2002 and 2007), among others, detect a "regime shift" in the early 1990s when IBES began adjusting their version of actual earnings to remove items not included in analysts' forecasts. Their analysis suggests that large IBES forecast errors, especially large negative FEs, may be associated with large discrepancies between GAAP and IBES-adjusted earnings metrics prior to this regime shift. Disregarding the FE distribution's tails will therefore mitigate the impact of this inconsistency in the definition of earnings on the results. Additionally, this restriction could eliminate many observations where the bias adjustment, ϕ_{it}, is relatively small compared to the magnitude of the forecast error. Hence, we potentially enhance the power of our test by isolating the role of forecast bias. In addition to full sample regression results we also report results for a restricted sample comprised of FEs from -0.0005 to +0.0005. This comprises the 22nd through 70th percentiles of the empirical distribution of FEs. We repeat the analysis for the range from –0.001 to +0.001, comprising the 17th through 81st percentiles. Since the results for both these restricted samples are essentially the same, we only report results from the former.

Equation (4) explicitly controls for the magnitude of forecast error. Under the bias adjustment hypothesis, investors anticipate that the expectation error as measured by FE_{it} is biased downward (upward) due to an optimistic (pessimistic) forecast and the announcement-period market response to the observed *FE* is systematically larger (smaller) due to the implicit earnings-expectation adjustment. Therefore, we expect, for a forecast error of given magnitude *FE*, that $(\alpha_1 + \beta_1 FE) > (\alpha_2 + \beta_2 FE) > \ldots > (\alpha_5 + \beta_5 FE)$.

<Table 3 here>

Table 3 presents our regression results. Consistent with Lopez and Rees (2002) and others, we find that the ERCs for positive forecast errors (beating forecasts) are greater than the ERCs for negative forecast errors (missing forecasts). From Panel A of Table 3, which gives estimates for equation (3) separately for each category, we find that this result holds across each of the five of the bias categories. Panels B and C gives estimates for equation (4) for the full and restricted samples, respectively, with restrictions that the control variable coefficients, γ_1, γ_2, and γ_3, are equal across the bias categories. Both panels show higher ERCs for positive forecast errors. With the exception of the P5 subset in Panels B and C, we find that after controlling for cross-portfolio differences in the magnitudes of the forecast error and other control variables, the absolute magnitude of the market penalty for misses is greater than the magnitude of the gain from beating a forecast (e.g., from Panel B, the P1 intercept for negative FEs is –0.44%, but +0.72% for positive FEs). This result is also broadly consistent with results reported by Lopez and Rees (2002).

<Table 4 here>

Panels A and B (full sample) and C and D (restricted sample) of Table 4 summarize tests for systematic cross-category differences for intercept (α_n) and slope (β_n) coefficients, for subsets of observations that miss forecasts (negative FEs) and beat forecasts (positive FEs), respectively.

The left side of these panels contains the results of F-tests for all ten pairwise comparisons of the intercept coefficients. Considering observations with negative forecast errors for the full sample (Panels A), we find that, seven out of ten pairwise differences are directionally consistent with the bias adjustment hypothesis, though none of the differences is statistically significant. For the restricted sample (Panel C) we find nine out of ten pairwise differences are consistent with the bias adjustment hypothesis, though only two are statistically significant (one each at the 0.10 and 0.05 levels) in the direction indicated by the hypothesis, and one difference is significant in the opposite direction (α^-_1 vs. α^-_2 is significant at 0.10). Considering Panels B and D for positive FEs, the full sample results in Panel B indicate that eight out of ten pairwise differences are directionally consistent with the bias adjustment hypothesis, with six being significant (four at the 0.01 level and two at the 0.05 level). For the restricted sample, Panel D reveals that seven out of ten pairwise differences are directionally consistent with the bias adjustment hypothesis. Six of which are significant (three at the 0.01 level, one at the 0.05 level, and two at the 0.10 level). In summary, though the results for intercepts overall tend to favor the bias adjustment hypothesis, the results are statistically less than conclusive.

We also evaluate differences in the ERCs across the bias categories. For observations where actual earnings missed the forecasts, the right-hand side of the panels contains the results of F-tests for the ten pairwise comparisons of the slope (i.e., earnings response) coefficients. The results exhibit no consistent pattern in slopes across the bias categories. The full sample results for observations that do not meet forecasts are in Panel A of Table 4. Only one of the pairwise comparisons (β^-_4 vs. β^-_5) is significant in the direction that would yield a higher predicted CAR for the more optimistic forecasts of a given FE, and five of the tests actually support higher ERCs for more pessimistic forecasts. Restricted sample results in Panel C are qualitatively similar, with six of the ten coefficients directionally inconsistent with the bias adjustment hypothesis, though none is significant.

Similar to the results in Panel A, full sample results in Panel B for observations that beat forecasts show that seven of the ten comparisons indicated higher ERCs for more pessimistic forecasts, of which five are statistically significant (two at the 0.01 level, two at the 0.05 level, and one at the 0.10 level), and that only one of the pairwise differences (β^{+}_{1} vs. β^{+}_{2}) is significant in the direction that would yield a higher predicted CAR for more optimistic forecasts. Restricted sample results in Panel D show higher ERCs for more pessimistic forecasts in seven of the ten comparisons, of which three are statistically significant (one at the 0.05 level, and two at the 0.10 level), and that none of the pairwise differences is significant in the direction that would yield a higher predicted CAR for the more optimistic forecasts.

In summary, there does not appear to be a case for higher ERCs for firms with a historically more optimistic forecast bias. Indeed, the results appear to suggest that firms categorized as having more pessimistic historical forecast records tend to have higher ERCs. The observed pattern in ERCs, given the findings with respect to alpha coefficients, would tend to confound our interpretation of the impact of bias assessment on the market response to forecast errors.

