Lured by the Consensus: The Implications of Treating All Analysts as Equal

Roni Michaely, Amir Rubin, Dan Segal, and Alexander Vedrashko*

Sep. 24, 2017

Abstract:

We find that the market's focus on the consensus earnings forecast and not differentiating among analysts according to their quality has significant negative economic implications. We classify analysts into high and low quality (HQ and LQ) categories based on their forecast accuracy and find that the market overweighs the information content of the consensus forecast. HQ analysts' superior forecasting ability is persistent across stocks they cover, as well as recommendations they issue. The market does not fully utilize price-relevant information in the forecasts and recommendations of the HQ analysts. In particular, the HQ analysts' recommendation changes and forecast dispersion predict the firm's stock return and return volatility next month. In addition, the PEAD phenomenon is present only when the HQ analysts are relatively uncertain about the firm's performance. At the aggregate level, recommendation changes of the HQ analysts predict future industry and market returns, while the consensus recommendation changes do not, and market volatility is higher following periods of greater uncertainty among the HQ analysts. Overall, our results indicate that investors' fixation on the consensus can lead to less accurate forecasts and inefficient prices.

Keywords: Analyst quality, Forecasts, Consensus, PEAD, recommendations

*Michaely is from Johnson@Cornell Tech—Cornell University (rm34@cornell.edu). Rubin is from Simon Fraser University and IDC (arubin@sfu.ca). Segal is from IDC and Singapore Management University (dsegal@idc.ac.il). Vedrashko is from Simon Fraser University (awv@sfu.ca). We thank seminar participants at NFA 2017 conference, Tel-Aviv University and the University of Vienna for helpful comments and suggestions.

Investors and academics alike use analysts' consensus forecasts as the measure of market expectations of firms' future earnings. The perceived importance of consensus earnings estimates has greatly increased in recent years, to the extent that even companies' investor relations departments tend to follow it on a continuous basis (Consensus earnings estimates report, 2013). The high publicity of the consensus forecast and investor's fixation on the mean of analysts' forecast distribution can be described as an instance of central fixation bias, which is people's tendency to fixate their vision at the center of a group of objects and which can be optimal for initial information processing (Tatler, 2007). However, investors' continuous fixation on the consensus can have negative economic implications. For example, the consensus, by construction, ignores the possibility that analysts may have different abilities and, consequently, varying forecast accuracy due to their varying experience (Mikhail, Walther, and Willis, 1997; Clement, 1999; Hirst, Hopkins, and Wahlen, 2004), aptitude (Jacob, Lys and Neale, 1999), education (Maines, McDaniel, and Harris, 1997; De Franco and Zhou, 2009), brokerage house association (Clement, 1999), proximity to firm (Malloy, 2005), or work habits (Rubin, Segal, and Segal, 2017) among others. Given the evidence on differences in analysts' ability, there is no *a priori* reason to believe that the consensus forecast is the best estimate of the market's expectations or that the consensus recommendation (e.g., Jegadeesh et al., 2004) is the best signal to follow. The market's fixation on a simple average of analysts' forecasts or recommendations that disregards differences in analyst forecasting ability motivates us to examine whether it leads to the market's reliance on less accurate forecasts, inefficient pricing, and suboptimal use of information.

Our findings can be summarized as follows: Investors do not sufficiently recognize quality differences among analysts and react to the consensus forecast rather than the more accurate forecast generated by the high quality (HQ) analysts. This inefficient handling of information in analysts' forecasts suggests market mispricing around earnings announcements. One can exploit this inefficiency by predicting earnings surprises using the difference between the HQ analysts' average forecast and the consensus to generate positive returns.

With the same analyst ranking used to uncover investors' fixation on the consensus, we investigate other economic consequences of the market's lack of awareness of the superior ability of the HQ analysts. We observe that investors do not sufficiently react to recommendation revisions of the HQ analysts, which allows for predicting stock returns based on the HQ analysts' recommendation revisions. Next, because the HQ analysts' forecasts contain relatively more information, the dispersion of their forecasts also contains more information. We find that, unlike the dispersion of all analysts' forecasts, the HQ analysts' forecast dispersion before annual earnings announcements strongly predicts return volatility for the firm one month ahead. Further, our methodology of ranking analyst quality based on their forecasting performance enables us to provide a new insight on the post-earnings announcement drift (PEAD) phenomenon. We find that the PEAD exists only when the HQ analysts are relatively uncertain (compared to all analysts following the firm) about the firm's prospective earnings.

Finally, the HQ analysts' superior ability to forecast individual firms' earnings might translate into better forecasting of the industry and market performance as well. Our results indeed indicate that aggregating forecasts and recommendation revisions of the HQ analysts across all firms during the earnings announcement month provides valuable information about the future stock market and industry returns and market volatility, which is not found for aggregated forecasts and recommendations provided by all analysts. Taken together, we find that the consensus fixation phenomenon and our other findings on recommendations, return volatility, and the PEAD share the same economic mechanism causing investors to systematically underweight quality differences among analysts and the information output of the HQ analysts. Relying on the forecasts and recommendations of analysts with persistently high-quality outputs results in better decisions for the investing public both at the firm and aggregate market levels.

We start by examining a key necessary condition implicit in the principle of differentiating analysts in terms of their quality—that analysts' quality measured by forecast accuracy is persistent. We define high and low quality (HQ and LQ) analysts as those, respectively, above and below the median in the accuracy ranking in the previous year.[1] Two earlier studies (Stickel, 1992; Sinha, Brown, and Das, 1997) report the persistence in forecast accuracy in different subsamples of analysts.[2] Our empirical design is more general in that it analyzes all analysts across time and firms. We find that analysts who are categorized as HQ in a given year tend to be ranked as HQ in the following year as well and that analysts who are HQ in one firm are also likely to be HQ in the other firms they follow. The persistence in forecasting ability across time and firms implies that forecasting performance captures analysts' quality.

The superiority of the HQ analysts suggests that it is more optimal for investors to use their average forecast rather than the consensus forecast and, consequently, the market should react more vigorously to earnings surprises that are measured based on the average forecast of the HQ analysts. However, we find the earnings response coefficient on the standardized unexpected earnings (SUE) based on the consensus forecast to be higher than that of SUE based on the average of the HQ analysts' estimates. This finding implies the market pays too much attention to the consensus forecast and fails to fully incorporate the information embedded in the forecasts of the HQ analysts. Indeed, a trading strategy based on differences between the mean forecast of the high quality (HQ) analysts and the consensus yields economically and statistically significant abnormal returns over the announcement day and the following trading day.

[1] A ranking based on the last year's forecast accuracy is also used in Loh and Mian (2006) and supported by Sinha, Brown, and Das (1997), who find it to be superior to rankings based on more years, and Carpenter and Lynch (1999), who find it to be relatively less exposed to survivorship bias. We conduct further sensitivity tests indicating that our findings are not affected by different classifications of analysts into the HQ and LQ categories.

[2] Stickel (1992) analyzes forecast revisions by analysts who are members of the All-American Research Team, where the All-American status is based on both the past forecasts' accuracy and other criteria. Sinha, Brown, and Das (1997) rank analysts into three categories based on their annual forecast errors in the previous years and find persistence for the top category. Brown (2004) finds that these two models built on past forecasting performance predict analysts' forecasting accuracy as well as a model based on analysts' individual characteristics (Clement, 1999).

Our finding that analysts have different quality suggests these differences should manifest themselves in other aspects of analysts' informational output as well. First, similar to the earnings announcement results above, we find that when investors are fixated on the consensus forecast the market does not fully impound the information associated with recommendation revisions of the HQ analysts. Specifically, we observe that only the HQ analysts' recommendation revisions in the earnings announcement month predict stock returns next month; a strategy that is long (short) in stocks where the HQ analysts on average provide an upgrade (downgrade) produces a statistically significant 0.9% return in the following month. Importantly, these results do not hold for analysts classified as LQ.

Second, given our finding that the forecasts by the HQ analysts are more informative about the future level of earnings than the consensus or the LQ analysts' forecasts, we consider whether the second moment of the HQ analysts' forecasts is also more strongly associated with uncertainty regarding future firm performance than that of the consensus or the LQ analysts' forecasts. We find a weak relation between the forecast dispersion of all analysts following the firm and its stock return volatility next month. However, when we examine the relation separately for HQ and LQ analysts we find that the dispersion of the HQ analysts' forecasts is a strong predictor of the firm's stock return volatility in the month following the annual earnings announcement month, whereas the LQ analysts' forecast dispersion does not predict return volatility.

The finding that it is the HQ analysts' forecast dispersion that captures the firm's uncertainty allows us to consider a possible role of analysts' forecast quality in the relation between forecast dispersion and the post-earnings announcement drift (PEAD).[3] We find that the PEAD is

[3] The most closely related studies are the following. The model in Abarbanell, Lanen, and Verrecchia (1995) suggests that when forecast dispersion is high investors delay their complete response to earnings announcements, which we suggest could lead to a greater PEAD. Zhang (2006a) finds that analysts' forecast dispersion predicts the price drift following analysts' forecasts (the relation to the PEAD is not tested). Hung, Li, and Wang (2014) do not consider analysts' forecast dispersion but find a causal relation between uncertainty about firm performance and the PEAD.

higher when the HQ analysts are more uncertain relative to all analysts covering the firm. Specifically, the standard PEAD strategy is that of buying (shorting) shares when SUE is positive (negative). We implement the strategy for the subsamples in which the forecast dispersion of HQ is greater (lower) than the dispersion of all analysts. We find a significantly greater PEAD (annualized 9.4% after 11 months) if the HQ analysts are relatively uncertain. During most of this forecast horizon, the PEAD is not statistically different from zero in the sample where the HQ analysts are relatively less uncertain, implying that the long-puzzling PEAD phenomenon arises primarily during the periods of high uncertainty among the HQ analysts. Overall, these findings indicate that the superior information in the HQ analysts' forecasts predicts not only the immediate reaction to earnings announcements but also the long-term market response.

Finally, having established that the HQ analysts' recommendations and forecast dispersion are better predictors than the consensus at the firm level, we next explore their predictive ability at the market-wide level. Because individual HQ analysts' recommendation changes predict the firm's returns, the HQ analysts' average recommendation change across all firms in the market should predict the market return. The argument for the average dispersion of analysts' forecasts predicting market volatility is similar. Indeed, we find that by relying on the average recommendation changes of the HQ analysts, one can predict the market and industry returns in the following month, in contrast to the average recommendation changes of all analysts or the LQ analysts following the firm. For example, a long-short strategy based on the direction of the HQ recommendation revisions produces a 7.9% annualized return in the post-announcement month. We also find that the HQ analysts' normalized dispersion is associated with a higher market volatility as measured by the VIX next month. In contrast, analysts' forecast dispersion at the consensus level or that of the LQ analysts do not have a relation to the VIX. This implies our measure of uncertainty based on the HQ analysts' normalized dispersion has a systematic risk component. Given that the VIX is often interpreted as the "fear index", investors should be worried about the economy and the stock market performance when the HQ analysts become relatively uncertain compared to all other analysts.

Our contribution to the literature consists of two levels—the core findings and the implications. Our study belongs to a recent literature suggesting that the average of analysts' estimates can be inaccurate and can be improved upon (So, 2013; Buraschi, Piatti, and Whelan, 2017). The consensus fixation is also related to the limited attention literature (Hirshleifer, Lim, and Teoh, 2009) in that our findings suggest that investors may prefer a single number of the consensus to spending their cognitive effort on assessing analyst quality.

