

Surprising absence of scale for forecast error magnitudes and forecast dispersion

Foong Soon Cheong
Rutgers University
fscheong@rutgers.edu

Jacob Thomas
Yale University
jake.thomas@yale.edu

Current Version: October 20, 2009

We received helpful comments from an anonymous referee, Jeff Abarbanell, Mark Bagnoli, Larry Brown, Jim Ohlson, Phil Shane, Frank Zhang and seminar participants at AAA Annual meetings, Arizona, FARS Midyear meetings, Georgetown, HKUST, London Business School, Miami, Purdue, Rutgers, and Yale. We thank Shawn Thomas for graciously providing us the data that allowed us to illustrate our conjectures about the impact of price deflation.

Surprising absence of scale for forecast error magnitudes and forecast dispersion

Abstract

While levels of actual and consensus forecast earnings per share (EPS) vary with scale (measured typically by share price), magnitudes of the difference do not vary with scale. That is, forecast errors within a certain range (e.g., ± 5 cents per share) are equally likely for both high-price and low-price shares. We also find a similar lack of variation with scale for forecast dispersion, representing magnitudes of the difference between individual forecasts and the consensus (mean) for that firm-quarter. The prior literature has assumed that magnitudes of forecast errors (representing predictability) and forecast dispersion (representing disagreement across analysts) vary naturally with scale and has deflated both variables accordingly. We show that such scaling is likely to cause biased estimates, and recommend that scaling not be used, unless called for by theory. Regardless, a scale variable should be included as an additional regressor.

Surprising absence of scale for forecast error magnitudes and forecast dispersion

1. Introduction

Since both actual earnings per share (EPS) and consensus forecasts vary with the scale of individual shares, where scale is typically measured as price per share, conventional wisdom is that magnitudes of the difference should also vary with scale (see Appendix B for examples). That is, the distribution of forecast errors for high-price shares should be associated with larger absolute forecast errors and larger measures of variability, such as the variance and interquartile range. To investigate the validity of this intuition we examine the distributions of forecast errors for deciles of share price, where errors are measured as actual quarterly EPS according to I/B/E/S less the most recent consensus analyst forecast available. We find, much to our surprise, that measures of forecast error variability are similar across different price deciles.

To explore further why forecast error variability does not increase with share prices, we investigate the dispersion of individual forecasts around the consensus. Prior research (e.g., Barron et al. [1998]) has emphasized differences between the two constructs: whereas the variability of forecast errors is an across-firm-quarter measure of *predictability*—the ability of consensus forecasts to accurately predict actual EPS—forecast dispersion captures *disagreement* across analysts around the consensus for the same firm-quarter. Despite these differences, conventional wisdom (see Appendix B for examples) holds that disagreement should also vary with scale, similar to predictability. Disagreement across analysts, measured in cents per share, must surely be higher for higher priced shares with larger values of forecast EPS. Again, much to our surprise, we find that forecast dispersion is also similar across price deciles.

Our first objective is to document the two surprising empirical regularities and to confirm that they are robust. Our second objective is to illustrate the extent to which our findings affect results in prior research that uses deflated predictability and disagreement. Because predictability

and disagreement do not in fact vary with scale, deflation induces a strong negative correlation with scale. As a result, using deflated measures as dependent (independent) variables could bias results if the included independent (dependent) variables happen to be correlated with price.

Before summarizing our results, we provide some background and describe labels we use for different constructs (see Appendix B). I/B/E/S data are based on forecasts of EPS (F_{ijt}) made by analyst j for firm i in quarter t . The mean (μ_{it}) and standard deviation (σ_{it}) of the distribution of individual forecasts for each firm-quarter are referred to as the consensus forecast and forecast dispersion, respectively. Consensus forecasts are subtracted from the actual EPS reported for that firm-quarter to generate forecast errors (FE_{it}). These forecast errors are then pooled across firm-quarters to generate a second set of distributions, described by parameters such as the mean (μ) and standard deviation (σ) representing forecast bias and predictability, respectively.¹

We collect a sample of firm-quarters from the I/B/E/S files, each with a forecast error (FE_{it}) and a dispersion value (σ_{it}). We then group observations into deciles every calendar year, based on beginning-of-year share prices, and generate distributions for forecast error and dispersion for those price deciles. We consider three measures of forecast error variability when investigating predictability: standard deviation, interquartile range, and the range between the 5th and 95th percentiles. We also consider mean/median *absolute* forecast errors as an alternative measure of predictability, but use those results only for purposes of illustration since absolute errors overstate variability slightly because mean errors are systematically different from zero, especially for more extreme price deciles. When investigating disagreement we focus on two

¹ We refer to these second set of distributions for illustrative purposes, since they may not actually arise in practice. For example, in analyses based on regressions of forecast errors on its determinants, bias may refer to the intercept and predictability may refer to variability of the error term.

measures of central tendency for the dispersion distributions—mean and median—since dispersion at the firm-quarter level is already a measure of variability.

The following statistics illustrate how substantially our findings deviate from the intuition in the literature. To provide context, the mean (median) beginning-of-year share price of our lowest price decile is approximately \$5 (\$5), which is less than one-tenth the mean (median) share price of our highest price decile of approximately \$70 (\$60). And the magnitudes of forecast and actual EPS vary proportionately with this substantial variation in the scale of share prices across price deciles. Despite this substantial variation across price deciles in the scale of actual and forecast EPS, disagreement and predictability vary only slightly with share price. The mean (median) dispersion of forecasts around the consensus forecast is 3 (1) cents for the lowest priced decile versus 5 (2) cents for the highest priced decile, and the standard deviation (interquartile range) of forecast errors for the lowest price decile is 24 (5) cents versus 26 (5) cents for the highest price decile.

As a first approximation our evidence suggests that forecast dispersion and forecast errors of a particular amount (say within ± 5 cents) are as likely for a \$5 stock as they are for a \$70 stock. We conduct a host of sensitivity analyses and confirm that both findings are robust. Similar results are observed in different time periods, across industry partitions, and different measures of scale. We also confirm that similar lack of variation for predictability and disagreement is observed across different variables commonly used in analyst forecast research, including analyst coverage, forecast staleness, and return volatility.

Prior research, apparently unaware of these results, has relied on the common wisdom regarding predictability and disagreement varying with scale and deflated both variables before using them as dependent or independent variables. In our sample, the standard deviation

(interquartile range) for the distribution of price-deflated forecast errors declines sharply from 12 (1) percent of price for the lowest price decile to 0.32 (0.08) percent for the highest price decile. Similarly, the mean (median) price-deflated dispersion declines from 0.79 (0.32) percent of price for the lowest price decile to 0.07 (0.04) for the highest price decile. This large negative correlation between price and price-deflated measures of predictability and disagreement could bias coefficient estimates from analyses based on deflated measures.

To illustrate the potential for such biases, we extend the investigation in Thomas [2002] that relates variation in price-deflated predictability and disagreement to the degree of firm diversification, which happens to be correlated to share price. Our results confirm the presence of substantial biases and suggest the following general approach for future research. Avoid deflation by scale, unless it is called for by theory. If links between theory and the need to scale predictability and disagreement are unclear, report results for both unscaled and scaled versions, but include price and inverse of price, respectively, as an additional regressor (see, for example, Barth and Kallapur [1996] for similar suggestions).

We emphasize two proximate areas of prior research that this paper does not cover. First, since we study forecast error variability, not the *level* of forecast errors, we do not investigate whether our suggestions regarding price-deflation apply to research based on the level of forecast errors.² Second, although we investigate the impact of price-deflation for predictability and disagreement, we do not offer general recommendations such as those provided by the literature on scaling (e.g., Kothari and Zimmerman [1995], and Barth and Kallapur [1996]). Our paper is closer to Durtschi and Easton [2005], which shows that inferences from the distribution of EPS and forecast errors depend on whether or not those variables are scaled by lagged price.

² Deflation of forecast errors, rather than forecast error variability, is often undertaken to mitigate heteroskedastic errors (see for example, Kothari [2001, p. 154]). Our results suggest that such deflation potentially increases rather than mitigates heteroskedasticity, which results in biased standard errors.

We note that scaling generally refers to deflating firm-level data, not the share-level data considered in this paper. Share-level data can be viewed as firm-level data that have already been scaled by number of shares. But number of shares is an odd deflator to use since it does not completely deflate for scale at the firm level, and it is also an endogenous choice variable that can be arbitrarily increased or decreased by splitting or reverse splitting shares.³ In the specific area of analyst forecast research, however, there are reasons to view share-level rather than firm-level data as the original undeflated data and then consider the impact of deflating by share-level measures of scale, such as share price. Analysts appear to focus their earnings forecasts at the share level⁴. The investment community and popular financial press emphasize earnings at the per share level; e.g., reported EPS are described as beating share-level forecasts by a penny or being a penny short. Finally, academic research typically studies forecast-related variables at the share level, rather than at the firm level.

The remainder of this paper is organized as follows. The sample and descriptive statistics are described in Section 2, and Section 3 contains our main findings and a summary of robustness tests. Section 4 discusses the results of extending Thomas [2002] to illustrate the potential for biases when predictability and disagreement are deflated by price. Section 5 concludes.

³ Our results suggest that firm-level forecast error variability and dispersion increase with size (e.g., market capitalization, which equals share price * number of shares), but not proportionately so. Scaling firm-level variables by size causes forecast error variability and dispersion to decrease with size, similar to the patterns observed in the right blocks of Panels A and C in Figure 1.

⁴ Figure 1 of Herrmann and Thomas [2005] suggests that analysts exhibit a strong tendency to round EPS forecasts since the frequency of forecasts at multiples of nickels (e.g., 5 cents, 10 cents, etc.) is on average about five times that for forecasts that are not multiples of nickels. This pattern would not be observed if analysts forecast firm-level data and then computed per-share forecasts mechanically, by dividing by number of shares.

2. Sample and descriptive statistics

2.1. Sample selection

We include all U.S. firms on the unadjusted I/B/E/S files (with data not adjusted for stock splits) that have fiscal quarters ending in the 14 calendar years from 1993 to 2006. We drop years before 1993 because of concerns about a shift around the early 1990's in the methodology used to compute "actual" EPS as reported by I/B/E/S, which includes adjustment by I/B/E/S for items analysts did not forecast.⁵ We exclude Canadian firms, and American Depository Receipts (ADRs) from our sample. For each firm-quarter included we obtain the actual quarterly EPS (*IBESACTUAL*), the most recent consensus (mean) forecast EPS (*FORECAST*) and the standard deviation of individual forecasts around that consensus (*DISPERSION*).⁶ We also obtain other variables such as the number of analysts issuing forecasts (*COVERAGE*), and the age of individual forecasts as of the date of the consensus forecast (*STALENESS*). To increase the reliability of consensus forecasts, we delete firm-quarters for which the consensus is based on fewer than three analysts. Finally, we require that stock price (*BEGPRICE*) is available on I/B/E/S as of January before the fiscal quarter end.