To test the hypothesis of a higher market response to a given forecast error for more forecast optimism, this study simultaneously considers the impact of bias category on both alpha and beta (ERC) coefficients. We do this by using the estimated regression parameters from equation (4) to determine the range of forecast errors over which the bias adjustment hypothesis is (is not) supported. To this end, we examine the combined impact of cross-category differences in both the intercept and ERC coefficients at various percentiles of the forecast errors distribution. The cross-category comparisons implicitly control for the effects of the conditioning variables by evaluating the predicted market response (CAR) for firms of average book-to-market ratios ($diffBtoM_{it}$=0), average size ($diffSize_{it}$=0), and with prior period earnings that just meet analysts' expectation ($FE_{i,t-1}$=0). The estimated market impact for bias category n at the pth percentile value of FE, $\overline{CAR}(FE^p, n)$ under these conditions is:

$$\overline{CAR}(FE^p, n) = \alpha_n^- + \beta_n^- \ FE^p \text{ for FE} < 0 \text{ and,}$$

$$\overline{CAR}(FE^p, n) = \alpha_n^+ + \beta_n^+ \ FE^p \text{ for FE} > 0.$$

For each percentile of FE under consideration, we use the Kendall rank-order correlation statistic to test the hypothesis:

$$\overline{CAR}(FE^p, 1) > \overline{CAR}(FE^p, 2) > . \ . \ . \ > \overline{CAR}(FE^p, 5)$$

<Table 5 here>

Panels A and B of Table 5 present the results for the full sample. These panels reflect parameter estimates for negative and positive FEs, respectively, at selected percentile values. Since FE=0 falls at the 40th percentile of the sample, this is the percentile value that separates the negative FEs in Panel A and positive FEs in Panel B, and is a limiting case in each panel. The first two columns of each panel of the table give the percentile, p, of the distribution of FEs and the corresponding percentile value, FE^p. The next five columns provide the predicted CAR for the five bias categories, followed by the Kendall rank-order statistic, S, defined as the number of pairwise comparisons (out of ten) that are consistent with the hypothesis, less the number that are not (e.g., if six of ten are consistent with the hypothesis, then S=6-4=2). The *p*-value associated with S, using small sample probabilities, is from Appendix Table Q in Siegel (1956). The final column lists the difference between the CARs for the extreme optimism category (P1) and the extreme pessimism category (P5), which indicates the potential magnitude of the change in market reaction induced by forecast bias.

From Panel A of Table 5, we find that at both the 5th and 10th percentiles of the FE distribution, nine of the ten pairwise comparisons (yielding S=9-1=8) indicate that the market response to historically optimistic analysts' forecasts is significantly larger (less negative) than the price response associated with more pessimistic analysts' forecasts. This information supports the notion that investors are aware that a given shortfall from the consensus earnings forecast is a relatively better outcome for observations in P1, where analysts have historically tended to overestimate earnings. For example, at the 5th percentile FE, the market response is -1.547% for category P5, but -0.814% for P1, a difference of 0.733% in favor of the category with historically optimistic forecasts. The results of a Kendall rank-order test for both of these percentiles indicate significance in the rank-ordering of CARs from P1 to P5, with a *p*-value of 0.042. At the 20th, 30th, and 40th percentiles, seven of the ten pairwise comparisons are consistent with the hypothesis that the market's reaction to historically optimistic forecasts is higher (or less negative) than the market's response to historically pessimistic forecasts, though the Kendall rank-order statistic *p*-value increases to 0.242, and the differential in CARs also declines.

For positive forecast errors in Panel B of Table 5, the number of pairwise comparisons that favor the bias adjustment hypothesis (i.e., have a positive S-value) outnumbers those that do not through the 80th percentile of the distribution of FEs, although the Kendall test statistic declines steadily. None of these Kendall statistics is significant at the 0.10 level. Comparing the CARs for the extreme categories, P1 and P5, we see the predicted CAR differences declining monotonically through the 80th percentile, from 0.462% to 0.327%.

We performed similar tests for the restricted sample predictions in Panels C and D of Table 5. Since the sample restrictions eliminate the tails of the distribution of FEs, we present results for increments of 0.0001 in FE, from -0.0005 to +0.0005. The restricted sample results are more consistent with the bias adjustment hypothesis, especially for observations which missed their forecasts. Panel C of Table 5 shows that for FEs from the 22nd to the 33rd percentiles (from -0.0005 to -0.0001), the Kendall S-statistic of ten represents a monotonically declining predicted CAR over the range from optimism to pessimism. The *p*-value associated with the rank-ordering of bias categories and predicted CARs is less than 0.01. The rank order statistic is still significant at a 0.05 *p*-value up to an FE of zero. The difference in CARs ranging from the most optimistic to the most pessimistic bias category is remarkably consistent (0.45% to 0.47%). For observations where the firm beats the consensus forecast in Panel D of Table 5, the Kendall S-statistic is four at every FE value in the range from 0.000 to +0.0005 (i.e., seven of the ten pairwise comparisons of CARs are consistent with the notion that the market's reaction to the more optimistic bias category is higher or less negative than the market's response to the more pessimistic category). Although the difference in CARs (the last column) still favor optimistic forecasts by the same order of magnitude as those observed in Panel C, the small-sample *p*-value for the Kendall rank-order statistic increases to 0.242 in these cases.,

In summary, for firms that miss analysts' forecasts, the full sample results are weakly consistent with the hypothesis that investors take historical forecast bias into account when reacting to earnings announcements. Also, this study confirms that a given forecast error results in a larger (less negative) market response for historically more optimistic forecasts. For firms that beat analysts' forecasts, the results for the full sample are even weaker, especially beyond the 50th percentile. The results are considerably strengthened when we consider a restricted range of forecast errors where the FE distribution's tails are truncated. Truncation makes sense on two levels: (1) possible non-linearity in the market's response to forecast errors and (2) isolation of cases where relatively more of the forecast error may be related to historical bias in forecasts.[8]

We can only speculate as to why the historical bias categories approach appears to be more relevant for observations that missed earnings forecasts. Kasznik and McNichols (2002, p.757) offer a competing hypothesis regarding the meaning of the historical record of forecast errors. They find that ". . . the market rewards firms that consistently meet expectations with a distinct market premium incremental to their higher future earnings." They go on to speculate that investors perceive such firms to be less risky and have a lower cost of equity capital. If Kasznik and McNichols are correct, then a record of consistently beating earnings may signal lower risk instead of (or in addition to) a pessimistic bias in analysts' forecasts, thus offsetting or replacing the negative market response differential associated with pessimistic bias with a positive market response differential. Their findings would diminish, even possibly reverse, the implication of this study's categorical bias instrument for those observations we classified as having a relatively pessimistic earnings forecast bias. Indeed, the two hypotheses about how investors interpret the forecast error record are not mutually exclusive, and the fact that the results for negative forecast errors are more consistent with the bias adjustment hypothesis makes some sense in light of the competing hypothesis. For a firm that has a consistent historical record of beating forecasts that investors have interpreted as signifying low risk, the failure to meet a forecast might result in the discarding of the prior belief that the firm is less risky, in favor of a belief that historical analysts' forecasts have exhibited pessimistic bias. This new perception would diminish the risk factor's role vis-à-vis that of bias perception and improve the statistical relevance of the latter hypothesis.