The implications of our core contribution touch four different strands of literature and are driven by the insight that investors' fixation on the consensus forecast is associated with their inefficient use of analysts' other informational outputs. First, our findings directly contribute to the literature differentiating analysts in terms of the value of their recommendations (Sorescu and Subrahmanyam, 2006; Loh and Stulz, 2011) and are most closely related to Loh and Mian (2006) and Ertimur, Sunder, and Sunder (2007), who measure analyst quality with forecast accuracy and analyze the quality of recommendations with returns. Our study is different from them in that it generalizes the use of a single measure of analyst quality across multiple information outputs by analysts and also finds a predictive rather than contemporaneous relation with returns, which more clearly indicates the market's insufficient attention to analyst quality differences.

Second, we contribute to a relatively underinvestigated topic of a relation between analysts' forecast dispersion and stock return volatility, for which our method allows us to find a predictive rather than contemporaneous relation reported in Ajinkya and Gift (1985).

Third, our findings limit the extent of the PEAD anomaly's challenge to market efficiency (Fama 1998) in that the PEAD is restricted to the times with high uncertainty about the firm's prospects. We also contribute to the discussion about rational and behavioral explanations to the

PEAD (Brav and Heaton, 2002). Our finding that our measure of uncertainty is associated with both systematic risk and the PEAD points in the direction of a rational explanation to the PEAD.

Fourth, we advance the literature that aims to find a relation between analyst outputs and industry and market-level variables (Park, 2005; Boni and Womack, 2006; Kadan et al., 2012). We are the first to find the predictability of market returns and the VIX based on the aggregation of the HQ analysts' firm-level recommendations and forecast dispersion, respectively. This result neither assumes nor indicates that the HQ analysts have superior macroeconomic knowledge or ability to predict market-level developments (e.g., Hutton, Lee, and Shu, 2012).

2. Data and variables

We use the sample of annual EPS estimates and earnings announcements in I/B/E/S during the period from January 1992 to December 2015 for companies with daily return data in CRSP.[4] The starting year of 1992 is chosen because some of the analyses require analysts' recommendation data, which begins in 1993. Earnings estimates and actual earnings are adjusted for splits using the daily cumulative adjustment factor from CRSP (Glushkov and Robinson, 2006). Each year, we rank analysts based on the closest absolute forecast error, which is the absolute difference between an analyst's earnings forecast closest to the earnings announcement (but made at least one day prior to the announcement) and actual earnings, divided by the share price at the beginning of the calendar year. From the initial sample, we generate 861,349 firm-year-analyst rankings based on the closest forecast error. This number drops to 804,003 observations once we require firms to have Compustat data. Next, to avoid small sample bias in our ranking when the number of analysts following the firm is small, we exclude firm-years with less than four analysts

[4] We focus on annual rather than quarterly earnings for two main reasons. First, fewer analysts provide quarterly forecasts than annual forecasts. Second, annual earnings announcements are typically more informative, combined with a conference call, and followed by recommendation changes.

following, which reduces the sample to 750,295 observations. In addition, for the analyst-level persistence analyses (Figures 1 and 2; Tables 1 and 2), analysts must appear in the data in two consecutive years for a given annual announcement, reducing the sample to 485,815 observations. Rankings based on past performance are common not only for analyst forecasting persistence studies (e.g., Stickel, 1992; Sinha, Brown, and Das, 1997) but also in the mutual fund and pension fund literature studying performance persistence (e.g., Hendricks, Patel, and Zeckhauser, 1993; Carhart, 1997; Tonks, 2005).[5]

In the firm-level regressions, we control for the following firm characteristics—size, annual stock return, book-to-market ratio, number of analysts following, and leverage. *Firm size* is the market value of the firm's equity at the end of the month prior to the earnings announcement month. *Annual stock return* is measured based on monthly equity returns in the 12 months prior to the earnings announcement month. The *book-to-market ratio* is computed as stockholder equity minus preferred stock plus deferred taxes at the end of the fiscal year for which the earnings are announced divided by firm size. The *number of analysts* is the number of analysts who made at least one earnings forecast for this announcement. *Leverage* is the book values of total liabilities divided by total assets at the end of the fiscal year for which the earnings are announced. Some of the regression models also control for analyst characteristics. Specifically, *Overall tenure* is the number of years since the analyst first appeared in the I/B/E/S file. *Firm-specific tenure* is the

[5] An alternative ranking procedure would be to rank analysts in a given year by averaging their forecast errors across firms they follow. There are several advantages for this alternative ranking procedure. Analysts follow 15 firms on average, which implies that this procedure could avoid small sample bias when a firm is followed by too few analysts and, perhaps, achieve a higher level of persistence in analyst ranking. It would also avoid losing the observations of the first year when an analyst begins covering a firm because we could rely on the analyst's ranking in the previous year in other firms. However, this year-level ranking approach also has several pitfalls. First, an aggregated ranking across firms can be misleading if analysts' ability to predict earnings is mainly firm- or industry-specific. Second, with the year-level ranking, we end up with some firms followed almost exclusively by high or low quality analysts, and, as we find, populated by just one analyst-quality type. This would undermine our study's objective because we compare the average estimate of the HQ analysts to the consensus estimate in each firm. While the cross-firm ranking is not suitable for this study, we analyze the relation between an analyst's forecast accuracy in a given firm and all other firms covered by the analyst and find it supporting our time-dimension ranking measure. See Section 3.2 below.

number of years since the analyst began covering the company in the I/B/E/S file. *Brokerage house size* is the number of analysts employed by the brokerage firm. *Firm coverage* is the number of firms covered by the analyst.

In the models predicting industry and market returns and volatility, most of the controls we use follow Li, Ng, and Swaminathan (2013) and are for the month prior to the dependent variable's month. The *earnings-to-price ratio* and *dividend-to-price ratio* are calculated from the S&P 500 dividend, earnings, and price data on Robert Shiller's website. The one month *T-bill rate* and *30-year Treasury yield* are obtained from WRDS. *Term spread* is the difference between AAA rated corporate bond yields obtained from the FRED (Federal Reserve Bank in St. Louis) database and the one-month T-bill yield. The *default spread* is the difference between BAA and AAA rated corporate bond yields, obtained from the FRED. *Inflation* is the change in CPI (all urban consumers) obtained from FRED. Following Da, Engelberg, and Gao (2015), our VIX regressions also control for the perceived economic policy uncertainty (EPU), which is a news-based measure provided by Baker, Bloom, and Davis (2015). *EPU change* is the percentage change in the monthly average of the daily EPU for the month prior to the dependent variable's month. The VIX index is from WRDS.

3. Persistence in analysts' forecasting ability

We partition analysts into the high and low quality categories based on their absolute forecast error and then analyze whether this classification of analysts persists in the following year. We define HQ (LQ) analysts based on whether their absolute closest forecast error for the firm-year is below (above) the median absolute forecast error for the firm-year. We choose the median as the cutoff due to its advantage that the numbers of analysts in the high and low quality groups are equal in year t-1 and, consequently, remain relatively close in year t. This mitigates a possible effect of the number of analysts on the comparisons between the groups. In the robustness section, we discuss our findings for other cutoff values defining the HQ and LQ analysts.[6]

Figure 1 shows the mean absolute forecast errors of HQ analysts and the consensus during the 300 days prior to the earnings announcement. We observe acceleration in the reduction of the mean forecast error around quarterly earnings announcements at 90, 180, and 270 day marks. The graph shows that the mean absolute forecast error of all analysts is higher than the mean absolute forecast error of the HQ analysts in all days prior to the earnings announcement. The mean absolute forecast errors of the consensus and HQ analysts decrease over time, respectively, to around 0.012 and 0.0115 one day before the earning announcement. This difference of 4.17% ($\frac{0.0115}{0.0120} - 1$) is economically meaningful and statistically significant (p-value<0.01). Also notable is that the HQ analysts' accuracy 30 days before the announcement is already higher than the consensus accuracy at the announcement.

Table 1 analyzes how the classification of analysts to low or high quality is associated with various analyst characteristics and the persistence of the classification over time. Panel A provides univariate comparisons. We find that the HQ analysts tend to be more experienced, are employed by larger brokerage firms, and cover more firms. To analyze the persistence of analysts' forecast accuracy, we compare the HQ and LQ analysts' forecast errors in the year after they were ranked. The absolute forecast errors of the HQ analysts remain smaller than those of the LQ analysts—the difference is 9% (0.0081/0.0089) and statistically significant. In the last line of the panel, we find

[6] The literature on optimally combining forecasts to minimize the out-of-sample combined forecast performance is vast (Clemen, 1989). Our equal-weighting forecasts of the best performing subset of analysts is also similar to the approach investigated, for instance, in Aiolfi and Timmerman (2006). Obviously, there can be methods combining forecasts that are more accurate than our HQ analysts' average forecast, although simple averaging of expert forecasts is found to be more optimal or almost equivalent to more sophisticated weighting methods for various economic series (Genre et al 2013). Our study does not attempt to contribute to this literature and does not require an analyst combination that beats the consensus by the biggest margin. Instead, using a parsimonious approach of combining analysts, our goal is to consider the economic implications of the market ignoring significant variation in analyst quality.

that both the HQ and LQ analysts have an optimism bias in their forecasts (the average forecast errors are significantly different from zero, with untabulated p-values<0.01), but there is no statistical difference in optimism bias between them.[7]

The analysis in Panel B of Table 1 examines the persistence in the quality classification of analysts. In the probit models in columns (1) and (2), the dependent variable is an indicator that equals one if the analyst is of HQ and zero otherwise. In columns (3) and (4), the dependent variable is the absolute forecast error, a continuous variable, which allows us to control for firm fixed effects in the regression. In columns (1) and (3), we control for firm characteristics, and in columns (2) and (4), we control for both firm and analyst characteristics. The main coefficient of interest is the HQ classification in year t-1. The results show that the coefficient on *HQ analyst indicator (t-1)* is highly significant (p-value<0.01) in all specifications, indicating that analysts' rankings and forecast accuracy are persistent in consecutive years. For example, the unconditional probability of belonging to the HQ group is approximately 50%, and accounting for the HQ status in the previous year increases this likelihood by approximately 4.1% according to columns (1) and (2). Columns (3) and (4) show that HQ analysts continue to have lower absolute forecast errors in the following year. Their average absolute forecast error is 8.2% lower (0.00072/0.0085) than the average absolute forecast error for all analysts.

We next conduct cross-firm tests to examine whether forecasting performance is persistent not only through time but also across stocks the analyst follows. This analysis is not only important in its own right but its affirmative findings will add confidence to the concept that some analysts are indeed better than others. We define an analyst's performance in the other firms as that of high (low) quality if the analyst is classified in high (low) quality category in the majority of the other

[7] For robustness, in untabulated tests, we distinguish between firms with high (more than 10 analysts following) and low analyst coverage, which also approximates large and small firms. On the whole, the full sample relations hold for both types of firms, indicating that the differences between HQ and LQ are not associated with firm size.

firms he or she follows during the year (excluding this firm).[8] Panel A of Table 2 reports that HQ analysts in a given firm are also ranked as HQ in the other firms that they follow 54.4% of the time, which is statistically greater than the unconditional percentage of HQ analysts in a given firm, 48.3%.[9] LQ analysts in a given firm also tend to be LQ in the other firms they follow; LQ analysts in a given firm are LQ in 57.6% of the other firms that they follow. Panel B tests whether ranking as a HQ analyst in the other firms in year t-1 can predict an analyst's forecasting performance in year t over and above the HQ classification in year t-1 in the same firm. We estimate two probit models where the dependent variable is the HQ analyst indicator in firm j in year t. The independent variables of interest include the HQ indicator of the same analyst in firm j in year t-1, and the *HQ in other firms* indicator that is equal to one if this analyst is also HQ in the majority of other firms she followed in year t-1. We find that analysts who were of HQ in the majority of other firms they followed in year t-1 are 5.1% (p-value<0.01) more likely to be HQ in a given firm in year t. The coefficient on the firm specific HQ designation in year t-1 remains positive and significant (p-value<0.01), consistent with Table 1. Hence, the cross-firm findings in Table 2 suggest that analysts' forecasting performance transcends across stocks they follow and, further, that the HQ analysts are indeed better than their peers in a persistent manner.