Our "full sample" that satisfies these requirements contains 142,708 firm-quarters. We then sort firm-quarters into price deciles each calendar year, based on the beginning-of-year price (*BEGPRICE*). For our supplementary analyses, we create additional variables, derived from Compustat and CRSP. Details of all variables are provided in Appendix A. No variables have

⁵ Cohen et al. [2007, p. 272] state that "prior to the early 1990s, I/B/E/S did not always adjust actual earnings to exclude items not forecasted by analysts, thereby creating a mismatch between its actual (realized) and forecasted (expected) earnings." We find, however, that our main findings remain unchanged when we include data from before 1993.

⁶ The most recent forecast is typically from the same month as the month of earnings announcement, or the prior month if the earnings announcement has already been made before I/B/E/S' cut off date for that month. In a few cases, we go back up to 90 days before the earnings announcement to find an available consensus forecast.

been Winsorized or truncated.⁷ Since our robustness investigation includes a comparison of the predictability and disagreement distributions across industry sectors, we excluded the “Miscellaneous/Undesignated” sector since it had only a handful of firm-quarters.

Investigation of time-series variation in sample size suggests a general increase in the number of observations in each year through time, though there is a temporary decline in the years 1999 to 2003. Investigation of across-sector variation suggests quite some variation across sectors, although all sectors include a reasonable number of firm-quarters in each sector: Technology has the most observations (28,005) while Transportation has the fewest (3,397).

2.2. Summary statistics for sample

Table 1 presents descriptive statistics for several key variables, sorted in alphabetical order. Most results are consistent with prior research.⁸ *FORECAST* distributions in Panel A are similar to those for *IBESACTUAL*, suggesting that consensus forecasts are reasonably accurate. That inference is confirmed by the relatively tight interquartile range (from –1 cent to +3 cents) for forecast errors (*FORECASTERR*) and the relatively small values for absolute values of forecast error (*ABSFE*). The distribution for actual EPS according to Compustat (*COMPACTUAL*) is similar to that for *IBESACTUAL*, except that it is more left-skewed, and the corresponding forecast error (*COMPFE*) distributions are similar to those based on *IBESACTUAL*. Disagreement across analysts, measured by the dispersion of individual forecasts around consensus forecasts (*DISPERSION*) just prior to earnings report dates, is fairly narrow, indicated by mean (median) *DISPERSION* of 3 (2) cents per share.

⁷ One firm, Berkshire Hathaway (I/B/E/S ticker is BKHT), is deleted from our sample because it had an unusually large forecast error for the quarter ending December 2006 (the forecast error of \$406.64 per share arises from an *IBESACTUAL* of \$1859 versus a *FORECAST* of \$1452.36). This error is so large that it skews some of our descriptive statistics (the next highest forecast error in our sample is below \$44).

⁸ For example, Abarbanell and Lehavy [2003] also find that the mean forecast error is lower than the median forecast error and that the fraction of positive forecast errors exceeds the fraction of negative forecast errors.

The mean and median share prices (*BEGPRICE*) for our sample are \$27.1 and \$22.5, implying that the distribution of share prices is right-skewed. The mean number of analysts covering stocks in our sample (*COVERAGE*) is about 7 and the mean and standard deviation of the age of forecasts as of the consensus date is captured by *MEANSTALE* and *SDSTALE*, respectively. We deflate forecast error, absolute forecast errors, and dispersion by share price to generate *DEFLFE*, *DEFLABSFE*, and *DEFLDISP*, respectively, which are used later to investigate the impact of price deflation by comparing them with *INVBEGPRC*, which is the inverse of beginning share price. *VOL* is a measure of fundamental uncertainty, derived from the standard deviation of daily returns over a prior 200 day window.

Panel B of Table 1 provides Pearson and Spearman correlations between different pairs of these variables. We limit our attention at this stage to a few correlations with share price, and consider some other correlations later. The level of forecasts (*FORECAST*) and actual earnings (*IBESACTUAL* and *COMPACTUAL*) are strongly positively related to *BEGPRICE*. And yet variability of forecast error, as captured by *ABSFE*, is only weakly positively related to *BEGPRICE*. Similarly, the dispersion of forecasts is also weakly related to share price.⁹ More details regarding the variation of *ABSFE* and *DISPERSION* with *BEGPRICE* are explored in the next section. Because *ABSFE* and *DISPERSION* are only weakly related to price, deflating them by price creates a strong negative (positive) relation with *BEGPRICE* (*INVBEGPRC*).¹⁰

Panel C of Table 1 provides the means and medians for selected variables across different price deciles. The results show considerable variation in share price around the overall mean/median that is reported in the right-most column for purposes of comparison. Mean and

⁹ The relatively high Pearson correlation of 0.15 between *DISPERSION* and *BEGPRICE* appears to be due to extreme values because it declines to 0.11 when we Winsorize the extreme 1 percent of the distributions.

¹⁰ As mentioned in the prior footnote, the relatively low negative Pearson correlations observed for *DEFLABSFE* and *DEFLDISP* appear to be related to extreme values, since they increase to values close to the Spearman correlations when we Winsorize the extreme 1 percent of observations.

median values of *BEGPRICE* for the highest decile are well over ten times those for the lowest decile. This variation in the scale of share price is mirrored in corresponding variation in the magnitudes of consensus EPS forecasts and both measures of actual EPS, since the means and medians reported for *FORECAST*, *IBESACTUAL*, *COMPACTUAL* for decile 10 are over ten times those reported for decile 1.

3. Main findings

As described in Section 1, we believe that the common practice of deflating variability of forecast errors and forecast dispersion by price or level of actual/forecast earnings is based on the intuition that the ability of consensus forecasts to predict reported EPS and the disagreement across analysts when forecasting EPS for a particular firm-quarter both vary with scale. It appears that price deflation is preferred over deflation by levels of actual or forecast earnings because of the potential for distortion when those levels are close to zero or negative. Regardless of which scaling variable is used, common wisdom is that deflation should improve comparability across shares of different scale.

3.1. Evidence of how predictability and disagreement vary with price

Figure 1 provides a more comprehensive perspective on the relation between prices and variability of forecast errors/dispersion than that provided by the correlations in Panel B of Table 1. Each vertical bar represents the distribution for a price decile, and the different marks identify the location of the mean, median, and 5th, 25th, 75th, and 95th percentiles of the pooled forecast error distribution. Numerical values for the mean, median, standard deviation and interquartile ranges for these distributions are provided in the corresponding panels of Table 2.

The left block of Figure 1, Panel A describes the distribution of forecast errors (*FORECASTERR*) for different price deciles, where forecast errors are measured relative to the actual EPS as reported by I/B/E/S. There is a concern that this proxy for the “core” earnings

number that analysts attempt to forecast may be biased in unexpected ways since I/B/E/S adjusts it after observing the price reaction to announced earnings.¹¹ To alleviate those concerns, the middle block of Panel A describes the distribution of *COMPFE*, forecast errors measured relative to actual EPS as reported by Compustat. The left blocks in Panels B and C provide the distributions for absolute forecast errors (*ABSFE*), an alternative measure of forecast error variability, and dispersion in analyst forecasts around the consensus for each firm-quarter (*DISPERSION*). The right blocks in Panels A, B, and C provide the distributions for *DEFLFE*, *DEFLABSFE* and *DEFLDISP*, which are price-deflated values of forecast errors, absolute forecast errors, and dispersion, respectively.

The main finding from the left block of Figure 1, Panel A, and Panel A1 of Table 2 is that variability of forecast error distributions does not increase substantially with share price. The spread between the 5th and 95th percentiles, the interquartile range, and the standard deviation all suggest a shallow U-shaped relation between the variability of forecast error and share price, with the right end of the U (firms with higher priced shares) being slightly taller than the left end of the U (firms with lower priced shares). To illustrate the surprising absence of scale implied by these results, consider for example the relative lack of variation in the interquartile range across the price deciles. Even though firms in decile 1 are on average considerably smaller in scale than firms in decile 10 (in terms of price and actual and forecast EPS), the interquartile range of 5 cents for decile 1 is quite similar to the 5 cents reported for decile 10. In essence, holding aside systematic variation in forecast biases that are captured by differences in the mean/median forecast error across the ten share price deciles, consensus forecasts are almost equally accurate

¹¹ The Wharton Research Data Services (Glushkov [2007, p. 27]) provides the following description: “IBES observes the market reaction to the earnings announcement prior to choosing exactly which earnings components to include in street earnings. This leads to a potential ex post selection bias.” Bradshaw and Sloan [2002, p. 42] define street earnings as the “numbers announced by corporations in their press releases and tracked by analyst estimate clearinghouse services, such as I/B/E/S.”

regardless of whether the EPS being forecasted is only a few cents (for firms in decile 1) or almost a dollar (for firms in decile 10).

The results in the middle block of Figure 1, Panel A, and Panel A2 of Table 2 confirm that the observed lack of scale exhibited by predictability is not sensitive to whether forecast errors are computed using actual EPS according to I/B/E/S or Compustat. While the measures of variability for *COMPFE* in the middle block of Figure 1, Panel A, and Panel A2 of Table 2 are systematically higher than those for *FORECASTERR*, variation across the price deciles continues to be described by a shallow U-shaped relation, rather than a sharply rising one.

The results reported for absolute forecast errors in the left block of Panel B in Figure 1 and Panel B1 of Table 2 confirm the first finding, since magnitudes of forecast errors do not increase much with scale. Mean and median levels of *ABSFE* for deciles 1 (10) are 0.07 (0.09) and 0.02 (0.03), respectively. Note that absolute values overstate true variability when the means/medians are not zero. And since the mean/median forecast errors in Panel A1 of Table 2 indicate a systematic pattern of more negative (positive) bias as one moves from the middle price deciles toward lower (higher) price deciles, the degree of variability overstatement increases for more extreme price deciles. As a result, we prefer to describe predictability in terms of measures of variability of *FORECASTERR*, such as the standard deviation and interquartile range, rather than mean/median values of *ABSFE*.