Other potential reasons for lack of stronger empirical support for the bias adjustment hypothesis are easy to imagine. The first, and probably most relevant, is that any test relies on the

[8] To the extent that our results are consistent with the bias adjustment hypothesis, the findings are similar to those of Lopez and Rees (2002, p. 181), who conclude that their findings are ". . . consistent with the market identifying and discounting the systematic component of unexpected earnings."

researcher's ability to assess bias in the same manner as investors. The historical bias categories approach is but one intuitively plausible means of quantifying systematic forecast bias. Undoubtedly, real-world bias assessment is a more complex process. For example, factors such as the demographic characteristics and work experience of a particular firm's analysts probably affect the accuracy of their forecasts. A second impediment to finding empirical evidence consistent with the bias adjustment hypothesis is misspecification related to interpretation of the bias categories. In particular, the pattern of predicted CARs, as well as the raw mean CARs displayed in Table 2, indicates that bias category P3 contains the most favorable market reaction to earnings announcements. Further, since P3 displays the least bias among the five categories, we speculate that there is a market premium for unbiased earnings announcements. Thus, we suspect that higher investor confidence (lower risk) is associated with the information embodied in the announcements. This speculation is consistent with the coefficients reported in Panel C of Table 3. For observations with positive FEs, P3 has both the highest intercept and largest ERC among the five bias categories. Third, the results displayed in Table 4 are based on predicted CAR values from a regression model that explains only about 5% of the variation in CARs, a typical outcome for such models. The most obvious explanation for the weak evidence in favor of the bias adjustment hypothesis is that other phenomena tend to obscure the role of the bias adjustment hypothesis under various circumstances.

6. Robustness checks

6.1. Robustness to regime shift

According to Abarbanell and Lehavy (2007), a regime shift occurred in the IBES database in the early 1990s when IBES began systematically adjusting actual earnings numbers to remove items not forecasted by the majority of analysts.[9] They report evidence that the reason early studies find an optimistic bias in analysts' forecasts is because prior to the early 1990s, the actual IBES earnings metric does not adjust for items not forecasted. To avoid the potential systematic bias in forecast error measurement, we re-estimate the restricted sample regression equation (4) over the period 1996-2001, where their analysis indicates that the earnings reported by IBES are more consistent with analysts' earnings definitions. The re-estimation results (not presented) do not alter our primary results over the whole period 1990-2001, although significance levels for the negative FE range are slightly lower. Thus, we conclude that our results are not attributable to a systematic bias in the IBES data.

6.2. Robustness to scaling

To ensure that the results are not attributable to scaling, we estimate equation (4) for the restricted sample after scaling the forecast error by the standard deviation among individual analyst's forecasts of a firm's earnings for each quarter. Our reason for performing the additional test is because higher dispersion in analysts' forecasts attenuates the "surprise" in a forecast error of given magnitude. The scaled forecast error is UFE_{it} / STD_{it}, where UFE_{it} is the raw analysts' earnings forecast error for firm i at quarter t (not scaled by stock price), and STD_{it} is the standard deviation among analysts' forecasts for firm i at quarter t. Again, the results are essentially unaffected by this method change.

[9] We are greatly indebted to an anonymous referee who made this point. Systematic exclusion of some recurring items not included in analysts' earnings forecasts could influence the bias classification. However, the MFFE method, by using a relatively long series of forecast errors, will probably mitigate the impact of one-time items, thus partially mitigating the impact that any IBES regime shift would have on our results.

6.3. Robustness to unadjusted IBES data

Payne and Thomas (2003) point out potentially serious rounding issues associated with stock split adjustments in the adjusted IBES data. As a result, researchers now generally use the unadjusted detail data. Not only is it not stock-split adjusted, but it has two additional decimal places of precision. To check robustness to this problem, we follow Payne and Thomas's further recommendation by replicating the restricted sample results using the unadjusted data computed by applying IBES's Cumulative Split Factor.[10] These new results are broadly consistent with those calculated using the adjusted earnings data, though significance levels are lower.

6.4. Other checks

We repeat the regression analysis using two alternative market returns metrics. First, using market-adjusted returns in computing the CARs, and second, with market-model adjusted returns based upon (-1, +1) and (0, +1) event windows. The main results are not affected by these alternative methods.

7. Summary and conclusions

This paper investigates whether investors take analysts' forecasts at face value when assessing firms' earnings announcements. Such behavior leads to the prediction that naïve investors' reactions to earnings announcements, as captured by the announcement-period abnormal return associated with a given forecast error (holding other factors constant), do not vary systematically with the estimated bias in analysts' forecasts. By contrast, if investors do

[10] These data were assembled prior to the general availability of IBES's unadjusted data. Rather than reassemble the data, we simply use the split factors included in our data to reverse the split adjustment. Payne and Thomas (2003, p.1066) also suggest that an alternative to using the unadjusted data is to "recalculate I/B/E/S consensus statistics using the I/B/E/S adjusted data. The detail data are rounded to four decimal places, which allow a more accurate recalculation of the actual amount using the split factor provided by I/B/E/S."

possess the ability to use historical experience to assess forecast bias, we would expect them to apply a negative earnings-expectation adjustment to forecasts characterized by historical optimism, and a positive earnings-expectation adjustment to forecasts having historical pessimism.