Our finding that the HQ analysts as a group tend to provide more accurate earnings forecasts than the consensus leads us to a question whether investors should always heed the HQ analysts' forecasts and disregard the consensus forecast. The extent to which the average of the HQ analysts' forecasts is more accurate than the consensus may depend on the number of the HQ analysts. While Appendix A provides a more formal derivation, the intuition is simple. The greater the number of forecasts (analysts following the firm), the smaller is the forecast error and, hence,

[8] If the number of high and low quality rankings of the analyst in the other firms is the same, this analyst-year-firm observation cannot be categorized as either high or low quality in the other firms and, thus, is excluded from this analysis (approximately 9% of the observations).

[9] There are slightly fewer HQ analysts than LQ analysts in year t-1 because in firms with an odd number of analysts, the analyst at the median is designated as a LQ analyst.

the more accurate is the consensus. A HQ analyst has on average a smaller forecast error to begin with, and the forecast error of the HQ analysts as a group also decreases in the number of these analysts for the firm. As the number of the HQ analysts increases, investors are more likely to obtain a more accurate forecast than the consensus by following the average estimate of the HQ analysts. Table 3 empirically investigates this issue. It provides statistical tests comparing the absolute SUE of consensus with the absolute SUE of HQ analysts.[10] We find that as the number of HQ analysts following the firm increases, the HQ analysts as a group eventually become more accurate than the consensus, confirming the prediction of the analysis in Appendix A. Further, when the number of HQ analysts is four or greater the absolute forecast error of the HQ analysts is smaller than the consensus. It is in these firms that investors seeking more accurate earnings forecasts should switch from using the consensus forecast to the average of the HQ analysts' forecasts. For the same reason, we use the sample of firms with four or more HQ analysts when we examine whether the market is aware of differential analyst quality in the analysis of recommendation changes, forecast dispersion, and the PEAD.

4. Is the market aware of high quality analysts?

The previous section demonstrates that with the HQ analysts' earnings forecasts, one can generate an earnings forecast superior to the consensus forecast. We next test whether the market is aware of this empirical regularity. To this end, we analyze the immediate market reaction to three earnings surprise measures based on the consensus, HQ, and LQ analysts' average forecasts. We examine whether the reaction to the earnings surprise based on the mean forecast of the HQ

[10] We note that because some analysts can stop covering the firm after year t-1 and new, thus unranked, analysts can commence coverage, the numbers of HQ and LQ analysts in year t can become too small or too different relative to each other (e.g., five HQ and one LQ analyst or vice versa), leading to small sample bias and a lack of robustness when the average accuracies of the HQ and LQ analysts as groups are compared in the firm-level analysis. To mitigate this concern, we restrict the sample in all firm-level analyses (Tables 3-8) to firms in which the numbers of HQ and LQ analysts are not too different in year t. Specifically, we require that neither of these groups exceeds 75% of all analysts providing forecasts for a given announcement.

analysts is greater than the earnings surprise based on the consensus forecast and, separately, surprise based on the mean forecast of the LQ analysts.

Table 4 reports regression results in which the dependent variable is the buy-and-hold cumulative abnormal return (BHAR) for the earnings announcement day and the following trading day, based on the 4-factor model (Fama and French, 1993; Carhart, 1997). The main variables of interest are the coefficients on the SUE based on the consensus, HQ analysts, and LQ analysts. While a greater reaction to the consensus than the HQ analysts' forecast can be expected and is efficient for firms with fewer than four HQ analysts, it is not so for firms with four or more HQ analysts according to Table 3. Table 4 shows that the reaction to the SUE of the consensus is greater than the reaction to the SUE of HQ analysts irrespective of the number of HQ analysts following the firm, with a highly statistically significant differences between the coefficients of 0.103 based on the chi-squared test in the full sample and a slightly smaller difference of 0.060 in the sample of firms with four or more HQ analysts. The coefficient on the SUE of HQ analysts is greater and statistically different than the coefficient on the SUE of LQ analysts, which suggests the market is aware to some extent of the accuracy differences among analysts. Overall, the results indicate that the market does not sufficiently recognize quality difference because it reacts to the consensus forecast even when the average of the HQ analysts is more accurate.

The finding that the market does not give sufficient weight to the HQ analysts' forecasts may have meaningful economic implications. To gauge their magnitude, we first construct a simple measure earnings surprise based on the difference between the HQ analysts' mean forecast and the consensus forecast, labeled *predicted surprise*. The intuition is to replace the actual earnings in the SUE formula with the HQ analysts' mean forecast,

$$Predicted\ surprise = \frac{Avg.Forecast^{HQ} - Avg.Forecast^{consensus}}{Price_{t-1}}, \tag{1}$$

so that *predicted surprise* defined this way can be used to predict the SUE of consensus. Investors aware of the quality differences among analysts can use this measure to predict the immediate market reaction to earnings announcements. Given that the HQ analysts are more accurate than the consensus, but the market over-weights the consensus forecast when it reacts to earnings surprise, one can expect positive (negative) abnormal returns to the earnings announcement when the mean forecast of HQ analysts is greater (smaller) than the consensus. A simple trading strategy is to buy (short) the stock when the predicted surprise is positive (negative). Additionally, we consider a definition for *predicted surprise* equal to the normalized difference between the HQ and LQ analysts' mean forecasts:

$$Predicted\ surprise = \frac{Avg.Forecast^{HQ} - Avg.Forecast^{LQ}}{Price_{t-1}} \quad (2)$$

It also reflects the idea that the market does not sufficiently react to the HQ analysts' estimates and, thus, overweighs the LQ analysts' estimates.

We report the empirical results in Table 5. The trading strategy is based on two variations of the signal based on *predicted surprise*: *Positive predicted surprise* and *Big predicted surprise* indicators. *Positive predicted surprise* is equal to one if *predicted surprise* is positive and zero otherwise. A stronger signal, *Big predicted surprise* indicator, is one (zero) depending on whether *predicted surprise* is above (below) the median of its positive (negative) values in the previous year. Using the values of *predicted surprise* measured in the previous year ensures our analysis is out-of-sample. We regress the two-day cumulative BHAR on each of these indicators and control variables. The coefficients on the predicted surprise indicators are positive and significant in all specifications, reaching 0.0019 in column (3), and the statistical significance of the predicted surprise indicators is greater for the definition based on the difference between the HQ and LQ analysts' forecasts. The last line of the table reports the two-day abnormal returns of a trading strategy that is long if the predicted surprise indicator tested in that column is equal to 1 and short if it is equal to 0. All returns are statistically significant and reach 0.24% for *Big predicted surprise* based on the difference between the HQ and LQ analysts' forecasts. These returns can be high enough relative to transaction costs (Novy-Marx and Velikov, 2016) because *predicted surprise* achieves its highest values when the HQ analysts are most accurate, i.e., in firms with many analysts following, implying relatively small transaction costs for these larger firms.

The overall conclusion from Tables 4 and 5 is that the market seems to overreact to the actual earnings' deviations from the consensus compared to deviations from the HQ analysts' average estimate. Another way to view these findings is that the market overreacts to the LQ analysts and underreacts to the HQ analysts. The simplicity of this strategy and the magnitude of its abnormal return suggest that fixation bias in the case of consensus forecast may be rather pervasive and deep rooted in investors' behavior.

5. Stock recommendations, forecast dispersion, and implications for the PEAD

The persistence in analysts' forecasting performance through time and across stocks suggest that HQ analysts have superior ability and, thus, it is possible that they issue superior stock recommendations. Further, given the HQ analysts are better in forecasting future earnings, the dispersion in their forecasts may contain more relevant information than the dispersion of the forecasts of the entire population of analysts following the firm. We empirically examine these predictions and their implication to the PEAD phenomena.

5.1. Stock recommendations

We examine the extent investors are aware of differences in analysts' forecasting ability when they react to recommendation revisions. We begin with analyzing whether the HQ analysts' recommendation revisions elicit stronger immediate market reaction. Then we address the key question of the speed with which prices incorporate any superior information contained in the HQ analysts' recommendations.

A recommendation is an integer between 1 and 5, where 1 is "strong buy", 5 is "strong sell", and 3 is "hold". For the ease of interpretation, we measure recommendation revisions as the negative of the current recommendation of the analyst minus the previous recommendation of the analyst, so that a positive (negative) recommendation revision is an upgrade (downgrade). The recommendation revision for the firm is the average of individual analysts' revisions. Our sample consists of recommendation revisions made during the month of the annual earnings announcement. This has several advantages. First, the month with the annual announcement has the most information for analysts to process during the year because of the announcement, information in earnings announcements and potential subsequent mispricing have a major influence on recommendation revisions (Yezegel, 2015), and analysts of both quality types face the same information set, in contrast to recommendations at random dates during the year. Hence, this setting allows for a direct and uniform link between analyst quality and recommendation quality. Second, this month investors obtain an updated analyst quality classification, as of year t rather than t-1. This also allows for a slighter greater number of firm-year observations in Table 6 than in Table 4 (columns 4-6) because when the ranking is based on year t the sample does not require that at most 75% of forecasts are made by one analyst type. Finally, and perhaps most importantly, the earnings announcement month is when we find that the market is fixated on the consensus forecast and does not produce a proper recognition to the HQ analysts; thus, we expect this pattern to be prominent for the HQ analysts' stock recommendations during this month as well. These considerations make the earnings announcement month the best time frame to examine whether the market efficiently incorporates its knowledge on analyst quality into reaction to recommendations.

The regressions of the immediate market reaction on the HQ and LQ analyst indicator cross-terms with individual analyst recommendation revisions (untabulated) yield results that are consistent with the finding for earnings announcements in Table 4—that investors appear to recognize, at least to an extent, the more accurate forecasters by reacting stronger to the HQ analysts' recommendation revisions. However, the important question is whether the market *fully* incorporates quality differences into prices at the time of the recommendation revision announcements. Table 6 shows that this is not the case. We examine this issue by regressing equity returns in the calendar month after the month of the recommendation revision on the interaction of the revision with the HQ and LQ variables, respectively. Our analysis of equity returns in the calendar month following a recommendation revision month allows for using all the revisions during that month because investors have learned the updated analyst quality classification by the end of the month. The investment delay from the revision date to the end of the revision month provides the investors with sufficient time to react to the revision and makes our monthly return estimates conservative because such a delay reduces the returns (Barber et al., 2001).[11]

The regression results in Table 6 reveal that the cross-term of recommendation revisions with the HQ analyst indicator is positive and significant, while the cross-term with the LQ analyst indicator is not. A one step recommendation upgrade by the HQ analysts during the month of the earnings announcement predicts the firm's stock return will be 0.25% greater next month. The HQ analysts' recommendation revisions are also the only ones generating value for investors that are aware of and utilize analyst quality differences in their investment decisions. To this end, we examine the returns of a long-minus-short strategy in the month following the revisions, where the long (short) position is in the firms for which the mean recommendation revision is positive

[11] The predicted monthly return results in Table 6 are unaffected by using the subsamples of recommendation revisions made before the earnings announcement, coinciding with the announcement, and after the announcement during the announcement month. We also reach the same conclusions by conducting event-time analysis for returns over the periods (2,32) and (2,62) days following the revision.