Our second finding regarding the lack of scale associated with analyst disagreement is described in the left block of Panel C in Figure 1 and Panel C1 of Table 2. As with *ABSFE*, the focus is not on the spreads of these distributions, but on the means and medians, since the variable (*DISPERSION*) already measures spread across individual forecasts. As with variability of forecast errors, the mean/median level of dispersion exhibits a shallow, asymmetric U-shaped

relation, that is taller for high price deciles, rather than the proportional relation expected in prior research. This counterintuitive finding suggests that disagreement across analysts, measured in cents per share, does not vary much across the price deciles even though the level of forecasted EPS varies substantially.

The impact of the common practice of price deflation on these variables is described by the right block in Panels A, B, and C of Figure 1 and Panels A3, B2, and C2 of Table 2. Examining measures of variability for *DEFLFE* and means/medians for *DEFLABSFE* and *DEFLDISP* suggests that price deflation causes variability of forecast errors and forecast dispersion to decline sharply with price. Given the very mild evidence of a positive relation between share price and undeflated variability of forecast errors and forecast dispersion, scaling by price reverses that mild positive relation and creates a strong negative relation with share price. To be more accurate, price deflation induces a strong positive relation with the inverse of share price.

3.2. Are the findings robust?

Panels A and B of Figure 2 offer a more detailed look at the distributions of forecast error and dispersion, respectively, to determine whether the distributional statistics reported in Figure 1 mask some unusual aspects. The histograms reported show the fraction of the sample represented by each cent of forecast error and dispersion. For brevity, we only report histograms for three price deciles: deciles 1, 5, and 10, representing low, medium, and high share price firms, respectively. Scrutiny of these histograms reveals interesting patterns, such as a) the frequency of large negative forecast errors (less than – 30 cents per share) is high for both low- and high-price shares, but low for medium-price shares, b) the frequency of large positive forecast errors (greater than 30 cents per share) is high only for high-price shares, consistent with right-skewness observed in Figure 1, Panel A, c) the fraction of observations in the “just missed”

category (forecast errors of -1 and -2 cents) is substantially lower for high-price shares, and d) the fraction of observations with dispersions of 0 and 1 cents decreases with share price, whereas the fraction with dispersion of 2 cents or more increases slightly with share price.

While these histograms suggest a variety of interesting differences related to share price, especially regarding middle and tail asymmetries in the forecast error distributions, the relevant conclusion is that we see no evidence to contradict our inferences from statistics reported in Figure 1 and Table 2 regarding a systematic lack of scale for predictability and disagreement.¹² It is not the case, for example, that observed lack of scale for predictability is driven by most observations having zero forecast error, where actual EPS exactly meets the consensus forecast.

We repeated the *FORECASTERR* and *DISPERSION* plots in Figure 1 for each year in our sample period. In addition, we computed the following statistics for each price decile: a) standard deviation/interquartile range for forecast error, and b) mean/median dispersion of each price decile. Our results (untabulated) confirm that the full sample findings regarding predictability and disagreement are observed in most years.

We conducted a similar analysis across each of the 11 sectors. There are interesting patterns in the levels of predictability and disagreement in different sectors. For example, variability of forecast errors (predictability) and mean/median levels of dispersion (disagreement) are considerably lower in the health care and technology sectors, but considerably higher in the transport and utilities sectors. However, all sectors reflect the same general patterns of lack of variation in predictability and disagreement across price deciles that we noted in the full sample.

¹² Abarbanell and Lehavy [2003, p. 106] define [left] tail asymmetry as “a larger number and a greater magnitude of observations that fall in the extreme negative relative to the extreme positive tail of the forecast error distributions” and middle asymmetry as “a higher incidence of small positive relative to small negative forecast errors in cross-sectional distributions.”

Since information regarding the quarter's operations is known to management soon after the quarter end, this knowledge may affect the accuracy and dispersion we study, since the most recent forecasts before the earnings announcement that we consider here are collected typically after the quarter-end. To investigate any potential effects of information available to managers, we repeated the analysis for forecasts as of six and nine months before the quarter-end. Again, we find very little variation with scale for forecast error variability and forecast dispersion at both the six and nine-month horizons.

We also confirm that our findings remain qualitatively unchanged when we a) use the median of the individual forecasts each quarter, instead of the mean, to represent the consensus forecast, and b) use absolute values of forecast earnings and per share level of total assets as alternative measures of scale, instead of share price.

4. Replication of Thomas [2002] to illustrate potential biases caused by price deflation

Until we obtain a better understanding of the factors that explain how predictability and disagreement vary with scale, researchers investigating these attributes should exercise caution when deflating by share price. If predictability and disagreement do not vary naturally with scale, price deflation will bias coefficient estimates, as long as other included variables happen to be correlated with share price. As described in Appendix B, there are a number of studies that use predictability and disagreement as dependent or independent variables, and both variables are deflated in the primary analyses in all the studies we investigated.¹³

To illustrate these issues we extend the regressions of price-deflated predictability and dispersion on diversification reported in Thomas [2002], a study that investigates the relation between diversification and information asymmetry between managers and investors. Price-

¹³ In some cases, footnotes indicated that similar results were obtained with undeflated proxies for these two variables; examples include Barron et al. [1999, footnote 13] and Barron [1995, footnote 13].

deflated measures of predictability and disagreement are two of many attributes of information asymmetry considered in that study, and diversification is measured by the Herfindahl Index (*HERF*) computed for each firm-year based on segment assets. Our objectives are to determine the extent to which the results of that study are affected when a) inverse of price is included as an additional regressor, b) predictability and disagreement are not deflated by price, and c) price is added as an additional regressor to the undeflated specifications.¹⁴

Panel A of Table 3 contains Pearson and Spearman correlations among pairs of key variables from Thomas [2002] as well as other variables we created from the underlying data. To avoid confusion with similar variables used earlier in this paper, we choose our own labels for these variables. The dependent variables in the regressions estimated in Thomas [2002] are labeled *DEFLATAFE* and *DEFLATDISP*, which are price-deflated values of absolute forecast errors and forecast dispersion, where deflation is based on share price five days before the annual earnings announcement (*PRICE5*). We focus here only on two of the regressors, *HERF* and *RESIDVOL* considered in the different equations. *HERF*, which measures diversification, varies between 0 and 1, with lower values representing greater diversification across different segments. *RESIDVOL*, which measures the standard deviation of market model residuals, is included in the final specification in Thomas [2002] to control for potential relations between idiosyncratic volatility and predictability/dispersion. The variables we introduce are undeflated absolute

¹⁴ Our objective is not to question the final conclusions in Thomas [2002], but to illustrate how estimated coefficients are affected by these three extensions. The conclusions reached in Thomas [2002] are ultimately supported in analysis conducted on alternative measures of asymmetric information that are not subject to the scaling issues investigated in the present paper, i.e., abnormal returns to seasoned equity offerings (Hadlock et al. [2001]) and market microstructure metrics (Clarke et al. [2004]).

forecast errors (*AFE*), dispersion (*DISP*), and the inverse of share price (*INVPRICE5*).¹⁵ Key correlations are introduced where relevant in the discussion below.

Panels B and C of Table 3 contain the results of extending the analyses in Tables 3 and 4 of Thomas [2002], which explain variation in price-scaled absolute forecast error and dispersion, respectively. Specification I refers to the original results and equations (1) through (5), reported in the columns, refer to the corresponding equations estimated in Thomas [2002]. The main finding from the results for specification I in both panels that is relevant for our purposes is that the coefficient on *HERF* is positive and significant in equations (1) through (4), but that relation switches to a negative and significant coefficient in equation (5), when volatility is introduced. That is, lower diversification (larger *HERF*) is associated with higher variability of forecast errors and forecast dispersion, but that relation reverses when a control for idiosyncratic volatility is introduced in equation (5). Recall that the dependent variables in both panels are deflated by share price.

Specification II considers the impact of introducing the inverse of share price as an additional regressor. This extension would be appropriate if theory called for measures of predictability and disagreement to be scaled by share price, but there remained a concern whether that deflation might induce a spurious correlation with variables that are related to price. Panel A of Table 3 indicates that price-deflated absolute forecast errors and dispersion are strongly positively related to the inverse of price. Introducing the inverse of share price offers a simple way to mitigate such a concern. The main result in specification II for both panels B and C is that including *INVPRICE5* to the right hand side eliminates all of the significant positive

¹⁵ *AFE* is similar to *ABSFE* except that the consensus forecast is the median not the mean forecast for each firm-quarter, *PRICE5* the share price used for deflation is similar to *BEGPRICE* except that it is based on share price five days before the earnings announcement rather than beginning-of-quarter share price, and *RESIDVOL* is similar to *VOL* except that it focuses on idiosyncratic not total return volatility.

coefficients on *HERF* observed in the original results for equations (1) through (4). These results can be anticipated by the negative correlation between *HERF* and price in Panel A of Table 3. And the lower coefficient on *RESIDVOL*, relative to that in specification I, is likely related to the positive correlation between volatility and inverse of price.

Specification III is similar to the original specification, but the dependent variables are no longer deflated by price. As with specification II, no significant positive coefficients are observed on *HERF* in either Panel B or C. These results suggest that the significant positive coefficients observed on *HERF* for equations (1) through (4) in the original specification are likely due to the negative correlation between *HERF* and share price, which then induces a positive correlation between *HERF* and the price-deflated dependent variables. Introducing a variable that is related to share price, such as *RESIDVOL* in equation (5), as an additional regressor controls for this correlation between *HERF* and the price-deflated variables.

Specification IV adds share price as an additional regressor to specification III to control for the small positive relations observed between share price and undeflated measures of predictability and disagreement (caused by the right end of the U-shaped relation, for firms with high-price shares, being slightly taller than the left end). Panel A of Table 3 confirms that *AFE* and *DISP* are positively related to share price. Observing a positive coefficient on *PRICE5*, that is especially significant in Panel C, illustrates the importance of controlling for the small residual positive relation with share price that is observed for undeflated measures of predictability and disagreement.

In sum, the results generated by extending the analyses in Thomas [2002] suggest the following implications for research that employs measures of predictability and disagreement. First, unless called for by theory, these measures should either not be deflated or both sets of

results based on deflated and undeflated measures should be reported. Second, if deflated measures are used, it is important to include the inverse of price as an additional regressor, to confirm that the coefficients are not biased because of the strong negative relation between deflated measures and share price. Third, even if undeflated measures are used, it is important to include price as an additional regressor, to mitigate any bias due to the small positive relation between undeflated measures and share price.