Testing requires a characterization of earnings forecast bias. We follow Kwag and Shrieves (2006) by classifying a current forecast into five rank-ordered bias categories in an optimism-pessimism continuum. We employ a multiple regression to determine whether investors respond differently to forecast errors based on the rank-ordered bias categories. The regression model includes the contemporaneous forecast errors as an explanatory variable, along with control variables for lagged forecast errors, whether the observed error misses or beats expectations, firm size, and book-to-market characteristics of firms. Under the bias adjustment hypothesis, investors capable of even partial assessment of forecast bias will apply an adjustment to analysts' earnings estimates. As a result, the bias category intercept and earnings response coefficients will reflect whether the excess returns for a given forecast error are systematically larger for more optimistic forecasts.

The regression results for the full range of observed forecast errors provide only weak support for the bias adjustment hypothesis, and only in cases where firms missed their forecasts. The strength of the results improves considerably when we restrict the sample by excluding large forecast errors to focus only on those cases where the role of bias in explaining the magnitude of observed forecast error is expected to be larger. In this restricted sample, regardless of whether firms miss or beat forecasts, the market's reaction to earnings announcements is higher for firms with historically optimistic forecasts relative to those with historically more pessimistic forecasts. The potential magnitude of the change in market reaction induced by forecast bias is estimated as the difference in the three-day abnormal market response to a given forecast error between the top and bottom bias classification quintiles, and is about 0.4% in favor of forecasts classified as optimistic. We caution that while results are highly significant for firms that miss forecasts, they are statistically weaker for firms that beat consensus forecasts.

Over a wide range of forecast errors, the results suggest that investors are sufficiently informed to distinguish between optimistic and pessimistic forecast bias, especially for firms where the historical forecast error record suggests extreme bias. The results agree substantially with the conclusions of a growing number of studies that the functional form commonly used in the earnings response literature does not appropriately capture the effect of real unexpected earnings information (i.e., investors' expectation errors, as opposed to analysts' forecast errors) on stock returns.

Methodological and data limitations lead us to qualify our conclusions. First, the impact of survivorship bias associated with the sample selection process clouds the assessment of the relation between ex ante earnings forecast bias classification and stock returns. Second, more elaborate methods of bias classification could potentially reduce the measurement error that remains in our bias metric. This is important, since we could reject a hypothesis that investors rationally evaluate bias, even when it is true, if the bias classification procedure is poor. Potential factors that would enable a more accurate forecast classification according to bias include momentum effects and the characteristics of individual analysts such as experience, brokerage relationship, age, and reputation.

References

Abarbanell, J. S. and V. L. Bernard, 1992. Tests of analysts' overreaction / underreaction to earnings information as an explanation for anomalous stock price behavior, Journal of Finance 47, 1181-1207.

Abarbanell, J. S. and R. Lehavy, 2002. Differences in commercial database reported earnings: Implications for empirical research, *Working paper*, University of North Carolina and University of Michigan.

Abarbanell, J. S. and R. Lehavy, 2003. Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts, Journal of Accounting and Economics 36, 105-146.

Abarbanell, J. S. and R. Lehavy, 2007. Letting the "tail wag the dog": The debate over GAAP versus street earnings revisited, Contemporary Accounting Research 24, 657-674.

Ackert, L. F. and G. Athanassakos, 1997. Prior uncertainty, analyst bias, and subsequent abnormal returns, Journal of Financial Research 20, 263-273.

Ball, R. and S. P. Kothari, 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns, Journal of Financial Economics 25, 51-74.

Ball, R. and S. P. Kothari, 1991. Security returns around earnings announcements, Accounting Review 66, 718-738.

Bartov, E., D. Givoly, and C. Hayn, 2002. The reward to meeting or beating earnings expectations, Journal of Accounting & Economics 33, 173-204.

Bernard, V. and J. Thomas, 1989. Post-earnings-announcement drift: Delayed price response or risk premium?, Journal of Accounting Research 27, 1-36.

Brown, L. D. and M.S. Rozeff, 1978. The superiority of analyst forecasts as measures of expectations: Evidence from earnings, Journal of Finance 33, 1-16.

Brown, L. D., R. Hagerman, P. Griffin, and M. Zmijewski, 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings, Journal of Accounting and Economics 9, 61-87.

Chari, V. V., R. Jagannathan, and A. R. Ofer, 1988. Seasonalities in security returns: The case of earnings announcements, Journal of Financial Economics 21, 101-121.

Cohen, D. A., R. N. Hann, and M. Ogneva, 2007. Another look at GAAP versus the street: An empirical assessment of measurement error bias, Review of Accounting Studies 12, 271-303.

Conroy, R. and R. Harris, 1987. Consensus forecasts of corporate earnings: analysts' forecasts and time series methods, Management Science 33, 725-738.

Cornell, B. and W. R. Landsman, 1989. Security price response to quarterly earnings announcements and analysts' forecast revisions, Accounting Review 64, 680-692.

Dechow, P. M. and R. G. Sloan, 1997. Returns to contrarian investment strategies: Tests of naïve expectations hypotheses, Journal of Financial Economics 43, 3-27.

Fama, E. F. and K. R. French, 1992. The cross-section of expected stock returns, Journal of Finance 47, 427-465.

Fama, E. F. and K. R. French, 1993. Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 3-56.

Fama, E. F. and K. R. French, 1995. Size and book-to-market factors in earnings and returns, Journal of Finance 50, 131-155.

Freeman, R. N. and S. Y. Tse, 1992. A nonlinear model for security price responses to unexpected earnings, Journal of Accounting Research 30, 185-209.

Fried, D. and D. Givoly, 1982. Financial analysts' forecasts of earnings: A better surrogate for market expectations, Journal of Accounting & Economics 4, 85-107.

Givoly, D. and J. Lakonishok, 1984. The quality of analysts' forecasts of earnings, Financial Analysts Journal 40, 40-47.

Hughes, J. S. and W. E. Ricks, 1987. Associations between forecast errors and excess returns near to earnings announcements, Accounting Review 62, 158-175.