(negative) during the earnings announcement month. In particular, with the HQ analysts' recommendation revisions, the resulting return almost doubles, to 0.85%, and is highly statistically significant contrasting to recommendation revisions by all analysts, for whom the trading strategy yields a statistically not significant 0.36% (untabulated). Further, trading based on the LQ analysts' recommendation revisions does not generate statistically significant returns. Overall, we find that the predictable relation between analyst recommendation revisions and equity returns in the subsequent month is driven by the recommendations of the HQ analysts. Hence, our findings suggest that analyst quality measured based on earnings forecasts transcends to recommendation revisions, and the market does not fully incorporate differences in analyst skill. These conclusions are entirely consistent with the notion that treating all analyst as equal can lead to inefficient pricing.

5.2. Analysts' forecast dispersion

Analysts' forecast dispersion has been widely used as a proxy for uncertainty about firms' future prospects. We conjecture that just as the HQ analysts' superior earnings forecasts and recommendations indicate they have superior information concerning firm value, those analysts' forecast dispersion also contains more accurate information about future uncertainty. We examine whether disagreement about the firm's prospects among the HQ analysts (relative to the disagreement among all analysts) is a superior predictor of uncertainty surrounding the firm's future performance, measured by future return volatility.

Table 7 reports regression results of the returns' standard deviation during the month following the firm's annual earnings announcement month on the standard deviation of analysts' forecast errors before the earnings announcements. To avoid stale forecasts and make forecasts more comparable in terms of their proximity to the announcement, we use only forecasts in the 60 days prior to the announcement.[12] This explains the sample size reduction after we apply the requirement stated in section 3 that neither of HQ nor LQ analyst forecasts exceed 75% of all forecasts for a given announcement. We consider separately the dispersion of forecasts for all analysts, the HQ analysts, and the LQ analysts, whose indicators are the variables of interest. The HQ analysts' forecast dispersion is statistically significant, while the LQ analysts' forecast dispersion is not. The dispersion for all analysts, which combines the HQ and LQ analysts, is just marginally significant as a result. These findings suggest that only the HQ analysts' forecasts capture variation in uncertainty, which is associated with future equity volatility, over time in a given firm.[13]

5.3. Post-earnings announcement drift

We draw from several studies in the theoretical and empirical literature to generate our hypothesis that the PEAD anomaly should be greater during periods when the HQ analysts' forecast dispersion is high. A theoretical model in Abarbanell, Lanen, and Verrecchia (1995) predicts that when the dispersion in the consensus is high investors place less weight on the forecasts relative to their private information, resulting in investors reducing their response to earnings surprise. We take this argument further and note that as investors receive more information about the firm over time, they will be able to react to the earnings news more fully, resulting in a PEAD. The few prior empirical studies on uncertainty and the PEAD can also be helpful to motivate our hypothesis. Hung, Li, and Wang (2014) find that exogenously reduced

[12] The length of the forecast dispersion measurement period varies in the literature significantly. For instance, it can be one month (Krishnaswami and Subramaniam, 1999), four months (Zhang, 2006b), six months (Babenko, Tserlukevich, and Vedrashko, 2012), and up to one year since the previous earnings announcement (Diether, Malloy, and Scherbina, 2002). Our choice of 60 days is to ensure that we use only the annual earnings forecasts made after the last quarterly earnings announcement. Our results are not affected if we use a different period length.

[13] In untabulated results without firm fixed effects, the coefficients on both HQ and LQ variables are positive and significant, which implies both analyst types recognize differences in uncertainty across firms, though to a different extent. The chi-squared tests for the difference in coefficients between the regressions indicate that the coefficient on the dispersion of the HQ analysts is greater than that of the LQ analysts, with the p-value of 0.3%.

information uncertainty (due to a switch to different accounting rules) leads to a lower PEAD. Zhang (2006a) argues that investors underreact more to public information when uncertainty is high and finds that analysts' forecast dispersion predicts the price drift following analysts' forecasts. Francis et al. (2007) find a positive relation between the PEAD and uncertainty measured with the unexplained portion of working capital accruals. Hence, we examine whether the PEAD is indeed associated with a greater dispersion of the HQ analysts' forecasts, which measure firm-level uncertainty according to the previous subsection.[14] We calculate the PEAD using the calendar-time approach. To make our PEAD results comparable with the standard PEAD measurement in the literature, we use the consensus earnings surprise to assign announcing stocks to the long (short) portfolio each month if earnings surprise is positive (negative). The stocks are then held in the portfolios for horizons from 1 to 11 months to avoid overlapping with the following annual earnings announcement. The monthly PEAD is the alpha from regressing the monthly value-weighted portfolio returns on the four Fama-French-Carhart factors.[15] The cumulative PEAD is the monthly alpha multiplied by the number of months for which the stock is held in the long or short calendar time portfolio. The resulting relation between forecast dispersion and the PEAD is presented in Figure 2 and Table 8.

Figure 2 reports the long PEAD portfolio return minus the short PEAD portfolio return for the sample of announcements with high uncertainty, defined as announcements for which the HQ analysts' forecast dispersion is greater than that of all analysts, the full sample, and the low uncertainty sample, in which the HQ analysts have lower forecast dispersion than all analysts. The high uncertainty PEAD is clearly above the full-sample PEAD, and the low uncertainty PEAD is below the full sample PEAD. Table 8 reports the statistical significance of the returns on long,

[14] We note that because of our sample's requirement that four or more HQ analysts follow the firm, the sample consists of relatively large firms. This is an advantage for analyzing the PEAD because it makes the illiquid stock explanation to the PEAD (Sadka, 2006) not affect our findings.
[15] We obtain similar results using equal-weighted portfolios.

short, and long minus short strategies for the subsamples with high and low uncertainty announcements. We find that the low uncertainty PEAD (approximately 60% of the announcements) is not significant except for the short portfolio and only weakly significant for the long-minus-short portfolio at the 11-month horizon. In contrast, when the forecast dispersion of the HQ analysts is greater than dispersion of all analysts the long-minus-short PEAD is highly significant for all horizons except for 4- and 5-month horizons and especially large and statistically significant for the long portfolio. Overall, the PEAD is produced primarily during the periods of high information uncertainty determined based on the relation between forecast dispersions of the HQ analysts' and all analysts. Further, uncertainty is better proxied by the HQ analysts' forecast dispersion than all analysts' forecast dispersion, which is the standard measurement method in the literature (e.g., Diether, Malloy, and Scherbina, 2002).

6. The content of the HQ analysts' information output at the aggregate

Given our findings about the predictability of future equity returns (via the HQ analysts' recommendation revisions) and volatility (via the HQ analysts' forecast dispersion) at the firm level, we now consider whether these relations can be aggregated to the industry and market levels. Prior studies, e.g., Boni and Womack (2006), do not find predictability of relative returns for industries based on consensus recommendation changes. Park (2005) finds a contemporaneous non-causal relation between forecast dispersion and market volatility. Kadan et al. (2012) find that industry recommendations predict industry returns and that some analysts in large brokerage houses incorporate cross-industry information into their firm recommendations, which results in industry return predictability for their aggregated firm recommendations. Our method differentiates analysts by their ability, which allows for analyzing firm recommendations by all analysts. In Table 6, we find that recommendation changes by the HQ analysts predict stock returns at the firm level, so that averaging the change in recommendations across firms results in a testable relation between HQ analysts' average recommendation changes and future market returns. The aggregation argument works similarly for forecast dispersion. In Table 7, a relatively high forecast dispersion of the HQ analysts implies they are uncertain about the firm's prospects, and aggregating analysts' dispersion across all firms in the market results in a dispersion measure that reflects the degree of uncertainty in the market.

In Table 9, Panels A and B, we report the estimation results on the relation between revisions in stock recommendations and *future* industry and market returns, respectively. To this end, each month, we average recommendation changes of the HQ, LQ, and all analysts in each firm and then across all firms in each 2-digit SIC industry, resulting in an industry-month panel in Panel A. We also aggregate across all firms regardless of their industry affiliation, resulting in one monthly time series for the market in Panel B. The dependent variables are the monthly value-weighted industry returns and value-weighted market returns in the month following the month with the annual earnings announcement. The recommendation change variables are aggregates of the recommendation change variables used in Table 6. *All analysts' mean recommendation change* is the mean of all recommendation changes during the month in which the firm's earnings are announced. *HQ (LQ) mean recommendation change* are analogous variables that are based only on recommendation changes of the HQ (LQ) analysts. The control variables included follow Li, Ng, and Swaminathan (2013), which we interact with industry fixed effects in the industry-level regressions. In addition, we control for the previous month's industry or market return to account for the possibility of a momentum in these returns.

The regression results in Panel A reveal that the recommendation revisions by HQ analysts predict future industry returns, while the revisions by the LQ analysts do not. Specifically, the coefficient on HQ analysts' recommendation revision is positive and significant indicating that the recommendation revision by HQ analysts are not fully internalized by the market because they can predict the industry return in the following month. A cross-industry arbitrage strategy that has a long (short) position in the industries for which the HQ analysts' mean recommendation revision is positive (negative) yields a highly statistically significant average return of 0.66% in the month following the announcement month. The strategy based on the LQ analysts does not yield a statistically significant return. The predictive power of the HQ analysts' recommendation revisions is strong enough to produce a statistically significant coefficient for revisions by the full set of analysts and a statistically significant long-short industry strategy return. Finally, we report the results on calendar-time alphas based on regressions of these long-minus-short monthly returns on the market index. Only the HQ analysts' recommendation revisions generate a statistically significant alpha of 0.57% per month.

In Panel B, we repeat this analysis at the market level. In the regressions of market returns on mean recommendation revisions, the mean HQ analysts' recommendation revisions predict the market return next month, while the LQ analysts' coefficient is not statistically significant. The coefficient on all analysts is statistically significant, albeit smaller than the HQ analysts' coefficient, because it combines recommendations by both HQ and LQ analysts. We also provide the results of a trading strategy in the market index based on mean recommendation revisions. Because the market-level data is a monthly time series, the long and short trading signals are based on the historical variation in monthly mean recommendation changes as follows. If the mean recommendation revision this month is greater (smaller) than the median of the monthly mean recommendation revisions over the previous 24 months, i.e., the current recommendation revisions are more optimistic than they were in the recent past[16], we buy (short) the market value-weighted index and hold it for one month. We regress the monthly returns of this strategy on the market

[16] The results are unaffected by selecting a longer window up to five years. The shorter, 24-month window we use minimizes the number of months lost to initialize this out-of-sample analysis, while providing enough observations for a robust distribution of monthly mean recommendation changes.

return and report the alphas for all, HQ, and LQ analysts' recommendation revisions. Only the HQ analysts' recommendations produce a statistically significant alpha, 0.51% per month. These findings suggest that the HQ (LQ) analysts' recommendations are (are not) informative about the future state of the market.

Together, consistent with the firm-level findings, the industry and market results indicate that the market does fully incorporate the superior ability of the HQ analysts or, alternatively, does not fully distinguish among analysts based on their quality, which results in inefficient prices at the industry and macro levels.

Finally, in Panel C of Table 9, we examine the second market-level aggregate relation—whether the HQ analysts' forecast dispersion aggregated across all firms is associated with uncertainty about the economy and, thus, with current and/or future market volatility, which we measure with the VIX index. Because the HQ and LQ analysts' forecast dispersions have firm-specific predictive power according to the estimation results without firm fixed effects in Table 7, they need to be normalized to make them comparable across firms and years and, thus, usable as potential measures of market uncertainty.[17] We define a dispersion ratio for the HQ analysts equal to the standard deviation of the HQ analysts' forecasts divided by the standard deviation of all analysts' forecasts (i.e., the consensus). The dispersion ratio for the LQ analysts is defined analogously. A market-level measure of analyst uncertainty is created by firm value-weighting the forecast dispersion or dispersion ratios across firms each month.