5. Concluding Remarks

In this paper, we document surprising empirical findings related to two aspects of analyst forecasts: a) predictability of reported EPS, measured by variability of forecast errors, and b) disagreement across analysts' EPS forecasts, measured by dispersion of forecasts around the consensus forecast. Prior research has relied on the intuition that predictability and disagreement should vary proportionately with scale, typically proxied by share price. We agree that this is a reasonable presumption since levels of both actual and forecast EPS, which are the variables underlying predictability and disagreement, vary with scale. However, contrary to these expectations, we find that measures of variability of forecast errors as well as dispersion of individual forecasts around the consensus do not vary much with share price.

We confirm that this observed lack of variation with scale for predictability and disagreement is a robust finding. It is observed in each sample year and in each industry sector. Investigation of the histograms for the two distributions confirms that the lack of variation with scale is not due to quirks in the data.

While we do not investigate here potential reasons for our surprising findings, we investigate an important implication of those findings. Prior research has intuitively scaled predictability and disagreement by measures of scale, such as share price. Given that predictability and disagreement are relatively homogeneous across big and small shares,

deflating by scale creates a strong negative correlation with scale. As a result, coefficient estimates are potentially biased when predictability and disagreement appear as dependent (independent) variables and independent (dependent) variables happen to be correlated with scale. Overall, we believe that price deflation of these two variables be undertaken with caution. Researchers are encouraged to a) check whether the coefficients on variables of interest are robust to the use of deflated and undeflated measures of predictability/disagreement, and b) include price or inverse of price as an additional regressor where appropriate.

Appendix A
Variable definitions and sources
(Quarterly Compustat data items are provided in parentheses under Description)

Label	Description	Source
<i>ABSFE</i> (in dollars)	Absolute value of <i>FORECASTERR</i>	
<i>BEGPRICE</i> (in dollars)	Share price of firm at the beginning of year.	I/B/E/S Summary Actuals + Pricing unadjusted file (WRDS file name is ibes.actpsumu).
<i>COMPACTUAL</i> (in dollars)	Actual quarterly basic EPS as reported by Compustat, excluding extraordinary income and the after-tax effect of special items. $COMPACTUAL = [\#EPSPXQ - (1 - \text{tax}) * (\#SPIQ) / (\#CSHPRQ)] / \text{DilutionFactor}$. Scaling by DilutionFactor is necessary as <i>FORECAST</i> can be on a basic or diluted basis.	CRSP/COMPUSTAT Merged - Combined Industrial Quarterly file (WRDS filename is comp.fundq). ^a Tax rate is obtained from IRS: http://www.irs.gov/pub/irs-soi/02corate.pdf . DilutionFactor is obtained from I/B/E/S (WRDS filename is ibes.idsum).
<i>COMPFE</i> (in dollars)	$COMPACTUAL - FORECAST$	
<i>COVERAGE</i> (unit-free)	Number of estimates that constitute <i>FORECAST</i> .	I/B/E/S Unadjusted Summary Data (WRDS filename is ibes.statsumu).
<i>DEFLABSFE</i>	$ABSFE / BEGPRICE$	
<i>DEFLDISP</i> (unit-free)	$DISPERSION / BEGPRICE$	
<i>DEFLFE</i> (unit-free)	$FORECASTERR / BEGPRICE$	
<i>DISPERSION</i> (in dollars)	Standard deviation of the individual analyst's EPS forecast that constitute <i>FORECAST</i> .	I/B/E/S Unadjusted Summary Data (WRDS filename is ibes.statsumu).
<i>FORECAST</i> (in dollars)	Most recent consensus (mean) estimate of <i>IBESACTUAL</i> for the firm-quarter.	I/B/E/S Unadjusted Summary Data (WRDS filename is ibes.statsumu).
<i>FORECASTERR</i> (in dollars)	$IBESACTUAL - FORECAST$	
<i>IBESACTUAL</i> (in dollars)	Actual quarterly EPS as reported by I/B/E/S, after I/B/E/S has adjusted it "for comparability with estimates."	I/B/E/S Unadjusted Actuals Data (WRDS filename is ibes.actu).

<i>INVBEGPRC</i> (in 1/dollar)	Inverse of <i>BEGPRICE</i>	
<i>MEANSTALE</i> (in days)	The <i>mean</i> forecast age of (<i>effective</i>) individual forecast, measured between the issue date of the individual forecast and the date of the consensus forecast. What constitutes “ <i>effective</i> ” is explained in http://wrds.wharton.upenn.edu/ds/ibes/lib/IBES_Summary_from_Detail.pdf	I/B/E/S Unadjusted Detail Data (WRDS filename is ibes.detu).
<i>SDSTALE</i> (in days)	The standard deviation of (effective) individual forecast age. See description for <i>MEANSTALE</i> .	I/B/E/S Unadjusted Detail Data (WRDS filename is ibes.detu).
<i>VOL</i>	Standard deviation of stock returns over the period from day -210 to -11, relative to the fiscal quarter-end.	CRSP daily file (WRDS file name is crsp.dsf).

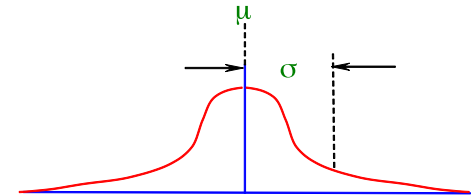
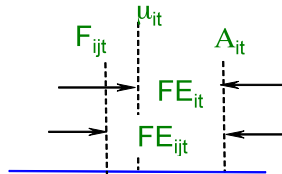
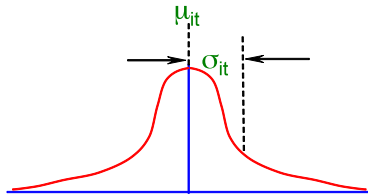
Appendix B

Each analyst j makes an eps forecast F_{ijt} for firm i in quarter t . That **consensus** forecast, μ_{it} , and which gives rise to the forecast distribution below.

The mean, μ_{it} , is the **consensus** forecast.

The standard deviation, σ_{it} , is referred to as **dispersion**, and measures **disagreement** across analysts.

Those forecast errors FE_{it} are pooled across firms and quarters to generate the forecast error distribution below. The mean, μ , is the **bias** in consensus forecasts. The standard deviation, σ , measures **predictability** of actual eps.



FIRM-QUARTER LEVEL

POOLED ACROSS FIRM-QUARTERS

VARIABLE

	Disagreement (σ_{it})	Predictability (σ or $\text{abs} FE_{it} $)
Dependent Variable	Case A	Case C
Independent Variable	Case B	Case D

To link our variables with the prior literature, we partition the prior literature into four cases based on whether the variable of interest is disagreement or predictability and whether that variable appears as a dependent or independent variable. The specific papers provided below are intended to be illustrative, and are not derived from an exhaustive search.

For Case A, we consider analyses where dispersion or disagreement among analysts (σ_{it}) is the dependent variable. As an example, Lang and Lundholm [1996, Table 6] investigate whether a firm's corporate disclosure policy affects its price-deflated dispersion. Other studies in this category include Barron et al. [1999, Table 5], Hope [2003, Table 4 and 5], Mozes [2003, Table 4], and Thomas [2002, Table 4].

For Case B, we consider analyses where dispersion is the independent variable. As an example, Zhang [2006a, Table III; 2006b, p.570] investigates the effect of information uncertainty (proxied using price-deflated dispersion) on stock returns. Other studies in this category include Ajinkya et al. [1991, Table 3 and p.393], Baber and Kang [2002, Table 4], Bamber et al. [1997, Table 3], Barron [1995, Table 3], Diether et al. [2002, Table II], Gu and Wu [2003, Table 3 and p.13], Imhoff and Lobo [1992, Table 4 and p.431], and Loh and Mian [2006, Table 7].

For Case C, we consider analyses where the predictability of EPS (σ or $|FE_{it}|$) is the dependent variable. As an example, Hope [2003, Table 4 and 5] investigates whether higher accounting disclosure results in greater predictability (measured using absolute value of price-deflated forecast error). Other studies in this category include Duru and Reeb [2002, Eq. 1 to 4], Haw et al. [1994, Eq 1 and Table 2], and Stickel [1993, Exhibit 2 and 4].

For Case D, we consider pooled analyses where the predictability of EPS is the independent variable. We did not find any relevant papers.