Kasznik, R. and M. F. McNichols, 2002. Does Meeting Earnings Expectations Matter? Evidence from Analyst Forecast Revisions and Share Prices, Journal of Accounting Research 40, 727-759.

Kross, W., B. Ro, and D. Schroeder, 1990. Earnings expectations: Analysts' information advantage, Accounting Review 65, 461-476.

Kwag, S. and R. E. Shrieves, 2006. Chronic bias in earnings forecasts, Financial Analysts Journal 62, 81-96.

Lopez, T. and L. Rees, 2002. The effect of beating and missing analysts' forecasts on the information content of unexpected earnings, Journal of Accounting, Auditing, and Finance 17, 155-184.

Mendenhall, R. R., 1991. Evidence on the possible underweighting of earnings-related information, Journal of Accounting Research 29, 170-179.

O'Brien, P. C., 1988. Analysts' forecasts as earnings expectations, Journal of Accounting and Economics 10, 53-83.

Payne, J. L. and W. B. Thomas, 2003. The implications of using stock-split adjusted I/B/E/S data in empirical research, Accounting Review 78, 1049-1067.

Shane, P. and P. Brous, 2001. Investor and (Value Line) analyst underreaction to information about future earnings: The corrective role of non-earnings-surprise information, *Journal of Accounting Research 39, 387-404.*

Siegel, S., 1956. *Nonparametric Statistics for the Behavioral Sciences*. (McGraw Hill Book Company, New York).

Table 1

Results of MFFE bias category classification

Q1 to Q5 are quintiles of MQFE scores (columns) and WNNFE scores (rows). MQFE and WNNFE are measures of optimism or pessimism in analysts' consensus forecasts. MQFE is the mean of the prior 20 calendar quarters' forecast errors (negative for optimism, positive for pessimism), and WNNFE is a weighted average of the number of negative forecast errors over the prior 20 quarters (a higher fraction of negative forecasts is associated with greater optimism). Mean category MQFE is increasing (going from negative to positive) and WNNFE is decreasing as the category index number increases. The rationale for category designations P1 to P5 is that observations in each category should dominate observations in the next higher numbered category on at least one of the two metrics, and be at least equivalent on the other. Figures in the parentheses denote the number of quarterly analysts' forecasts in each cell.

	MQFE rank					
WNNFE rank	Q1	Q2	Q3	Q4	Q5	Row total
Q1	P1 (3,277)	P1 (2,430)	(824)	(200)	(236)	6,967
Q2	P1 (1,861)	P2 (1,919)	P2 (1,631)	(850)	(767)	7,028
Q3	(1,033)	P2 (1,335)	P3 (1,829)	P4 (1,452)	(1,304)	6,953
Q4	(544)	(900)	P4 (1,486)	P4 (1,965)	P5 (2,016)	6,911
Q5	(247)	(378)	(1,192)	P5 (2,495)	P5 (2,639)	6,951
Column total	6,962	6,962	6,962	6,962	6,962	34,810

Table 2

Descriptive statistics
N is the number of observations in firm-quarters; CAR is the cumulative three-day (−2, 0) earnings-announcement abnormal return for quarter t; MQFE is the mean quarterly forecast error over one through 20 quarters prior to the quarterly announcement for quarter t; FE is the quarterly forecast error of firm i for quarter t (i.e., actual quarterly earnings less the most recent consensus analysts' earnings forecast of firm i for quarter t, scaled by stock price ten days prior to the associated earnings announcement date); Size is the market value of equity (MVE) in millions; logSize is the logarithm of Size; BtoM is the book value of common equity/market value of common equity; STD is the standard deviation among individual analysts' forecasts of earnings for firm i, quarter t; PCTPOS is the percentage of positive forecast errors; and PCTNEG is the percentage of negative forecast errors; ANALYSTS is the average number of analysts for a forecast; and Pn is a binary variable that equals one if an observation belongs to category n using the MFFE bias category classification, and zero otherwise, n = 1, …, 5. P1 denotes the most optimistic category; P5 denotes the most pessimistic category. NA indicates data not collected.

	Full sample	P1	P2	P3	P4	P5
N	19,745	5,075	3,787	1,457	3,896	5,530
CAR	0.00319***	0.00374***	0.00391***	0.00528***	0.00286**	0.00186**
MQFE	-0.00051***	-0.00251***	-0.00031***	-0.00001***	0.00011***	0.00063***
mean FE	-0.00019***	-0.00108***	-0.00012**	-0.000003	0.00009***	0.00034***
median FE	0.0001	-0.00013	0.00006	0.00005	0.00009	0.00023
ANA-LYSTS	NA	6.91	7.79	8.6	8.25	7.83
Size	6,548.77	2,558.53	5,269.62	9,684.4	9,612.16	8,102.3
logSize	7.4961	6.8194	7.4179	7.9235	7.9173	7.7611
BtoM	0.4951	0.6411	0.5109	0.392	0.4078	0.4391
STD	NA	0.34	0.17	0.09	0.1	0.11
PCTPOS	59	46	56	62	65	71
PCTNEG	40	54	44	38	35	29

***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

Table 3

Regression of CAR on signed forecast errors controlling for size and book-to-market ratio

Panel A reports separate parameter estimates for each bias category under the regression model:

$$CAR_{it} = \alpha^{+}POSDUM_{it} + \alpha^{-}NEGDUM_{it} + \beta^{+}(FE_{it} * POSDUM_{it}) + \beta^{-}(FE_{it} * NEGDUM_{it})$$
$$+ \gamma_1 FE_{i,t-1} + \gamma_2 diffSize_{it} + \gamma_3 diffBtoM_{it} + u_{it}$$

Panel B reports results where α_n and β_n coefficients are allowed to vary according to positive or negative forecast error, while restricting the remaining coefficients to be identical. Panel C further restricts the values of the forecast error to the range -0.0005 to +0.0005. The variables are defined as follows: CAR_{it} is the cumulative three-day (–2, 0) earnings-announcement abnormal return; Pn_{it} is a binary variable that equals one if an observation belongs to bias category n and zero otherwise, n = 1, …, 5; FE_{it} is the analysts' earnings forecast error for firm i at quarter t, scaled by stock price ten days prior to the earnings announcement date; $diffSize_{it}$ is the difference between $\log(MVE_{it})$ and the grand mean of $\log(MVE_{it})$ where MVE_{it} is market value of equity in millions; $diffBtoM_{it}$ is the difference between $BtoM_{it}$ and the grand mean of $BtoM_{it}$, where $BtoM_{it}$ is the ratio of book value of equity to market value of equity; and u_{it} is a random error term. $POSDUM_{it}$ and $NEGDUM_{it}$ are binary variables which are unity for observations with positive and negative, forecast errors, respectively, and zero otherwise. The numbers in parentheses are heteroskedasticity-adjusted *p*-values. We use the generalized method of moments (GMM) to generate heteroskedasticity-adjusted *p*-values.