Panel C reports regressions of monthly VIX returns (percentage changes in the VIX)[18] on the dispersions of all, HQ, and LQ analysts' forecasts and the HQ and LQ analysts' dispersion ratios measured before the earnings announcement in the previous month. The results indicate that a higher dispersion ratio for the HQ analysts predicts a greater VIX return in the following month.

[17] In the panel regression model in Table 7, we control for the firm-specific component using firm fixed effects.
[18] We obtain very similar results using the VIX values in place of returns on the VIX.

Consistent with our argument, not normalized measures of forecast dispersion indeed do not forecast the VIX. The LQ dispersion ratio also does not capture the level of uncertainty in the market. The last line of Panel C reports the performance of a trading strategy for the VIX index identical in design to the trading strategy for the market index. If the forecast dispersion or dispersion ratio in the corresponding column in a given month is above (below) the median of its distribution during the previous 24 months, we buy (short) the VIX index and record its percentage change over the next month. To benchmark the performance of this strategy, we regress its return on the market return and report the alphas. The only statistically significant (p-value<0.01) abnormal return is for the trading signal based the HQ analysts' dispersion ratio; the strategy yields a 3.32% alpha per month. We conclude that when the HQ analysts on average tend to be more uncertain than all analysts about the prospects of firms they follow, investors should expect an increase in market-wide uncertainty and volatility.

7. Robustness

Our robustness analysis considers whether our results are sensitive to different definitions of the HQ and LQ analysts. The definition used throughout the paper, splitting analysts into two groups at the median based on the accuracy of their closest estimate to the annual EPS announcement, is only one of many ways of generating an alternative to the consensus forecast and ranking analyst quality. Other ranking methods include using a different forecast accuracy cutoff between the two groups, giving more weight to forecasts made closer to the announcement, introducing observable analysts' characteristics to improve the classification of the HQ and LQ analysts, and ranking based on multiple year forecasting performance. We examine the first two possibilities in this section. The last two options would not improve our ranking procedure. Brown (2004) finds that a model based on observable analysts' characteristics has similar power to capture analyst accuracy persistence as the measure we use. If the ranking relied on a longer history of analysts' estimates, for example, we would face biased persistence tests, as Carpenter and Lynch (1999) find that a single-year past performance ranking is less likely to suffer from survivorship bias that induces spurious performance reversals if the survival depends on multi-year past performance, which is the case for analysts. With a multi-year ranking, we would lose forecasts of analysts that could not be ranked because of their short history of following the firm. Such observation dropping would be especially problematic for comparing the HQ analysts to the consensus forecast, whose figure used by the market includes the forecasts of all analysts. The overall advantage of our approach is that it provides us with a high degree of power and preserves observations to avoid potential biases.

We start with analyzing how alternative divisions of analysts into the high and low quality groups affects the persistence in analyst forecasting performance. The ranking procedure sorts analysts in a given firm-year based on their absolute forecast error. In general, HQ analysts are those who are ranked in the top p percent of analysts, while LQ analysts are those in the bottom $(1-p)$ percent. If analysts' forecasting performance were uncorrelated across years, the fractions of analysts who preserve their ranking in two consecutive years as HQ and LQ would be p^2 and $(1-p)^2$, respectively, or $p^2+(1-p)^2$ of all analysts.

Figure 3 plots the fraction of analysts that retain their rankings in consecutive years and the expected fraction assuming no performance correlation across years. We find that with almost all cutoff values of p, the actual fraction of persistent forecasting performance is above the expected fraction, and all these differences are statistically significant (p-value<0.01). For example, when we classify the top 10% of analysts following a firm in a given year as high quality (p=10%) and the bottom 90% as low quality, the expected fraction given random assignment is $0.9^2+0.1^2=0.82$. The figure shows that the actual fraction is greater than that at 0.843. The exception is the relaxed definition of HQ analysts as the best 95%. Nevertheless, the overall finding is that for almost all of the cutoff values, there is a sizeable persistent component, so that it should makes little difference for accuracy persistence which exact cutoff value we choose to partition HQ and LQ analysts.

We next analyze how our findings are affected by different cutoffs dividing analysts into the HQ and LQ categories. We define the HQ analysts symmetrically around the definition used in the paper, as those in the top 70% (the HQ/LQ proportion is 70%/30%) and the top 30% (the HQ/LQ proportion is 30%/70%) of forecast accuracy, respectively. Because the rankings are done in year t-1, and the proportion of HQ and LQ analysts following the firm can change dramatically by the year t announcement, we follow the a similar sample restriction as the one used in the paper to avoid small sample bias in the groups. It requires that at year t, the proportion of HQ and LQ analysts does not change by more than a 20% margin. This would assure we avoid a situation with too few analysts in both HQ and LQ analysts groups. For the definition of the HQ analysts as the top 70% in year t-1, they are in the range of 50% to 90% of all analysts covering the firm in year t; and when the HQ analysts are the top 30% in year t-1, they are in the range of 10% to 50% in year t.

The third alternative definition for the HQ/LQ analysts is based on value-weighted absolute forecast errors. The value-weighted absolute forecast error is computed based on all, rather than the most recent, forecasts by the analyst during the 300 days prior to the annual earnings announcement. It is computed as follows,

$$VWFE_t = \frac{FE_{300} \times d_1 + \sum_{j=2}^{n}(FE_j \times d_j)}{300} \tag{3}$$

where $VWFE_t$ is the value-weighted absolute forecast error of the analyst in year *t*; FE_{300} is the absolute forecast error based on the forecast outstanding on the 300th day prior to the earnings announcement; d_1 is the number of days this forecast is outstanding (from the 300th day prior to the earnings announcement to the earliest of the earnings announcement day or the following earnings forecast revision day); *n-1* is the number of estimates issued by the analyst between the

299th day prior to the earnings announcement and the earnings announcement day; FE_j is the absolute forecast error of forecast j; and d_j is the number of days the forecast has been outstanding. The advantage of the value-weighted measure is that it captures the analyst ability over a four-quarter period, instead of a one-time shot just before the annual earnings announcement. However, the disadvantage of the value-weighted measure is that it requires us to have an analyst's forecast at least 300 days prior to the annual earnings announcement day, in contrast to the closest forecast error measure used throughout the paper that can generate rankings for any analyst following the firm. Further, we have no knowledge whether the lack of early forecasts is due to the analyst's poor ability to forecast or a benign reason, such as common practice in the industry or the firm. Using the value-weighted absolute forecast error reduces the sample by approximately 65%, which reduces the power of our empirical analysis.

Table 10 repeats the key firm and industry level analyses of the paper with these three alternative measures. We first check whether the consensus forecast is inferior to the average of HQ forecasts (Table 3) and whether the market is fixated on the consensus and fails to incorporate this into prices at earnings announcements (Tables 4 and 5); then we retest other consequences of ignoring analyst quality differences—for recommendation revisions (Table 6) and the dispersion of forecasts (Table 7).

If we were to define the HQ as the top 70% of analysts, the cutoff for when the HQ analysts' average forecast accuracy is superior to the consensus forecast is with 3 or more HQ analysts in column (1) of Table 10 (instead of 4 or more in Table 3). The results in Table 10 corresponding to Tables 4 and 5 with this HQ analyst measure show that the market reacts more to the consensus than to the average of HQ analysts' forecasts, and there is a mispricing at the announcement associated with that. The results replicating Tables 6, 7, and 9.A are affected little and remain highly statistically significant. In fact, the trading strategy return in column (4) of Table 10 is greater than the corresponding return in column (2) of Table 6.

When we define HQ as the top 30% of analysts, the cutoff for when the average forecast of the HQ analysts is more accurate than the consensus increases to 6, a substantial number of HQ analysts in a firm. The reaction to the consensus is still greater than that of the average of the HQ analysts, and the difference is greater than in the line above, but it is not statistically significant. Nevertheless, the trading strategy results in column (3) are still weakly significant. The findings of Tables 6, 7, and 9.A hold.[19]

Finally, with the value-weighted measure we have a small sample and do not find statistical significance for the tests whether the market reacts more to the consensus than that the average of the HQ analysts' forecasts. We note, however, that if the market were not fixated on the consensus, it would react more to the average of HQ estimates than the consensus, which is not the case even for this definition. The trading strategy return in column (3) corresponding to Table 5 is also not statistically significant, but the results of Tables 6, 7, and 9.A hold. Overall, we interpret Table 10 as providing robust evidence that changes to the definitions of the HQ analysts have a rather small effect on the evidence of mispricing. The few differences in Table 10 from the earlier tables are unsurprising in that they follow the changes in sample size due to the alternative definitions of the HQ analysts relative to the definition used in the paper.

8. Conclusion

[19] The tradeoff between the first two alternatives is the following. If the HQ group is large (for example, top 70%), fewer analysts following the firm are needed for the average of HQ forecasts to be more accurate than the consensus, while, at the same time, the difference between the consensus and the average of the HQ forecasts, which represent a large fraction of all analysts, becomes smaller. On the other hand, a narrow definition of the HQ analysts (for example, top 30%) requires a greater number of analysts following the firm until the average of the HQ forecasts becomes more accurate than the consensus, but the difference between the two forecasts is relatively large. Hence, defining the HQ analysts as the top 50%, as we do in the paper, balances this tradeoff.

Our results show that analysts' forecasting ability tends to persist over time and across firms they follow. Therefore, it can be considered an analyst characteristic. As a result, we find that when firms are followed by a sufficient number of HQ analysts the consensus forecast tends to be less accurate than the average forecast of the HQ analysts. Disregarding the LQ analysts' forecasts and other information they generate, such as stock recommendations and forecast dispersion, can be beneficial for investors.

However, the market does not seem to be aware of these shortcomings of the consensus: it reacts stronger to deviations from the consensus earnings estimate than to deviation based on the HQ analysts' forecasts. We demonstrate the inefficient market reaction with a simple trading strategy for the immediate market reaction based on the HQ analysts' forecasts. Following the principle that acknowledges differences in analyst ability allows for uncovering other mispricing phenomena using stock recommendations and forecast dispersion. Just as parsimonious models tend to perform well out-of-sample (e.g., DeMiguel et al. (2009) on the underperformance of the mean-variance optimization analysis relative to equal investment across assets), our rather simple analyst quality measure allows for uncovering predictive relations for analysts' outputs. The HQ analysts' stock recommendations and forecast dispersion predict the first two moments of firm and stock market returns. In short, the persistence of the analyst quality along multitude dimensions is not recognized by the market, resulting in inefficient pricing after earning announcements and stock recommendations changes.

Overall, our findings suggest that the market is not justified in focusing only at the consensus forecast and can utilize the forecasts and other information output of the HQ analysts. Investors' giving less or no weight to the LQ analysts may lead to improvements not only in the price efficiency but also in the security analyst market. The more talented analysts would have additional incentives to exert more effort, while the less talented analysts would move to activities in which they add value. The accuracy of forecasts and informativeness of recommendations can improve as a result.

Appendix

Let there be n_G analysts of type *G* (high quality) and n_B analysts of type *B* (low quality) following the firm. Each analyst receives an unbiased noisy signal about the true earnings μ. Analysts of type *G* receive signal $S_i^G = \mu + \varepsilon_i^G$, where ε_i^G are i.i.d. $N(0, \sigma_G)$, while analysts of type *B* receive signal $S_i^B = \mu + \varepsilon_i^B$, where ε_i^B are i.i.d. $N(0, \sigma_B)$, uncorrelated with the noise of the good analysts, and $\sigma_G < \sigma_B$. Analysts do not act strategically in that their forecasts are equal to their signals.