References

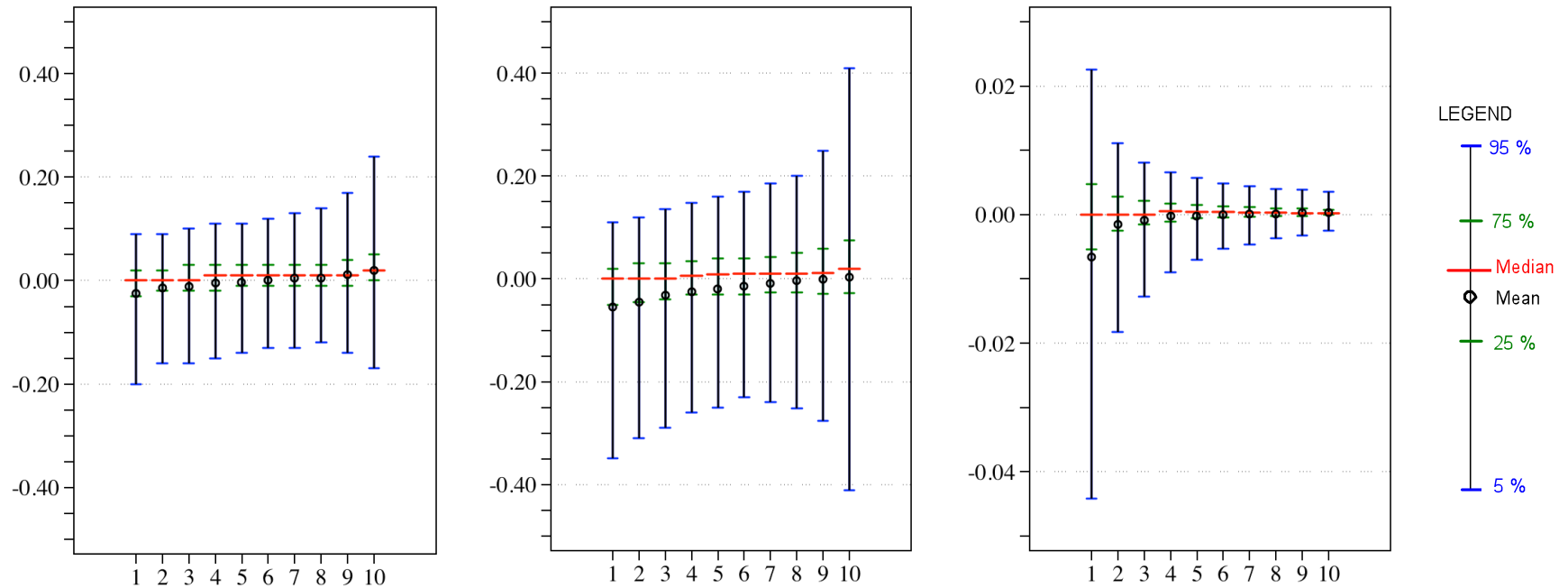
- ABARBANELL, J., AND R. LEHAVY. "Biased forecasts or biased earnings? The role of earnings management in explaining apparent optimism and inefficiency in analysts' earnings forecasts," *Journal of Accounting and Economics* 36 (2003): 105–146.
- AJINKYA, B.; R. ATIASE; AND M. GIFT. "Volume of trading and the dispersion in financial analysts' earnings forecasts," *The Accounting Review* 66 (1991): 389–401.
- BABER, W., AND S. KANG. "The Impact of Split Adjusting and Rounding on Analysts' Forecast Error Calculations," *Accounting Horizons* 16 (2002): 277–290.
- BAMBER, L.; O. BARRON; AND T. STOBBER. "Trading Volume and Different Aspects of Disagreement Coincident with Earnings Announcements," *The Accounting Review* 72 (1997): 575–598.
- BARTH, M., AND S. KALLAPUR. "The effects of cross-sectional scale differences on regression results in empirical accounting research," *Contemporary Accounting Research* 13 (1996): 527–567.
- BARRON, O. "Trading Volume and Belief Revisions that Differ Among Individual Analysts," *The Accounting Review* 70 (1995): 581–581.
- BARRON, O.; C. KILE; AND T. O'KEEFE. "MD&A Quality as Measured by the SEC and Analysts' Earnings Forecasts," *Contemporary Accounting Research* 16 (1999): 75–110.
- BARRON, O.; O. KIM; S. LIM; AND D. STEVENS. "Using Analysts' Forecasts to Measure Properties of Analysts' Information Environment," *The Accounting Review* 73 (1998): 421–433.
- BRADSHAW, M., AND R. SLOAN. "GAAP versus The Street: An Empirical Assessment of Two Alternative Definitions of Earnings," *Journal of Accounting Research* 40 (2002): 41–66.
- CLARKE, J.; C. FEE; AND S. THOMAS. "Corporate Diversification and Asymmetric Information: Evidence from Stock Market Trading Characteristics," *Journal of Corporate Finance* 10 (2004): 105–129.
- COHEN, D.; R. HANN; AND M. OGNEVA. "Another look at GAAP versus the Street: an empirical assessment of measurement error bias," *Review of Accounting Studies* 12 (2007): 271–303.
- DIETHER, K.; C. MALLOY; AND A. SCHERBINA. "Differences of Opinion and the Cross Section of Stock Returns," *The Journal of Finance* 57 (2002): 2113–2141.
- DURTSCHI, C., AND P. EASTON. "Earnings Management? The Shapes of the Frequency Distributions of Earnings Metrics Are Not Evidence Ipso Facto," *Journal of Accounting Research* 43 (2005): 557–592.
- DURU, A., AND D. REEB. "International Diversification and Analysts' Forecast Accuracy and Bias," *The Accounting Review* 77 (2002): 415–433.
- GLUSHKOV, D. "Working with Analyst Data: Overview and Empirical Issues," Unpublished paper, Wharton Research Data Services, 2007. Available at http://wrds.wharton.upenn.edu/news/sideitem/user2007/analyst_data.pdf

- GU, Z., AND J. WU. "Earnings skewness and analyst forecast bias," *Journal of Accounting and Economics* 35 (2003): 5–29.
- HADLOCK, C.; M. RYNGAERT; AND S. THOMAS. "Corporate Structure and Equity Offerings: Are There Benefits to Diversification?" *Journal of Business* 74 (2001): 613–635.
- HERRMANN D., AND W. THOMAS. "Rounding of Analyst Forecasts," *The Accounting Review* 80 (2005): 805–823.
- HAW, I.; K. JUNG; AND W. RULAND. "The Accuracy of Financial Analysts' Forecasts after Mergers," *Journal of Accounting Auditing and Finance* 9 (1994): 465–465.
- HOPE, O. "Accounting Policy Disclosures and Analysts' Forecasts," *Contemporary Accounting Research* 20 (2003): 295–321.
- IMHOFF, E., AND G. LOBO. "The Effect of Ex Ante Earnings Uncertainty on Earnings Response Coefficients," *The Accounting Review* 67 (1992): 427–439.
- KOTHARI, S. "Capital markets research in accounting," *Journal of Accounting and Economics* 31 (2001): 105–231.
- KOTHARI, S., AND J. ZIMMERMAN. "Price and return models," *Journal of Accounting and Economics* 20 (1995): 155–192.
- LANG, M., AND R. LUNDHOLM. "Corporate disclosure policy and analyst behavior," *The Accounting Review* 71 (1996): 467–492.
- LOH, R., AND G. MIAN. "Do accurate earnings forecasts facilitate superior investment recommendations?" *Journal of Financial Economics* 80 (2006): 455–483.
- MOZES, H. "Accuracy, usefulness and the evaluation of analysts' forecasts," *International Journal of Forecasting* 19 (2003): 417–434.
- STICKEL, S. "Accuracy improvements from a consensus of updated individual analyst earnings forecasts," *International Journal of Forecasting* 9 (1993): 345–353.
- THOMAS, S. "Firm diversification and asymmetric information: evidence from analysts' forecasts and earnings announcements," *Journal of Financial Economics* 64 (2002): 373–396.
- WHITE, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48 (1980): 817–838.
- ZHANG, X. "Information Uncertainty and Stock Returns," *The Journal of Finance* 61 (2006a): 105–137.
- ZHANG, X. "Information Uncertainty and Analyst Forecast Behavior," *Contemporary Accounting Research* 23 (2006b): 565–590.

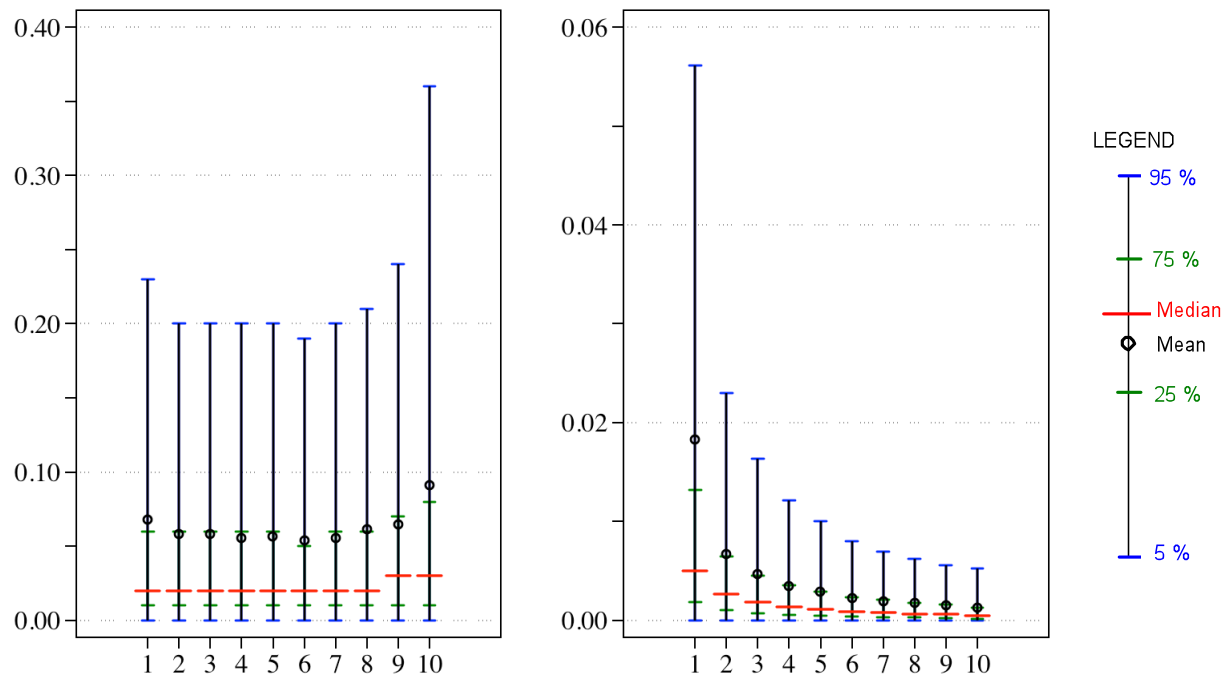
Figure 1. Distribution of forecast error and dispersion for different *BEGPRICE* deciles

The plots below describe key distributional statistics for measures of forecast error and forecast dispersion for different deciles of *BEGPRICE*, which is the beginning-of-year share price (in dollars). The price deciles are computed each calendar year, and the lowest (highest) price decile is denoted by 1 (10). The mean is indicated by the solid circle, the median by the long horizontal hash mark, and the remaining hash marks locate the 5th, 25th, 75th, and 95th percentiles of the pooled distributions for the different variables. *FORECASTERR* is defined as *IBESACTUAL* minus *FORECAST*, where *IBESACTUAL* is the actual quarterly EPS (in dollars) as reported by I/B/E/S, and *FORECAST* is the most recent consensus (mean) EPS forecast (in dollars) for that firm-quarter. *COMPFE* equals *FORECAST* minus *COMPACTUAL*, where *COMPACTUAL* is the actual quarterly EPS (in dollars) as reported by Compustat. *ABSFE* is the absolute value of *FORECASTERR*. *DISPERSION* is the standard deviation of the *individual* analysts' EPS forecasts around the consensus in that firm-quarter. *DEFLFE*, *DEFLABSFE*, and *DEFLDISP* are defined as *FORECASTERR*, *ABSFE*, and *DISPERSION* scaled by the beginning-of-year share price (*BEGPRICE*), respectively. All variables relate to firm-quarters, and are described in more detail in Appendix A.