Panel A: Separate regressions of CAR on signed forecast errors for each bias category

	Bias category				
Coefficients	P1	P2	P3	P4	P5
α^{+}	0.00722	0.01070	0.01275	0.00574	0.00356
	(<0.0001)	(<0.0001)	(<0.0001)	(0.0001)	(0.0031)
α^{-}	-0.00442	-0.00470	-0.00173	-0.00575	-0.00481
	(0.0002)	(0.0037)	(0.4551)	(0.0003)	(0.0170)
β^{+}	3.33089	2.39767	4.43159	6.09738	4.83383
	(<0.0001)	(0.0008)	(0.0465)	(<0.0001)	(<0.0001)
β^{-}	0.80961	0.73689	2.42806	3.98539	2.29165
	(0.0050)	(0.3320)	(0.1101)	(<0.0001)	(0.0637)
γ_1	-0.15582	-0.03455	-0.83853	-0.68814	-0.82481
	(0.1145)	(0.8483)	(0.0862)	(0.1627)	(0.0623)
γ_2	-0.00003	-0.00239	-0.00123	-0.00046	-0.00086
	(0.9660)	(<0.0001)	(0.1785)	(0.4726)	(0.1215)
γ_3	0.00850	0.00746	0.01045	0.00396	0.00808
	(0.0120)	(0.0600)	(0.0963)	(0.3421)	(0.0429)
N	4,878	3,749	1,439	3,890	5,547
F-value	51.11	29.23	15.11	32.74	28.94
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Adjusted R-squared	0.0671	0.0501	0.0642	0.0540	0.0341

Table 3 (*continued*)

Panel B: Regression results for full sample					
Bias category:	P1	P2	P3	P4	P5
α^+	0.0072	0.0114	0.0111	0.0060	0.0026
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(0.0156)
α^-	-0.0044	-0.0048	-0.0030	-0.0055	-0.0055
	(0.0003)	(0.0025)	(0.1418)	(0.0002)	(0.0051)
β^+	3.3755	1.9123	4.8991	5.8909	4.8399
	(<0.0001)	(0.0039)	(0.0236)	(<0.0001)	(<0.0001)
β^-	0.7813	0.9558	1.9255	3.9710	2.0962
	(0.0061)	(0.2006)	(0.1891)	(<0.0001)	(0.0862)
γ_1			-0.1984		
			(0.0215)		
γ_2			-0.0009		
			(0.0017)		
γ_3			0.0036		
			(0.1227)		
N			19,503		
F-value			45.16		
p-value			<0.0001		
Adjusted R-squared			0.0495		
Panel C: Regression results for restricted sample					
Bias category:	P1	P2	P3	P4	P5
α^+	0.00661	0.00739	0.00898	0.00460	0.00209
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(0.0301)
α^-	0.000232	-0.00321	-0.00362	-0.00353	-0.00426
	(0.8705)	(0.0294)	(0.1141)	(0.0166)	(0.0032)
β^+	3.03997	4.72137	6.61992	6.50557	4.88591
	(0.0001)	(<0.0001)	(0.0006)	(<0.0001)	(<0.0001)
β^-	2.82925	2.16061	1.73946	2.66829	3.15022
	(0.0001)	(0.0354)	(0.4132)	(0.0385)	(0.0022)
γ_1			-0.21453		
			(0.0043)		
γ_2			-0.00063		
			(0.0083)		
γ_3			0.00160		
			(0.2279)		
N			9,072		
F-value			28.23		
p-value			<0.0001		
Adjusted R-squared			0.0364		

Table 4

Pairwise tests for equivalence of bias category: Negative versus positive forecast error

For pairwise tests of the five bias categories, the following is estimated:

$$CAR_{it} = \alpha^{+}POSDUM_{it} + \alpha^{-}NEGDUM_{it} + \beta^{+}(FE_{it} * POSDUM_{it}) + \beta^{-}(FE_{it} * NEGDUM_{it})$$
$$+ \gamma_1 FE_{i,t-1} + \gamma_2 diffSize_{it} + \gamma_3 diffBtoM_{it} + u_{it}$$

The table summarizes the pairwise tests for systematic cross-category differences for intercept (α_n) and slope (β_n) coefficients, for subsets of observations that miss forecasts (negative FEs) and beat forecasts (positive FEs), respectively. The left side of both panels contains the results of F-tests for all ten pairwise comparisons of the intercept coefficients (α^{-}_{n} and α^{+}_{n}), while the right-hand side of the panels contains the results of F-tests for the ten pairwise comparisons of the earnings response coefficients (β^{-}_{n} and β^{+}_{n}). All variables are as defined in Table 3.