To obtain more accurate forecasts, closer to the true earnings μ, one would prefer the average forecast of type G analysts and ignore the forecasts of type *B* analysts if and only if the dispersion of the average signal of high quality analysts is less than that of low quality analysts:

$$var\left(\frac{1}{n_G}\sum \varepsilon_i^G\right) < \; var(\frac{1}{n_B}\sum \varepsilon_i^B) \tag{A.1}$$

This simplifies to

$$\frac{\sigma_G^2}{n_G} < \frac{\sigma_B^2}{n_B} \tag{A.2}$$

This means if a firm has relatively few high quality analysts and relatively many low quality analysts, the average forecast of the low quality analysts can be more accurate than the average forecast of the high quality analysts despite $\sigma_G < \sigma_B$. As the relative number of the high quality analysts increases, we will eventually prefer their average forecast over the low quality analysts' average forecast.

A similar logic applies to the consensus forecast, which averages across both low and high quality analysts. We should follow the average forecast of type G analysts rather than the consensus if and only if the dispersion of the average signal of high quality analysts is less than that for all analysts combined. This implies

$$\frac{\sigma_G^2}{n_G} < var(\frac{1}{n}\left(\sum \varepsilon_i^G + \sum \varepsilon_i^B\right)) \tag{A.3}$$

where n is the total number of analysts, $n_G + n_B$.

This simplifies to the following condition:

$$\sigma_G^2(1 + \frac{n}{n_G}) < \sigma_B^2 \tag{A.4}$$

The left-hand-side monotonically declines with n_G. Because the signal variances are unobserved, the model's testable predictions are based on the number of G-type analysts in the firm. As the number of high quality analysts increases, the inequality is more likely to hold, so that investors would prefer to consider the signals of only the *G*-type analysts, making it optimal to ignore the low quality analysts' and the consensus estimates.

References

Abarbanell, J.S., Lanen, W.N. and Verrecchia, R.E., 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. Journal of Accounting and Economics, 20(1), 31-60.

Aiolfi, M. and Timmermann, A., 2006. Persistence in forecasting performance and conditional combination strategies. Journal of Econometrics, 135(1), 31-53.

Ajinkya, B.B. and Gift, M.J., 1985. Dispersion of financial analysts' earnings forecasts and the (option model) implied standard deviations of stock returns. The Journal of Finance, 40(5), 1353-1365.

Babenko, I., Tserlukevich, Y. and Vedrashko, A., 2012. The credibility of open market share repurchase signaling. Journal of Financial and Quantitative Analysis, 47(5), 1059-1088.

Baker, S.R., Bloom, N. and Davis, S.J., 2015. Measuring economic policy uncertainty (No. w21633). National Bureau of Economic Research.

Barber, B., Lehavy, R., McNichols, M. and Trueman, B., 2001. Can investors profit from the prophets? Security analyst recommendations and stock returns. The Journal of Finance, 56(2), 531-563.

Bernard, V., and Thomas, J., 1989. Post-earnings- announcement drift: Delayed price response or risk premium? Journal of Accounting Research 27, 1-48.

Boni, L. and Womack, K.L., 2006. Analysts, industries, and price momentum. Journal of Financial and Quantitative analysis, 41(1), 85-109.

Brav, A. and Heaton, J.B., 2002. Competing theories of financial anomalies. The Review of Financial Studies, 15(2), 575-606.

Brown, L. 2004. How important is past analyst forecast accuracy? Financial Analysts Journal 57, 44-49.

Buraschi, A., Piatti, I., and P. Whelan, 2017, Expected Term Structures, University of Chicago working paper.

Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of finance, 52(1), 57-82.

Carpenter, J.N. and Lynch, A.W., 1999. Survivorship bias and attrition effects in measures of performance persistence. Journal of financial economics, 54(3), 337-374.

Clemen, R.T., 1989. Combining forecasts: A review and annotated bibliography. International journal of forecasting, 5(4), 559-583.

Clement, M.B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics, 27(3), 285-303.

Consensus earnings estimates, Special 12-page report, April 2013, IR Magazine, www.InsideInvestorRelations.com

Da, Z., Engelberg, J. and Gao, P., 2015. The sum of all fears investor sentiment and asset prices. Review of Financial Studies, 28(1), 1-32.

De Franco, G. and Zhou, Y., 2009. The Performance of Analysts with a CFA Designation: The Role of Human-Capital and Signaling Theories. The Accounting Review, 84(2), 383-404.

DeMiguel, V., Garlappi, L. and Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies, 22(5), 1915-1953.

Diether, K.B., Malloy, C.J. and Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. The Journal of Finance, 57(5), 2113-2141.

Ertimur, Y., Sunder, J. and Sunder, S.V., 2007. Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts. Journal of Accounting Research, 45(3), 567-606.

Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.

Francis, J., Lafond, R., Olsson, P. and Schipper, K., 2007. Information uncertainty and post-earnings-announcement-drift. Journal of Business Finance and Accounting, 34(3-4), 403-433.

Genre, V., Kenny, G., Meyler, A. and Timmermann, A., 2013. Combining expert forecasts: Can anything beat the simple average? International Journal of Forecasting, 29(1), 108-121.

Hendricks, D., Patel, J. and Zeckhauser, R., 1993. Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. The Journal of finance, 48(1), 93-130.

Hirshleifer, D., Lim, S. and S.H Teoh, 2009. Driven to distraction: Extraneous events and underreaction to earnings news. Journal of Finance 64, 2289-2325.

Hirst, D.E., Hopkins, P.E. and Wahlen, J.M., 2004. Fair values, income measurement, and bank analysts' risk and valuation judgments. The Accounting Review, 79(2), 453-472.

Hung, M., Li, X. and Wang, S., 2014. Post-earnings-announcement drift in global markets: Evidence from an information shock. Review of Financial Studies, 28 (4), 1242-1283.

Hutton, A.P., Lee, L.F. and Shu, S.Z., 2012. Do managers always know better? The relative accuracy of management and analyst forecasts. Journal of Accounting Research, 50(5), 1217-1244.

Jacob, J., Lys, T.Z. and Neale, M.A., 1999. Expertise in forecasting performance of security analysts. Journal of Accounting and Economics, 28(1), 51-82.

Jegadeesh, N., Kim, J., Krische, S.D. and Lee, C., 2004. Analyzing the analysts: When do recommendations add value? The Journal of Finance, 59(3), 1083-1124.

Kadan, O., Madureira, L., Wang, R. and Zach, T., 2012. Analysts' industry expertise. Journal of Accounting and Economics, 54(2), 95-120.

Krishnaswami, S. and Subramaniam, V., 1999. Information asymmetry, valuation, and the corporate spin-off decision. Journal of Financial economics, 53(1), 73-112.

Li, Y., Ng, D.T. and Swaminathan, B., 2013. Predicting market returns using aggregate implied cost of capital. Journal of Financial Economics, 110(2), 419-436.

Loh, R.K. and Mian, G.M., 2006. Do accurate earnings forecasts facilitate superior investment recommendations? Journal of Financial Economics, 80(2), 455-483.

Loh, R.K. and Stulz, R.M., 2011. When are analyst recommendation changes influential?. Review of Financial Studies, 24(2), 593-627.

Maines, L.A., McDaniel, L.S. and Harris, M.S., 1997. Implications of proposed segment reporting standards for financial analysts' investment judgements. Journal of Accounting Research, 35, 1-24.

Malloy, C.J., 2005. The geography of equity analysis. The Journal of Finance, 60(2), 719-755.

Mendenhall, R.R., 2002. Arbitrage risk and post-earnings-announcement drift. Journal of Business 77, 875-894.

Mikhail, M.B., Walther, B.R. and Willis, R.H., 1997. Do security analysts improve their performance with experience? Journal of Accounting Research, 35, 131-157.

Novy-Marx, R., and Velikov, M. 2016. A taxonomy of anomalies and their trading costs. Review of Financial Studies 29, 104-147.

Park, C., 2005. Stock return predictability and the dispersion in earnings forecasts. The Journal of Business, 78(6), 2351-2376.

Rubin, A., Segal, B. and D. Segal, 2017. The interpretation of unanticipated news arrival and analysts' skill, Journal of Financial and Quantitative Analysis, Forthcoming.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. Journal of Financial Economics 80, 309-349.

Sinha, P., L. Brown, and S. Das. 1997. A re-examination of financial analysts' differential earnings forecast ability. Contemporary Accounting Research 14, 1-42.

So, E., 2013. A new approach to predicting analyst forecast errors: Do investors overweight analyst forecast? Journal of Financial Economics 108, 615-640.

Sorescu, S. and Subrahmanyam, A., 2006. The cross section of analyst recommendations. Journal of Financial and Quantitative Analysis, 41(1), 139-168.

Stickel, S., 1992. Reputation and performance among security analysts. Journal of Finance 47, 1811-36.

Tonks, I., 2005. Performance persistence of pension-fund managers. The Journal of Business, 78(5), 1917-1942.

Yezegel, A., 2015. Why do analysts revise their stock recommendations after earnings announcements? Journal of Accounting and Economics, 59(2), 163-181.

Zhang, X., 2006a. Information uncertainty and stock returns. The Journal of Finance, 61(1), 105-137.

Zhang, X., 2006b. Information uncertainty and analyst forecast behavior. Contemporary Accounting Research, 23(2), 565-590.

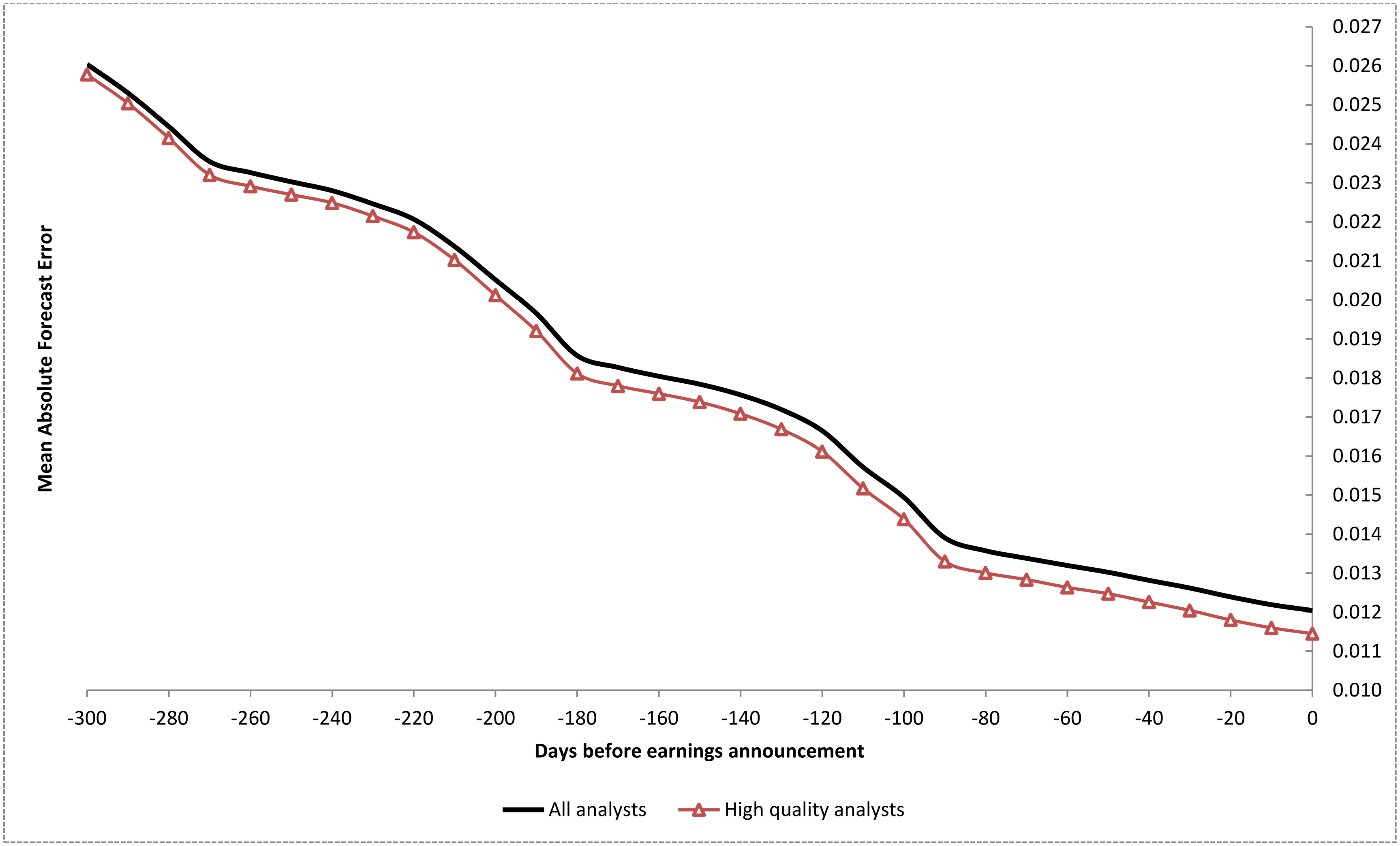

Figure 1: Absolute forecast error of analysts' estimates starting 300 days before the earnings announcement day. Absolute forecast errors at date t is calculated as the mean forecast error based on all forecasts outstanding as of day t prior to the earnings announcement date, averaged across firm-years and then averaged across firms for each pre-announcement day during 300 days prior to the announcement day. The high quality analysts are defined in Table 1.