Panel A: Distribution of *FORECASTERR*, *COMPFE*, and *DEFLFE* in each *BEGPRICE* decile



Panel B: Distribution of *ABSFE* and *DEFLABSFE* in each *BEGPRICE* decile



Panel C: Distribution of *DISPERSION* and *DEFLDISP* in each *BEGPRICE* decile

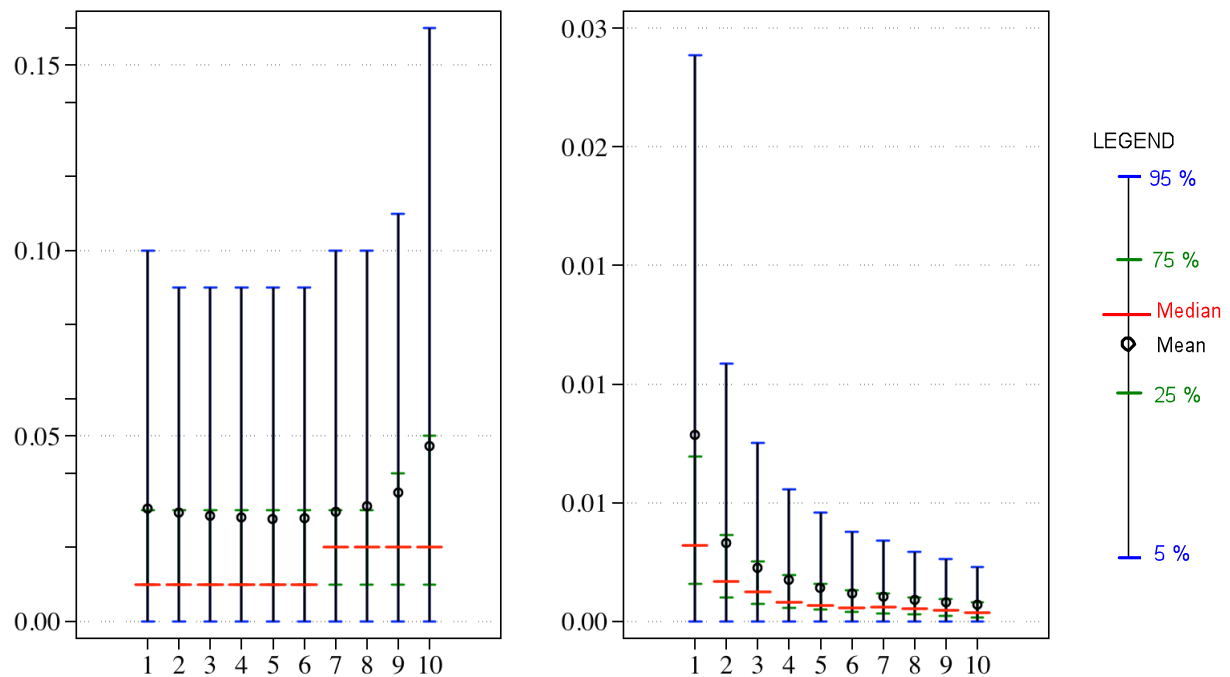
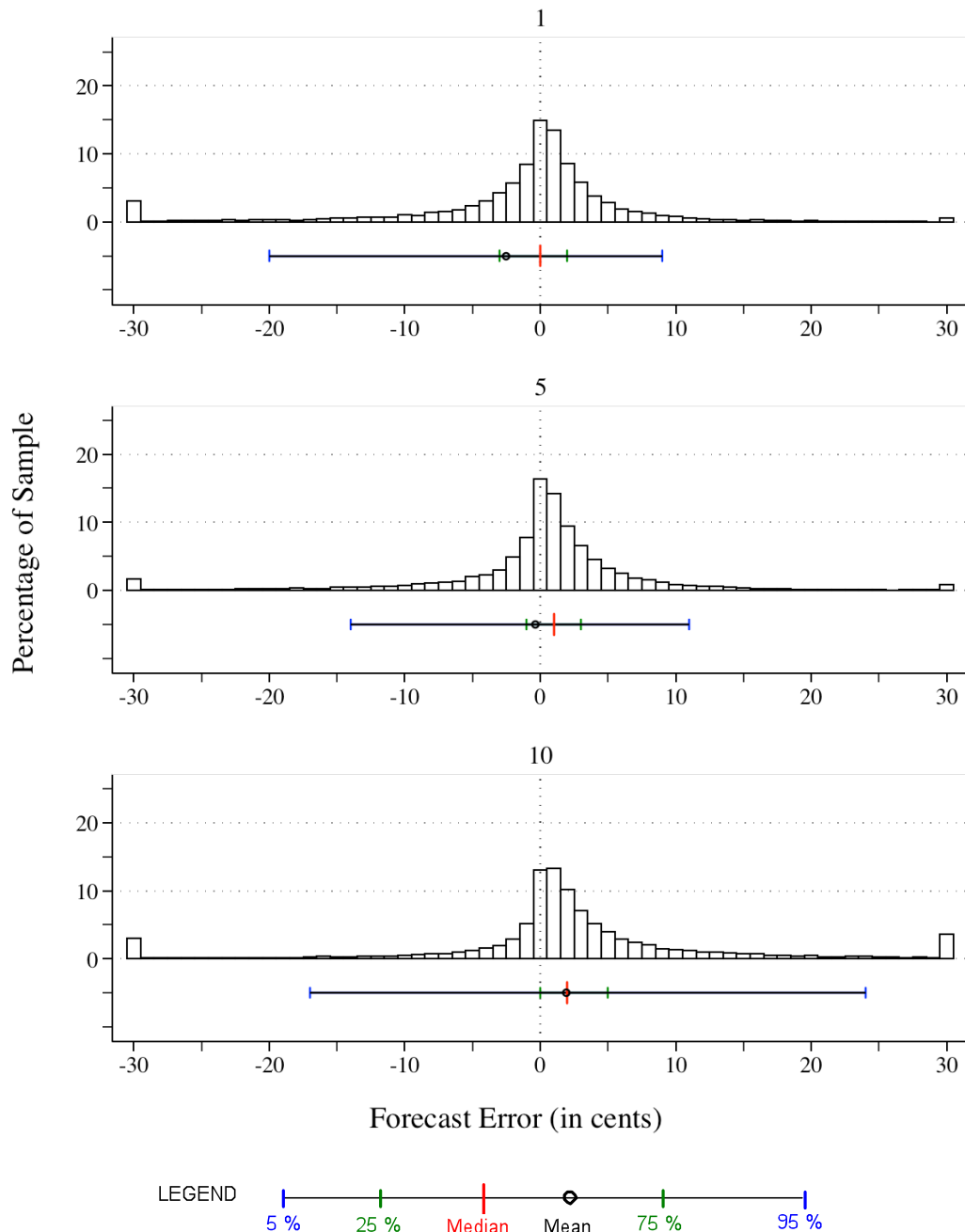


Figure 2. Histograms of forecast error and dispersion for selected *BEGPRICE* deciles

The histograms below for *FORECASTERR* and *DISPERSION* are provided for deciles 1, 5, and 10 of *BEGPRICE*, which is the beginning-of-year share price. Values below (above) -30 (30) cents are combined with observations in the -30 (30) cent group. The horizontal line below each histogram contains a solid circle to represent the mean, a long vertical hash mark for the median and hash marks for the 5th, 25th, 75th, and 95th percentiles. *FORECASTERR* is defined as *IBESACTUAL* minus *FORECAST*, where *IBESACTUAL* is the actual quarterly EPS (in dollars) as reported by I/B/E/S, and *FORECAST* is the most recent consensus (mean) EPS forecast (in dollars) for that firm-quarter. *DISPERSION* is the standard deviation of the *individual* analysts' EPS forecasts around the consensus in that quarter. All variables relate to firm-quarters, and are described in more detail in Appendix A.

Panel A: Histograms for *FORECASTERR* (for price deciles 1, 5, and 10)



Panel B: Histogram for *DISPERSION* (for price deciles 1, 5, and 10).