Tests for α^{-}_{n}				Tests for β^{-}_{n}			
Test	Test values	F-value	*p*-value	Test	Test values	F-value	*p*-value
Panel A: F-tests for equivalence of coefficients from the full sample for negative FE							
$\alpha^{-}_{1}=\alpha^{-}_{2}$	-0.0044=-0.0048	0.04	0.841	$\beta^{-}_{1}=\beta^{-}_{2}$	0.7813=0.9558	0.17	0.683
$\alpha^{-}_{1}=\alpha^{-}_{3}$	-0.0044=-0.0030	0.28	0.598	$\beta^{-}_{1}=\beta^{-}_{3}$	0.7813=1.9255	1.65	0.199
$\alpha^{-}_{1}=\alpha^{-}_{4}$	-0.0044=-0.0055	0.32	0.574	$\beta^{-}_{1}=\beta^{-}_{4}$	0.7813=3.9710	26.81	<0.0001
$\alpha^{-}_{1}=\alpha^{-}_{5}$	-0.0044=-0.0055	0.30	0.586	$\beta^{-}_{1}=\beta^{-}_{5}$	0.7813=2.0962	5.19	0.023
$\alpha^{-}_{2}=\alpha^{-}_{3}$	-0.0048=-0.0030	0.42	0.519	$\beta^{-}_{2}=\beta^{-}_{3}$	0.9558=1.9255	1.04	0.308
$\alpha^{-}_{2}=\alpha^{-}_{4}$	-0.0048=-0.0055	0.12	0.731	$\beta^{-}_{2}=\beta^{-}_{4}$	0.9558=3.9710	18.58	<0.0001
$\alpha^{-}_{2}=\alpha^{-}_{5}$	-0.0048=-0.0055	0.10	0.747	$\beta^{-}_{2}=\beta^{-}_{5}$	0.9558=2.0962	2.94	0.086
$\alpha^{-}_{3}=\alpha^{-}_{4}$	-0.0030=-0.0055	0.80	0.371	$\beta^{-}_{3}=\beta^{-}_{4}$	1.9255=3.9710	3.80	0.051
$\alpha^{-}_{3}=\alpha^{-}_{5}$	-0.0030=-0.0055	0.78	0.378	$\beta^{-}_{3}=\beta^{-}_{5}$	1.9255=2.0962	0.03	0.868
$\alpha^{-}_{4}=\alpha^{-}_{5}$	-0.0055=-0.0055	0.00	0.980	$\beta^{-}_{4}=\beta^{-}_{5}$	3.9710=2.0962	5.49	0.019
Tests for α^{+}_{n}				Tests for β^{+}_{n}			
Panel B: F-tests for equivalence of coefficients from the full sample for positive FE							
$\alpha^{+}_{1}=\alpha^{+}_{2}$	0.0072=0.0114	4.63	0.032	$\beta^{+}_{1}=\beta^{+}_{2}$	3.3755=1.9123	3.93	0.048
$\alpha^{+}_{1}=\alpha^{+}_{3}$	0.0072=0.0111	2.48	0.116	$\beta^{+}_{1}=\beta^{+}_{3}$	3.3755=4.8991	0.77	0.381
$\alpha^{+}_{1}=\alpha^{+}_{4}$	0.0072=0.0060	0.40	0.527	$\beta^{+}_{1}=\beta^{+}_{4}$	3.3755=5.8909	6.02	0.014
$\alpha^{+}_{1}=\alpha^{+}_{5}$	0.0072=0.0026	6.98	0.008	$\beta^{+}_{1}=\beta^{+}_{5}$	3.3755=4.8399	4.54	0.033
$\alpha^{+}_{2}=\alpha^{+}_{3}$	0.0114=0.0111	0.02	0.888	$\beta^{+}_{2}=\beta^{+}_{3}$	1.9123=4.8991	2.76	0.096
$\alpha^{+}_{2}=\alpha^{+}_{4}$	0.0114=0.0060	8.79	0.003	$\beta^{+}_{2}=\beta^{+}_{4}$	1.9123=5.8909	12.66	0.000
$\alpha^{+}_{2}=\alpha^{+}_{5}$	0.0114=0.0026	27.41	<0.0001	$\beta^{+}_{2}=\beta^{+}_{5}$	1.9123=4.8399	12.73	0.000
$\alpha^{+}_{3}=\alpha^{+}_{4}$	0.0111=0.0060	4.71	0.030	$\beta^{+}_{3}=\beta^{+}_{4}$	4.8991=5.8909	0.26	0.607
$\alpha^{+}_{3}=\alpha^{+}_{5}$	0.0111=0.0026	14.45	0.000	$\beta^{+}_{3}=\beta^{+}_{5}$	4.8991=4.8399	0.00	0.973
$\alpha^{+}_{4}=\alpha^{+}_{5}$	0.0060=0.0026	4.76	0.029	$\beta^{+}_{4}=\beta^{+}_{5}$	5.8909=4.8399	0.94	0.332

Table 4 (*continued*)