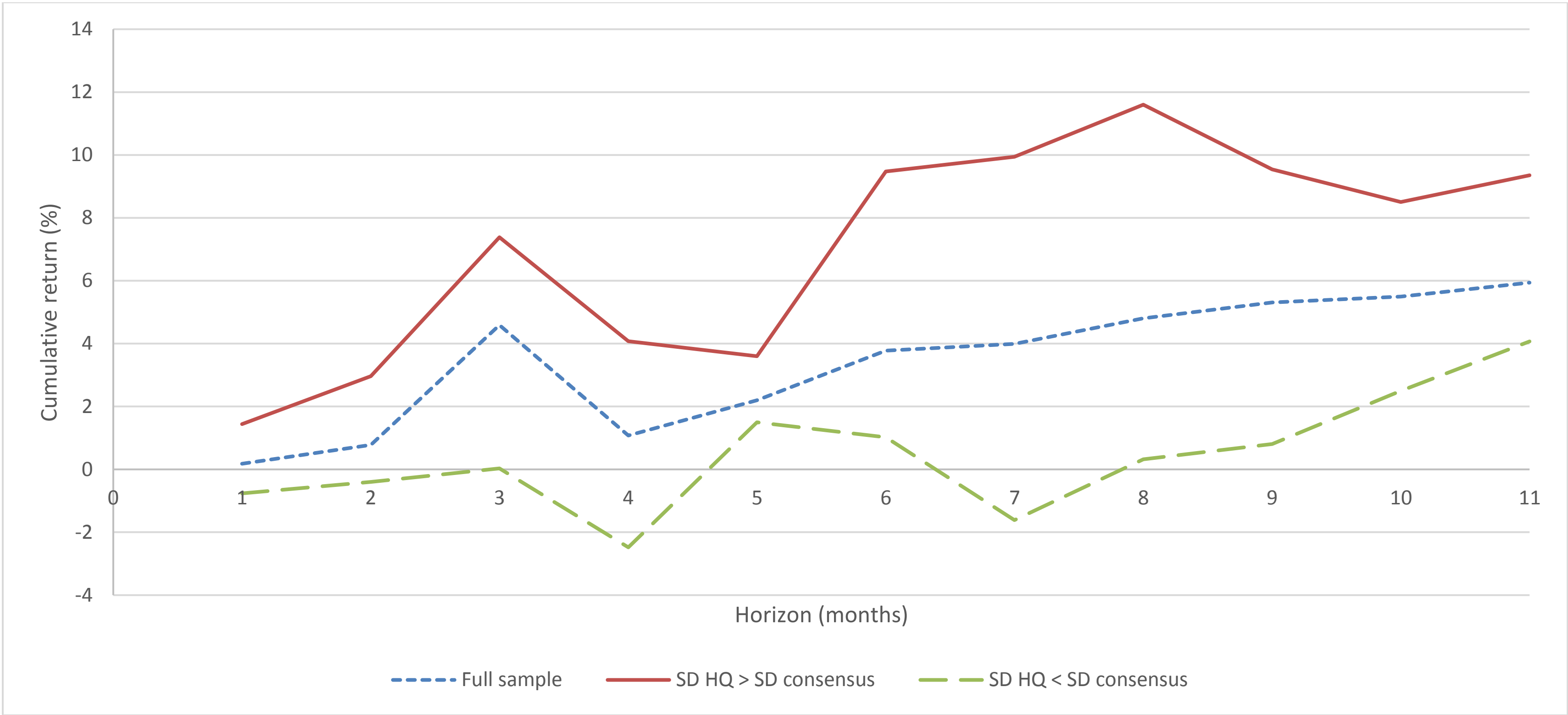

Figure 2. Cumulative post-announcement drifts depending on analysts' uncertainty
The figure shows the cumulative drift for 1 to 11-month horizons following earnings announcements. The horizontal axis is the drift's holding horizon, which is the number of months a stock is held in the calendar-time portfolios. Each month stocks enter a calendar-time long (short) portfolio depending on whether their earnings surprise is positive (negative), where earnings surprise is defined based on the consensus estimate. The long-minus-short value-weighted portfolio return is regressed on the four Fama-French-Carhart factors, and the intercept (monthly alpha) is multiplied by the portfolio's horizon to obtain the cumulative drift on the vertical axis. The graphs are for the full sample and two subsamples of firms in which the standard deviation of high quality analysts' forecasts (SD HQ) is greater or smaller than that of all analysts' forecasts (SD consensus). The high quality analysts are defined in Table 1.

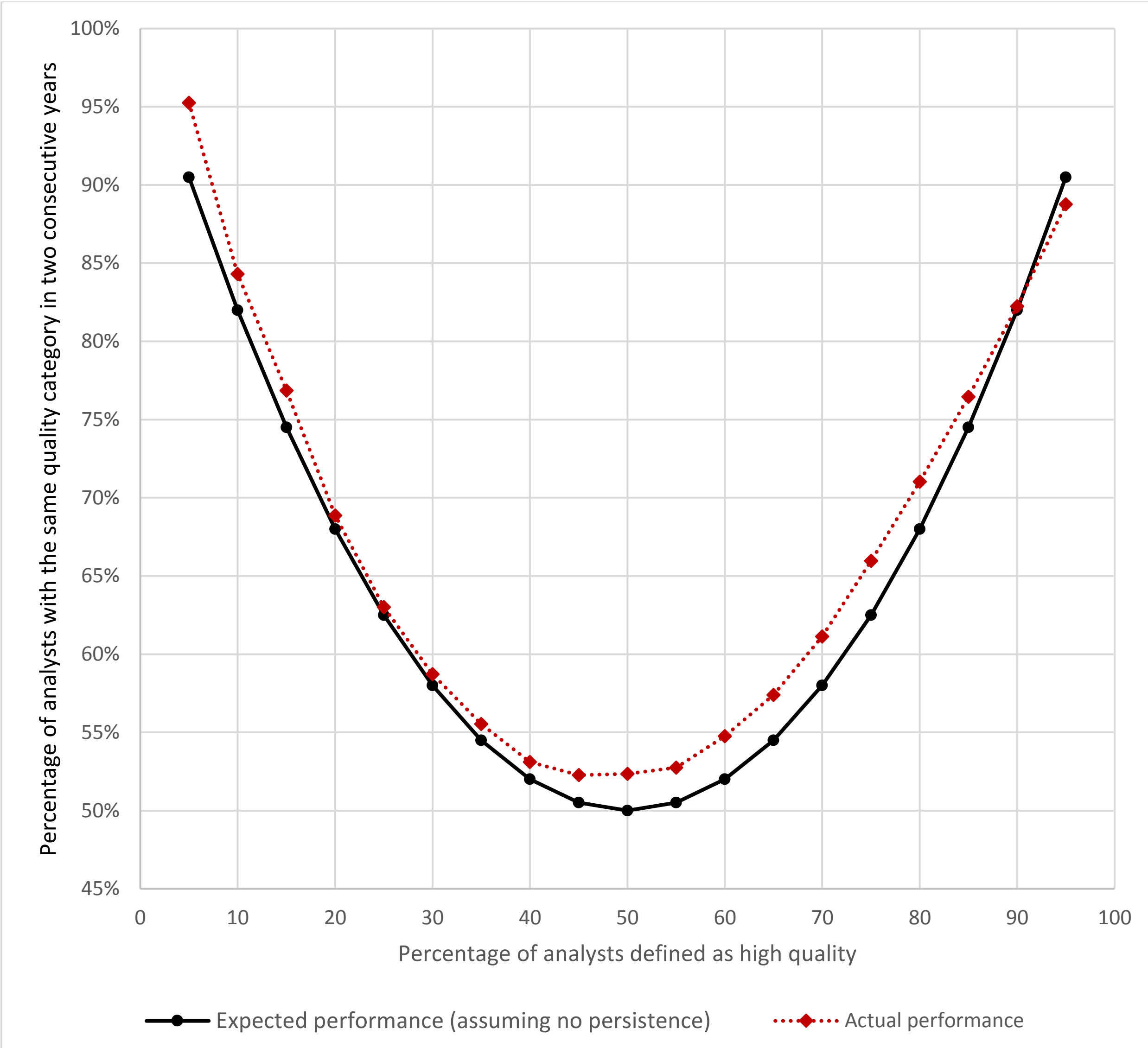

Figure 3: Persistence in analysts' forecasting performance. The figure depicts how the fraction of analysts retaining their ranking of either high or low quality of forecast accuracy in two consecutive years depends on the cutoff percentile in the definition of high quality analysts. High quality analysts are those whose closest absolute forecast errors are below the absolute forecast error at the cutoff percentile (horizontal axis) of the distribution of forecast errors for the firm's annual earnings announcement in year t-1. The closest absolute forecast error is the absolute difference between an analyst's forecast estimate closest to the earnings announcement prior to the announcement day and actual annual earnings, divided by the share price at the beginning of the calendar year. To rank analysts up to the decile precision, the sample of analysts ranked in consecurive years is constrained here only to firms that are followed by ten or more analysts. Expected performance assuming no persistence is the fraction of analysts who have the same forecast performance category in two consecutive years if their performance were uncorrelated between years.