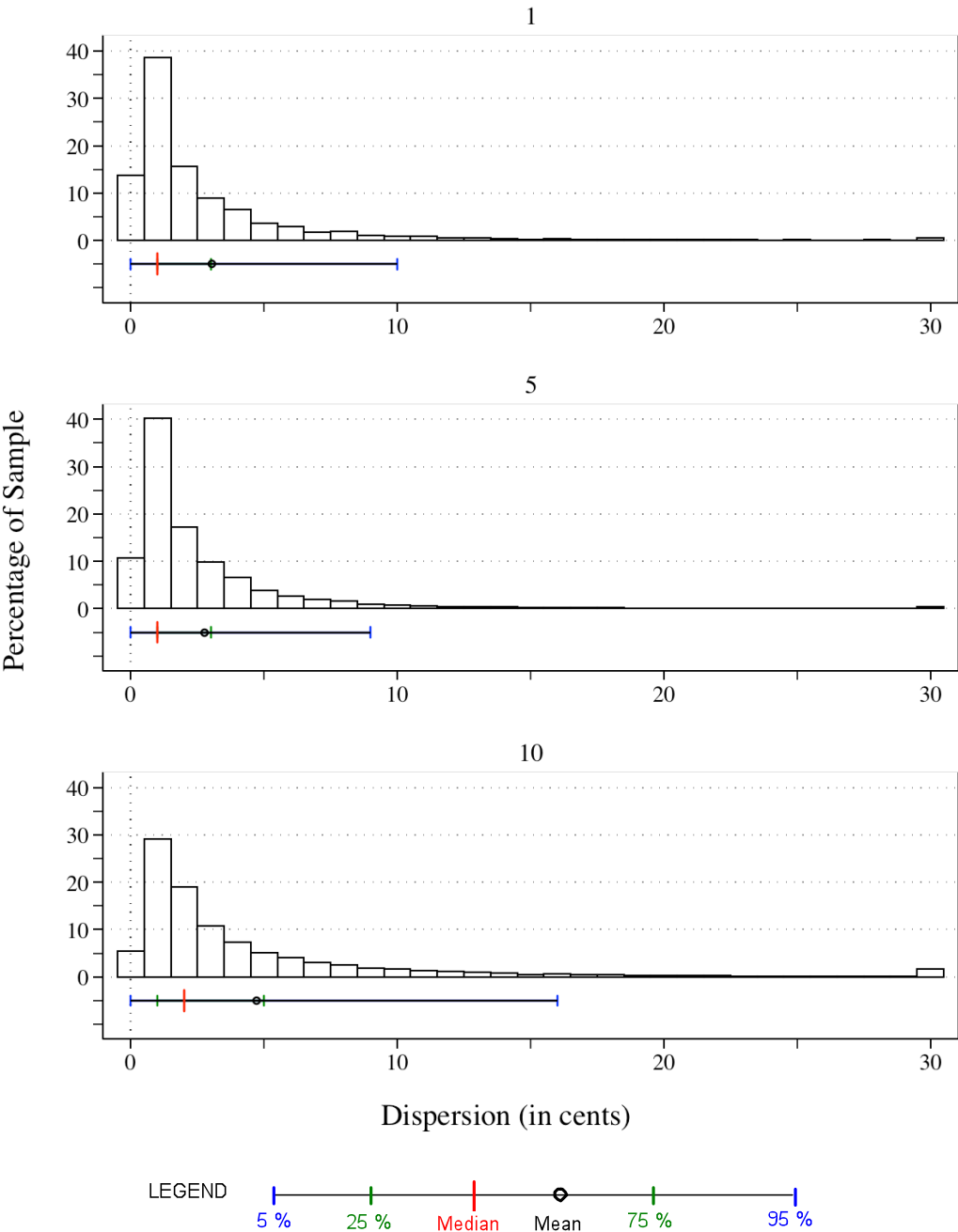


Table 1
Descriptive statistics

All variables relate to firm-quarters, and are described in more detail in Appendix A. *IBESACTUAL* is the actual quarterly EPS (in dollars) as reported by I/B/E/S and *COMPACTUAL* is the actual quarterly EPS (in dollars) as reported by Compustat. *FORECAST* is the most recent consensus (mean) estimate (in dollars) of *IBESACTUAL* for that firm-quarter, prior to the earnings announcement. *FORECASTERR* is defined as *IBESACTUAL* minus *FORECAST*. *ABSFE* is the absolute value of *FORECASTERR*. *DISPERSION* is the standard deviation of the individual analysts' EPS forecasts around the consensus in that quarter. *COVERAGE* is the number of estimates that underlie the consensus *FORECAST*. *BEGPRICE* is the share price (in dollars) at the beginning-of-year, and *INVBEGPRC* is the inverse of *BEGPRICE*. *MEANSTALE* and *SDSTALE* are the mean and standard deviation of forecast age (in days) of individual forecasts, respectively. *COMPFE* is *COMPACTUAL* minus *FORECAST*. *DEFLABSFE*, *DEFLDISP*, and *DEFLFE* are *ABSFE*, *DISPERSION*, and *FORECASTERR* scaled by *BEGPRICE*, respectively. *VOL* is the standard deviation of stock returns over the period from day -210 to -11, relative to its fiscal quarter-end.

Panel A: Distributional statistics

<i>Variable</i>	<i>N</i>	<i>Mean</i>	<i>StdDev</i>	<i>Min</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Max</i>
<i>ABSFE</i>	142,708	0.06	0.22	0.00	0.01	0.02	0.06	43.03
<i>BEGPRICE</i>	142,708	27.1	24.0	0.0	13.4	22.5	34.9	908.0
<i>COMPACTUAL</i>	139,800	0.28	0.65	-35.92	0.06	0.26	0.50	24.00
<i>COMPFE</i>	139,800	-0.02	0.38	-35.27	-0.03	0.01	0.04	23.52
<i>COVERAGE</i>	142,708	7.1	5.4	1.0	3.0	5.0	9.0	44.0
<i>DEFLABSFE</i>	142,708	0.0043	0.0362	0.0000	0.0004	0.0011	0.0032	8.6250
<i>DEFLDISP</i>	142,708	0.0020	0.0128	0.0000	0.0004	0.0008	0.0018	3.5339
<i>DEFLFE</i>	142,708	-0.0008	0.0364	-8.6250	-0.0006	0.0003	0.0015	3.8305
<i>DISPERSION</i>	142,708	0.03	0.06	0.00	0.01	0.02	0.03	4.30
<i>FORECAST</i>	142,708	0.31	0.49	-14.96	0.08	0.26	0.49	11.22
<i>FORECASTERR</i>	142,708	-0.00	0.22	-43.03	-0.01	0.01	0.03	8.84
<i>IBESACTUAL</i>	142,708	0.30	0.55	-46.46	0.08	0.27	0.50	12.21
<i>INVBEGPRC</i>	142,708	0.07	0.20	0.00	0.03	0.04	0.07	50.00
<i>MEANSTALE</i>	142,438	77.2	47.8	0.0	46.4	68.8	96.5	720.0
<i>SDSTALE</i>	140,001	46.9	39.2	0.0	20.1	37.4	64.3	523.9
<i>VOL</i>	142,679	0.030	0.017	0.002	0.018	0.026	0.038	0.300

Panel B: Pearson (lower diagonal) and Spearman (upper diagonal) correlation

	<i>ABSFE</i>	<i>BEGPRICE</i>	<i>COMPACTUAL</i>	<i>COMPFE</i>	<i>COVERAGE</i>	<i>DEFLABSFE</i>	<i>DEFLDISP</i>	<i>DEFLFE</i>	<i>DISPERSION</i>	<i>FORECAST</i>	<i>FORECASTERR</i>	<i>IBESACTUAL</i>	<i>INVBEGPRC</i>	<i>MEANSTALE</i>	<i>SDSTALE</i>	<i>VOL</i>
<i>ABSFE</i>		0.06	0.03	0.09	-0.08	0.87	0.36	0.13	0.46	0.05	0.17	0.03	-0.06	-0.15	-0.11	-0.03
<i>BEGPRICE</i>	0.08		0.58	0.13	0.44	-0.36	-0.45	-0.00	0.11	0.62	0.13	0.61	-1.00	-0.06	0.06	-0.45
<i>COMPACTUAL</i>	-0.10	0.39		0.43	0.24	-0.23	-0.28	0.19	0.05	0.93	0.26	0.95	-0.58	-0.01	0.08	-0.58
<i>COMPFE</i>	-0.17	0.02	0.67		0.05	0.03	-0.09	0.66	-0.02	0.19	0.69	0.30	-0.13	-0.01	-0.02	-0.12
<i>COVERAGE</i>	-0.03	0.34	0.14	0.00		-0.25	-0.18	0.03	0.05	0.26	0.09	0.26	-0.44	-0.02	0.14	-0.19
<i>DEFLABSFE</i>	0.42	-0.06	-0.14	-0.10	-0.05		0.52	0.14	0.37	-0.23	0.10	-0.24	0.36	-0.11	-0.12	0.17
<i>DEFLDISP</i>	0.14	-0.08	-0.14	-0.02	-0.05	0.44		-0.03	0.79	-0.27	-0.08	-0.28	0.45	-0.16	-0.09	0.16
<i>DEFLFE</i>	-0.31	0.02	0.11	0.15	0.02	-0.68	0.13		-0.05	0.05	0.96	0.20	0.00	0.01	-0.03	0.00
<i>DISPERSION</i>	0.33	0.15	-0.08	-0.05	-0.00	0.18	0.44	-0.00		0.09	-0.02	0.06	-0.11	-0.23	-0.07	-0.11
<i>FORECAST</i>	-0.00	0.51	0.81	0.11	0.19	-0.10	-0.16	0.03	-0.07		0.13	0.97	-0.62	-0.01	0.09	-0.60
<i>FORECASTERR</i>	-0.69	0.03	0.26	0.34	0.04	-0.31	-0.01	0.43	-0.09	0.08		0.29	-0.13	-0.01	-0.03	-0.05
<i>IBESACTUAL</i>	-0.28	0.47	0.82	0.23	0.19	-0.21	-0.15	0.20	-0.09	0.91	0.47		-0.61	-0.01	0.09	-0.59
<i>INVBEGPRC</i>	0.00	-0.17	-0.10	-0.02	-0.09	0.20	0.22	-0.10	-0.01	-0.12	-0.01	-0.12		0.06	-0.06	0.45
<i>MEANSTALE</i>	-0.02	-0.05	-0.03	-0.02	-0.05	0.01	-0.01	-0.02	-0.11	-0.03	-0.02	-0.03	0.03		0.63	0.04
<i>SDSTALE</i>	-0.03	-0.00	0.02	-0.01	0.03	-0.00	-0.00	-0.01	-0.04	0.03	-0.01	0.02	-0.00	0.66		-0.07
<i>VOL</i>	0.03	-0.20	-0.40	-0.12	-0.14	0.10	0.10	-0.05	0.00	-0.44	-0.05	-0.41	0.16	0.06	-0.04	

Panel C: Variation across *BEGPRICE* deciles in means and medians of selected variables, reported in the top and bottom halves of each row, respectively.

Variable	Stats	1	2	3	4	5	6	7	8	9	10	All
<i>BEGPRICE</i>	Mean	5.1	9.1	12.8	16.3	19.9	23.9	28.4	34.2	43.0	71.1	27.1
	Median	5.2	9.3	13.0	16.4	20.0	24.0	28.3	34.0	42.4	60.8	22.5
<i>COMPACTUAL</i>	Mean	-0.10	-0.01	0.07	0.14	0.22	0.29	0.36	0.42	0.54	0.79	0.28
	Median	-0.02	0.06	0.13	0.20	0.25	0.32	0.37	0.44	0.53	0.70	0.26
<i>COVERAGE</i>	Mean	4.1	4.7	5.3	5.8	6.3	6.9	7.5	8.3	9.5	11.9	7.1
	Median	3.0	4.0	4.0	4.0	5.0	5.0	6.0	7.0	8.0	11.0	5.0
<i>FORECAST</i>	Mean	-0.04	0.03	0.11	0.17	0.24	0.31	0.37	0.43	0.54	0.80	0.31
	Median	0.00	0.07	0.14	0.20	0.25	0.32	0.37	0.44	0.53	0.71	0.26
<i>IBESACTUAL</i>	Mean	-0.06	0.02	0.10	0.17	0.24	0.31	0.37	0.44	0.55	0.82	0.30
	Median	-0.01	0.07	0.14	0.20	0.26	0.32	0.37	0.44	0.54	0.72	0.27

Table 2

Distributional statistics for forecast error and dispersion in each *BEGPRICE* decile

This Table reports the mean, median, standard deviation (StdDev), inter-quartile range (QRange), and the number of observations (N) for distributions of forecast error, absolute forecast error, and forecast dispersion for different deciles of *BEGPRICE*, which is the beginning-of-year share price (in dollars). Price deciles are computed each calendar year, and the lowest (highest) price decile is denoted by 1 (10). *FORECASTERR* is defined as *IBESACTUAL* minus *FORECAST*, where *IBESACTUAL* is the actual quarterly EPS (in dollars) as reported by I/B/E/S, and *FORECAST* is the most recent consensus (mean) estimate (in dollars) of *IBESACTUAL* for that firm-quarter. *COMPFE* is *COMPACTUAL* minus *FORECAST*, where *COMPACTUAL* is the actual quarterly EPS (in dollars) as reported by Compustat. *ABSFE* is the absolute value of *FORECASTERR*. *DISPERSION* is the standard deviation of the *individual* analysts' EPS forecasts around the consensus in that firm-quarter. *DEFLFE*, *DEFLABSFE*, and *DEFLDISP* are defined as *FORECASTERR*, *ABSFE*, and *DISPERSION* scaled by *BEGPRICE*, respectively. Additional details for all variables are provided in Appendix A.