Tests for α^-_n				Tests for β^-_n			
Test	Test values	F-value	*p*-value	Test	Test values	F-value	*p*-value
Panel C: F-tests for equivalence of coefficients from the restricted sample for negative FE							
$\alpha^-_1=\alpha^-_2$	0.0002=-0.0032	2.83	0.0927	$\beta^-_1=\beta^-_2$	2.8292=2.1606	0.28	0.5944
$\alpha^-_1=\alpha^-_3$	0.0002=-0.0036	2.04	0.1533	$\beta^-_1=\beta^-_3$	2.8292=1.7394	0.24	0.6271
$\alpha^-_1=\alpha^-_4$	0.0002=-0.0035	3.38	0.0661	$\beta^-_1=\beta^-_4$	2.8292=2.6682	0.01	0.9133
$\alpha^-_1=\alpha^-_5$	0.0002=-0.0042	4.90	0.0269	$\beta^-_1=\beta^-_5$	2.8292=3.1502	0.07	0.7983
$\alpha^-_2=\alpha^-_3$	-0.0032=-0.0036	0.02	0.8820	$\beta^-_2=\beta^-_3$	2.1606=1.7394	0.03	0.8579
$\alpha^-_2=\alpha^-_4$	-0.0032=-0.0035	0.02	0.8768	$\beta^-_2=\beta^-_4$	2.1606=2.6682	0.10	0.7569
$\alpha^-_2=\alpha^-_5$	-0.0032=-0.0042	0.26	0.6091	$\beta^-_2=\beta^-_5$	2.1606=3.1502	0.47	0.4932
$\alpha^-_3=\alpha^-_4$	-0.0036=-0.0035	0.00	0.9760	$\beta^-_3=\beta^-_4$	1.7394=2.6682	0.14	0.7077
$\alpha^-_3=\alpha^-_5$	-0.0036=-0.0042	0.06	0.8093	$\beta^-_3=\beta^-_5$	1.7394=3.1502	0.36	0.5487
$\alpha^-_4=\alpha^-_5$	-0.0035=-0.0042	0.13	0.7213	$\beta^-_4=\beta^-_5$	2.6682=3.1502	0.09	0.7688
Tests for α^+_n				Tests for β^+_n			
Panel D: F-tests for equivalence of coefficients from the restricted sample for positive FE							
$\alpha^+_1=\alpha^+_2$	0.0066=0.0073	0.17	0.6809	$\beta^+_1=\beta^+_2$	3.0399=4.7213	1.90	0.1683
$\alpha^+_1=\alpha^+_3$	0.0066=0.0089	1.09	0.2958	$\beta^+_1=\beta^+_3$	3.0399=6.6199	2.93	0.0868
$\alpha^+_1=\alpha^+_4$	0.0066=0.0046	1.23	0.2671	$\beta^+_1=\beta^+_4$	3.0399=6.5055	6.06	0.0138
$\alpha^+_1=\alpha^+_5$	0.0066=0.0020	7.03	0.0080	$\beta^+_1=\beta^+_5$	3.0399=4.8859	2.91	0.0878
$\alpha^+_2=\alpha^+_3$	0.0073=0.0089	0.54	0.4630	$\beta^+_2=\beta^+_3$	4.7213=6.6199	0.78	0.3757
$\alpha^+_2=\alpha^+_4$	0.0073=0.0046	2.76	0.0966	$\beta^+_2=\beta^+_4$	4.7213=6.5055	1.44	0.2296
$\alpha^+_2=\alpha^+_5$	0.0073=0.0020	11.41	0.0007	$\beta^+_2=\beta^+_5$	4.7213=4.8859	0.02	0.8891
$\alpha^+_3=\alpha^+_4$	0.0089=0.0046	4.48	0.0343	$\beta^+_3=\beta^+_4$	6.6199=6.5055	0.00	0.9598
$\alpha^+_3=\alpha^+_5$	0.0089=0.0020	12.08	0.0005	$\beta^+_3=\beta^+_5$	6.6199=4.8859	0.70	0.4013
$\alpha^+_4=\alpha^+_5$	0.0046=0.0020	3.07	0.0799	$\beta^+_4=\beta^+_5$	6.5055=4.8859	1.39	0.2377

Table 5

Rank order tests of the bias adjustment hypothesis at various percentiles of the distribution of forecast errors

$$\overline{CAR}(FE^p,1) > \overline{CAR}(FE^p,2) > \ldots > \overline{CAR}(FE^p,5)$$

Under the hypothesis, CARs are predicted to be higher for more optimistic bias category than for a less optimistic category. The number in the column "S" is the number of pairwise comparisons (out of ten) that support the hypothesis, less the number that do not (e.g., if six of ten support the hypothesis, then S=6-4=2). One-tailed *p*-values in the last column are for small samples, taken from Table Q of the Appendix to Siegel (1956). Percentiles designated as 40^- and 40^+ are asymptotic limiting values approached from below and above zero, respectively (since the sample used excludes zero FEs).

Forecast error		Predicted CAR value					Kendall S		
Per-centile	value	P1	P2	P3	P4	P5	S	*p*-value	P1 less P5
Panel A: Full sample results for negative FE									
5	-0.00477	-0.00814	-0.00936	-0.01217	-0.02447	-0.01547	8	0.042	0.0073
10	-0.00223	-0.00615	-0.00693	-0.00727	-0.01437	-0.01014	8	0.042	0.0040
20	-0.00067	-0.00493	-0.00544	-0.00426	-0.00817	-0.00687	4	0.242	0.0019
30	-0.00018	-0.00455	-0.00498	-0.00333	-0.00625	-0.00586	4	0.242	0.0013
40^-	0.00000	-0.00441	-0.00480	-0.00298	-0.00552	-0.00547	4	0.242	0.0011
Panel B: Full sample results for positive FE									
40^+	0.00000	0.00721	0.01143	0.01109	0.00602	0.00259	6	0.117	0.0046
50	0.00010	0.00756	0.01163	0.01159	0.00662	0.00309	6	0.117	0.0045
60	0.00026	0.00807	0.01192	0.01234	0.00752	0.00383	4	0.242	0.0042
70	0.00049	0.00886	0.01237	0.01349	0.00891	0.00496	2	0.408	0.0039
80	0.00092	0.01031	0.01319	0.01559	0.01143	0.00704	2	0.408	0.0033
90	0.00194	0.01376	0.01514	0.02060	0.01745	0.01198	0	0.592	0.0018
95	0.00341	0.01872	0.01795	0.02779	0.02610	0.01909	-2	0.408	-0.0004
Panel C: Restricted sample results for negative FE									
22	-0.00050	-0.00118	-0.00429	-0.00449	-0.00486	-0.00584	10	0.008	0.0047
24	-0.00040	-0.00090	-0.00407	-0.00432	-0.00460	-0.00552	10	0.008	0.0046
26	-0.00030	-0.00062	-0.00386	-0.00414	-0.00433	-0.00521	10	0.008	0.0046
29	-0.00020	-0.00033	-0.00364	-0.00397	-0.00406	-0.00489	10	0.008	0.0046
33	-0.00010	-0.00005	-0.00343	-0.00379	-0.00380	-0.00458	10	0.008	0.0045
40^-	0.00000	0.00023	-0.00321	-0.00362	-0.00353	-0.00426	8	0.042	0.0045
Panel D: Restricted sample results for positive FE									
40^+	0.00000	0.00661	0.00739	0.00898	0.00460	0.00209	4	0.242	0.0043
49	0.00010	0.00691	0.00786	0.00964	0.00525	0.00258	4	0.242	0.0043
56	0.00020	0.00722	0.00833	0.01030	0.00590	0.00307	4	0.242	0.0042
62	0.00030	0.00752	0.00881	0.01097	0.00655	0.00356	4	0.242	0.0040
66	0.00040	0.00783	0.00928	0.01163	0.00720	0.00404	4	0.242	0.0038
70	0.00050	0.00813	0.00975	0.01229	0.00785	0.00453	4	0.242	0.0036