Panel A1: Distributional statistics for *FORECASTERR* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	-0.02	-0.01	-0.01	-0.00	-0.00	0.00	0.00	0.00	0.01	0.02	-0.00
Median	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01
StdDev	0.24	0.17	0.17	0.15	0.21	0.17	0.18	0.40	0.16	0.26	0.22
QRange	0.05	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.05	0.04
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Panel A2: Distributional statistics for *COMPFE* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	-0.06	-0.04	-0.03	-0.02	-0.02	-0.01	-0.01	-0.00	-0.00	0.00	-0.02
Median	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01
StdDev	0.34	0.49	0.32	0.30	0.28	0.36	0.36	0.32	0.36	0.57	0.38
QRange	0.07	0.08	0.07	0.06	0.07	0.07	0.07	0.08	0.09	0.10	0.07
N	12,256	13,700	13,633	13,941	14,010	14,175	14,356	14,390	14,636	14,703	139,800

Panel A3: Distributional statistics for *DEFLFE* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	-0.0066	-0.0015	-0.0009	-0.0003	-0.0002	0.0000	0.0001	0.0001	0.0003	0.0003	-0.0008
Median	0.0000	0.0000	0.0000	0.0005	0.0004	0.0004	0.0003	0.0003	0.0002	0.0002	0.0003
StdDev	0.1180	0.0211	0.0139	0.0092	0.0105	0.0076	0.0061	0.0113	0.0036	0.0032	0.0364
QRange	0.0102	0.0053	0.0036	0.0028	0.0021	0.0017	0.0015	0.0013	0.0012	0.0008	0.0021
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Panel B1: Distributional statistics for *ABSFE* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	0.07	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.06	0.09	0.06
Median	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.02
StdDev	0.23	0.17	0.16	0.14	0.21	0.17	0.17	0.40	0.14	0.25	0.22
QRange	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.06	0.07	0.05
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Panel B2: Distributional statistics for *DEFLABSFE* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	0.0183	0.0067	0.0047	0.0035	0.0029	0.0023	0.0020	0.0018	0.0015	0.0013	0.0043
Median	0.0050	0.0027	0.0018	0.0014	0.0011	0.0009	0.0008	0.0007	0.0006	0.0005	0.0011
StdDev	0.1167	0.0201	0.0132	0.0085	0.0101	0.0072	0.0058	0.0111	0.0033	0.0030	0.0362
QRange	0.0113	0.0054	0.0038	0.0029	0.0024	0.0019	0.0017	0.0014	0.0014	0.0011	0.0028
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Panel C1: Distributional statistics for *DISPERSION* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.05	0.03
Median	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02
StdDev	0.07	0.06	0.05	0.06	0.04	0.05	0.05	0.06	0.06	0.09	0.06
QRange	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.04	0.02
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Panel C2: Distributional statistics for *DEFLDISP* in each *BEGPRICE* decile

	1	2	3	4	5	6	7	8	9	10	All
Mean	0.0079	0.0033	0.0023	0.0017	0.0014	0.0012	0.0011	0.0009	0.0008	0.0007	0.0020
Median	0.0032	0.0017	0.0012	0.0008	0.0007	0.0006	0.0006	0.0005	0.0005	0.0004	0.0008
StdDev	0.0418	0.0068	0.0037	0.0037	0.0021	0.0020	0.0018	0.0016	0.0013	0.0012	0.0128
QRange	0.0054	0.0026	0.0018	0.0014	0.0011	0.0009	0.0008	0.0007	0.0007	0.0006	0.0015
N	12,458	13,878	13,886	14,159	14,301	14,446	14,638	14,760	14,968	15,214	142,708

Table 3
Extension of analyses in Tables 3 and 4 of Thomas [2002] to show price deflation effect

Panel A reports the Pearson (Spearman) correlation of selected variables from Thomas [2002] in the lower (upper) diagonal. Panel B (C) reports a partial view of the regression results based on Table 3 (Table 4) of Thomas [2002], which investigates the relation between absolute forecast error (forecast dispersion) and diversification. Absolute forecast error (*AFE*) and dispersion are measured as $|IBESACTUAL - median FORECAST|$ and standard deviation of analyst forecasts. When scaled by *PRICE5*, which is share price five days before the annual earnings announcement, we denote them as *DEFLATAFE* and *DEFLATDISP*. Diversification is measured by *HERF*, which is the Herfindahl Index, based on assets reported for different segments. A smaller value of *HERF* represents more diversification or more balanced asset investments spread across more segments. *RESIDVOL*, measured as the standard deviation of the market model residuals over the period from 210 to 11 days before the earnings announcement date, is a control variable that is included in equation (5) in both Panels. See Thomas [2002] for more details. Specification I in Panels B & C refers to the regressions estimated in the original study. Specification II includes the inverse of *PRICE5* (*INVPRICE5*) as an additional regressor. Specification III returns to specification I but considers undeflated values of the dependent variables. Specification IV adds price as an additional regressor to specification III. Associated White [1980] t-statistics are reported in parentheses below each coefficient estimate, and significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Panel A: Pearson (lower diagonal) and Spearman (upper diagonal) correlation

	<i>AFE</i>	<i>DEFLATAFE</i>	<i>DISP</i>	<i>DEFLATDISP</i>	<i>PRICE5</i>	<i>INVPRICE5</i>	<i>HERF</i>	<i>RESIDVOL</i>
<i>AFE</i>		0.89	0.45	0.48	-0.13	0.13	-0.10	0.12
<i>DEFLATAFE</i>	0.56		0.32	0.63	-0.54	0.54	0.03	0.39
<i>DISP</i>	0.50	0.14		0.75	0.14	-0.14	-0.25	-0.17
<i>DEFLATDISP</i>	0.27	0.53	0.49		-0.51	0.51	-0.03	0.30
<i>PRICE5</i>	0.10	-0.26	0.29	-0.25		-1.00	-0.29	-0.69
<i>INVPRICE5</i>	0.04	0.58	-0.09	0.52	-0.48		0.29	0.69
<i>HERF</i>	-0.09	0.06	-0.17	0.03	-0.26	0.18		0.40
<i>RESIDVOL</i>	0.05	0.42	-0.11	0.35	-0.47	0.64	0.34	

Panel B: Selected coefficients from regressions based on Table 3 of Thomas [2002].

Specification Dep. Var.	Variable	Equation				
		(1)	(2)	(3)	(4)	(5)
I <i>DEFLATAFE</i>	<i>HERF</i>	2.55 (8.65)***	0.91 (2.73)***	0.95 (2.88)***	0.86 (2.49)**	-1.02 (3.14)***
	<i>RESIDVOL</i>					4.37 (22.82)***
II <i>DEFLATAFE</i>	<i>HERF</i>	-1.40 (5.13)***	0.45 (1.59)	0.47 (1.65)*	0.35 (1.19)	-0.36 (1.20)
	<i>INVPRICE5</i>	67.01 (24.11)***	71.55 (22.73)***	70.38 (22.09)***	70.27 (22.05)***	56.28 (17.11)***
	<i>RESIDVOL</i>					1.89 (11.46)***
III <i>AFE</i>	<i>HERF</i>	-0.2701 (7.10)***	0.0139 (0.30)	0.0154 (0.33)	-0.0021 (0.04)	-0.0861 (1.84)*
	<i>RESIDVOL</i>					0.1955 (18.52)***
IV <i>AFE</i>	<i>HERF</i>	-0.1954 (4.35)***	0.0159 (0.33)	0.0209 (0.43)	0.0049 (0.10)	-0.0737 (1.56)
	<i>PRICE5</i>	0.0031 (3.10)***	0.0004 (0.36)	0.0011 (0.97)	0.0013 (1.16)	0.0038 (3.32)***
	<i>RESIDVOL</i>					0.2151 (16.98)***

Panel C: Selected coefficients from regressions based on Table 4 of Thomas [2002].

Specification Dep. Var.	Variable	Equation				
		(1)	(2)	(3)	(4)	(5)
I <i>DEFLATDISP</i>	<i>HERF</i>	0.29 (5.49)***	0.17 (2.66)***	0.18 (2.84)***	0.13 (2.01)**	-0.16 (2.83)***
	<i>RESIDVOL</i>					0.68 (25.54)***
II <i>DEFLATDISP</i>	<i>HERF</i>	-0.39 (7.85)***	0.09 (1.63)	0.09 (1.75)*	0.04 (0.76)	-0.04 (0.74)
	<i>INVPRICE5</i>	11.53 (21.12)***	12.57 (20.39)***	12.18 (19.80)***	12.12 (19.72)***	10.52 (15.51)***
	<i>RESIDVOL</i>					0.22 (7.75)***
III <i>DISP</i>	<i>HERF</i>	-0.1401 (16.02)***	-0.0062 (0.57)	-0.0060 (0.55)	-0.0124 (1.15)	-0.0228 (2.15)**
	<i>RESIDVOL</i>					0.0241 (12.31)***
IV <i>DISP</i>	<i>HERF</i>	-0.0766 (7.37)***	0.0017 (0.15)	0.0032 (0.28)	-0.0022 (0.20)	-0.0153 (1.43)
	<i>PRICE5</i>	0.0026 (11.61)***	0.0016 (6.07)***	0.0018 (6.83)***	0.0019 (7.15)***	0.0023 (8.82)***
	<i>RESIDVOL</i>					0.0360 (13.86)***