Table 1: Analyst characteristics and forecast accuracy persistence

Panel A conducts univariate analysis for high and low quality (HQ and LQ) analysts. HQ (LQ) quality analysts are those whose closest absolute forecast errors are below (at or above) the median closest absolute forecast error for the firm's earnings announcement. The closest *absolute forecast error* is the absolute difference between an analyst's forecast estimate closest to the earnings announcement prior to the announcement day and actual annual earnings, divided by the share price at the beginning of the calendar year. The rankings in all panels and the sample in Panel A are based on firms that have at least four analysts in year t-1. *Overall tenure* is the number of years since the analyst first appeared in the I/B/E/S file. *Firm-specific tenure* is the number of years since the analyst began covering the specific firm in the I/B/E/S file. *Brokerage house size* is the number of analysts in the analyst's brokerage house. *Firm coverage* is the number of firms covered by the analyst. Panel B reports probit model results for the *HQ analyst indicator* that equals one if the analyst is of high quality and zero otherwise (columns (1) and (2)) and regressions for the analyst's closest absolute forecast error in columns (3) and (4). *Firm size* is the log of the firm's market value of equity equal to the stock price times the number of shares outstanding at the end of the month prior to the annual earnings announcement. *Annual return* is the annual return of the firm's equity over the 12 months prior to earnings announcement month. *Leverage* is the book value of total liabilities divided by the book value of total assets, and *Book-to-market* is the book value of common equity divided by the market value of equity at the end of the fiscal year. *Number of analysts* is the number of analysts following the firm. All independent variables are measured prior to the announcement date. The probit coefficients are marginal probability effects. All models include the intercept. Robust standard errors are clustered by firm. *z*- and *t*-statistics are in parentheses in the first two and last two columns of Panel B, respectively. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A. HQ and LQ analyst characteristics

Analyst or announcement characteristic	HQ analysts	LQ analysts	Difference (t-statistic)
Overall tenure	7.07	7.00	0.07^{***} (4.73)
Firm-specific tenure	3.04	2.97	0.07^{***} (8.61)
Brokerage house size	65.76	63.04	2.72^{***} (19.14)
Firm coverage	17.60	17.55	0.05^{*} (1.79)
Absolute forecast error	0.0081	0.0089	-0.008^{***} (-12.83)
Forecast error	0.00185	0.00181	0.00004 (0.66)

Panel B. Predicting analysts' forecasting performance

	HQ analyst indicator		Absolute forecast error	
	(1)	(2)	(3)	(4)
HQ analyst indicator (t-1)	0.0414***	0.0407***	-0.00073***	-0.00072***
	(25.54)	(25.13)	(-16.14)	(-15.91)
Firm size	0.0034***	0.0011***	-0.00611***	-0.00612***
	(10.52)	(3.11)	(-25.12)	(-25.15)
Annual return	-0.0003	0.0005	-0.00086***	-0.00086***
	(-0.50)	(0.88)	(-5.68)	(-5.67)
Leverage	0.0006	0.0003	0.00573***	0.00570***
	(0.37)	(0.20)	(6.58)	(6.55)
Book-to-market	0.0001	0.00003	0.00002	0.00002
	(1.40)	(0.54)	(1.48)	(1.48)
Number of analysts	0.0007***	0.0007***	0.00022***	0.00022***
	(12.40)	(14.52)	(10.78)	(10.69)
Overall tenure		0.0007***		-0.00004***
		(4.33)		(-7.12)
Firm-specific tenure		0.0023***		0.00001
		(9.66)		(1.45)
Brokerage house size		0.0001***		-0.00000
		(10.63)		(-0.26)
Firm coverage		-0.0007***		0.00003***
		(-12.32)		(8.40)
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects			Yes	Yes
Observations	485,815	485,815	485,815	485,815
Adj. R-squared			0.344	0.344

Table 2: Analyst quality across firms

The table reports how analysts' forecasting quality in one firm is related to their quality in other firms covered by the analyst in the same year, with Panel A showing the contemporaneous and Panel B showing the predictive relations. High and low quality (HQ and LQ) analysts are defined in Table 1. *HQ (LQ) indicator* is one if the analyst is ranked HQ (LQ) and is zero otherwise. *High (low) quality analyst in other firms* equals one (zero) if the analyst is of high (low) quality in the majority of the other firms the analyst follows during the year; analysts who have equal numbers of other firms with HQ and LQ performance rankings are excluded (9% of the sample). Panel B reports probit regressions predicting the HQ analyst indicator in a given firm based on analysts' HQ status indicator in the other firms in the previous year. The other independent variables are defined in Table 1. All independent variables are measured prior to the announcement date, and all specifications include the intercept. The reported coefficients are marginal probability effects. Robust standard errors are clustered by firm. *z*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A. HQ and LQ analysts' forecasting performance in other firms

		Performance in other firms		Full sample	t-statistic HQ vs. Full sample
		HQ	LQ		
Performance in this firm	HQ	54.4%	42.4%	48.3%	70.6***
	LQ	45.6%	57.6%	51.7%	

Panel B. Probit model predicting the HQ analyst status in a given firm

	HQ analyst in other firms, year t-1	HQ analyst in this firm, year t-1	Overall tenure	Firm-specific tenure	Brokerage house size	Firm coverage	Number of obs.
Marginal probability (z-statistic)	0.0531*** (32.95)	0.0387*** (23.14)					443,262
Marginal probability (z-statistic)	0.0514*** (31.83)	0.0382*** (22.84)	0.0002 (1.12)	0.0030*** (12.90)	0.0001*** (7.23)	-0.0005*** (-8.63)	443,262

Table 3: The number of HQ analysts and improvement in forecast accuracy

The table compares the accuracy of the average forecast of the high quality (HQ) analysts and the consensus sorted by the number of HQ analysts following the firm in a given year. High quality analysts are defined in Table 1. SUE of Consensus (SUE of HQ analysts) is the difference between the actual earnings and the average forecast provided by all analysts (HQ analysts) normalized by the stock price at the beginning of the year. *Accuracy improvement* is the percentage reduction from the absolute SUE of the consensus to the absolute SUE of HQ analysts. *t*-statistics is for the difference in means between the absolute SUE of consensus and HQ analysts.

Number of HQ analysts	Absolute SUE of Consensus	Absolute SUE of HQ analysts	Accuracy improvement	t-statistics Abs. SUE difference
1 or more	0.00656	0.00678	-3.31%	-8.63***
2 or more	0.00589	0.00595	-1.08%	-3.19***
3 or more	0.00514	0.00513	0.19%	0.54
4 or more	0.00461	0.00455	1.17%	2.99***
5 or more	0.00422	0.00415	1.52%	3.51***
6 or more	0.00404	0.00396	1.96%	3.95***
7 or more	0.00386	0.00377	2.35%	4.47***
8 or more	0.00377	0.00367	2.69%	4.60***
9 or more	0.00355	0.00346	2.61%	3.91***
10 or more	0.00346	0.00337	2.59%	3.44***

Table 4: Immediate Reaction to Earnings News

The table reports the earnings response coefficients for measures of earnings surprise based on all analysts' forecasts and on the forecasts of the high and low quality (HQ and LQ) analysts defined in Table 1. The dependent variable is the buy-and-hold abnormal return (based on the four-factor Fama-French-Carhart model) for the earnings announcement day and the following trading day. SUE of Consensus and SUE of HQ and LQ analysts are defined in Table 3. All other variables are defined in Table 1. Columns (1) - (3) use the entire sample of earnings announcements, and columns (4) - (6) use the sample of earnings announcements by firms followed by at least four HQ analysts. All independent variables other than SUE are measured prior to the announcement date. The intercept and year fixed effects are included in all regressions. Robust standard errors are clustered by firm. *t*-statistics are provided in parentheses. The last two lines report p-values for chi-squared tests of the equality of the coefficients on SUE measures for the three analyst groups. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Full Sample			4 or more HQ analysts		
	(1)	(2)	(3)	(4)	(5)	(6)
SUE of Consensus	0.7245***			0.75260***		
	(13.62)			(5.20)		
SUE of HQ analysts		0.6211***			0.69274***	
		(12.93)			(5.12)	
SUE of LQ analysts			0.5691***			0.60441***
			(12.78)			(4.85)
Firm size	-0.0003	-0.0002	-0.0001	-0.00068	-0.00066	-0.00065
	(-0.71)	(-0.55)	(-0.37)	(-1.19)	(-1.15)	(-1.14)
Annual return	-0.0006	-0.0006	-0.0006	0.00066	0.00071	0.00069
	(-0.98)	(-0.89)	(-0.87)	(0.68)	(0.72)	(0.71)
Leverage	0.0059***	0.0057***	0.0057***	0.00301	0.00290	0.00275
	(3.73)	(3.55)	(3.56)	(1.24)	(1.19)	(1.12)
Book-to-market	0.00002	0.00001	0.00001	-0.00020***	-0.00018***	-0.00023***
	(0.38)	(0.29)	(0.28)	(-9.20)	(-7.67)	(-11.14)
Number of analysts	0.00002	0.00002	0.00001	-0.00004	-0.00003	-0.00004
	(0.25)	(0.28)	(0.08)	(-0.42)	(-0.40)	(-0.45)
Observations	44,709	44,709	44,709	20,221	20,221	20,221
Adjusted R-squared	0.0153	0.0134	0.0125	0.0108	0.0101	0.00887
p-value (SUE of HQ analysts vs. SUE of consensus)		0.000			0.009	
p-value (SUE of HQ analysts vs. SUE of LQ analysts)			0.02			0.01

Table 5: Abnormal return on earnings announcement day

The dependent variable is the buy-and-hold abnormal return (based on the four-factor Fama-French-Carhart model) for the earnings announcement day and the following trading day. High and low quality (HQ and LQ) analysts are defined in Table 1. *Predicted surprise* is equal to (HQ analysts' average forecast minus the consensus forecast) in columns (1) and (2) and (HQ analysts' average forecast minus LQ analysts' average forecast) in columns (3) and (4), normalized by the stock price at the beginning of the year. *Positive predicted surprise* indicator equals one if *Predicted surprise* is positive and zero if it is negative. *Big predicted surprise* equals one if *Predicted surprise* is greater than the median of positive values of *Predicted surprise* and zero if *Predicted surprise* is smaller than the median of negative values of *Predicted surprise* in year t-1. All independent variables are measured prior to the announcement date, and the regressions include the intercept and year fixed effects. Robust standard errors are clustered by firm. *t*-statistics are provided in parentheses. The last line of the table provides the two-day holding returns of a trading strategy that is long if the predicted surprise indicator variable in that column is equal to 1 and short if it is equal to 0. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Predicted surprise: HQ average – Consensus		Predicted surprise: HQ average – LQ average	
	(1)	(2)	(3)	(4)
Positive predicted surprise	0.0015*		0.0019**	
	(1.92)		(2.48)	
Big predicted surprise		0.0007*		0.0008**
		(1.71)		(1.98)
Firm size	0.00029	-0.00002	0.00028	-0.00002
	(0.76)	(-0.03)	(0.75)	(-0.04)
Annual return	-0.0002	-0.0006	-0.0002	-0.0010
	(-0.25)	(-0.67)	(-0.24)	(-1.15)
Leverage	0.0040**	0.0071***	0.0040**	0.0072***
	(2.45)	(2.89)	(2.45)	(2.90)
Book-to-market	0.00000	0.00001	0.00000	0.00000
	(0.03)	(0.15)	(0.03)	(0.11)
Number of analysts	-0.00000	-0.00006	-0.00000	-0.00002
	(-0.02)	(-0.51)	(-0.04)	(-0.15)
Observations	44,709	20,999	44,709	20,605
Adj. R-squared (%)	0.086	0.078	0.171	0.230
Two-day long-short strategy returns (%)	0.14*	0.20*	0.19**	0.24*
	(1.88)	(1.64)	(2.52)	(1.94)

Table 6: Returns following recommendation revisions

The dependent variable is the firm's stock return in the calendar month following the month with a recommendation revision. The sample consists of all recommendation revisions in the month when the annual earnings announcement is made by the firm. A recommendation is an integer from 1 to 5, where 1 is strong buy, 5 is strong sell, and 3 is hold. A recommendation revision is the negative of the difference between the current and the previous recommendations of an analyst, so that a positive (negative) recommendation revision is an upgrade (downgrade). The recommendation revision variable is the average of individual analysts' revisions for the firm during the earnings announcement month. The HQ and LQ indicators are for the HQ and LQ analysts, respectively, defined in Table 1. The other indepedent variables are defined in Table 1 and measured prior to the earnings announcement. All regressions include the intercept, and robust standard errors are clustered by firm. The last line reports long-minus-short portfolio returns in the calendar month following the month with the revision where the long (short) position is in the firms for which the mean recommendation revision is positive (negative). *t*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
Recommendation revision × HQ indicator	0.0025**		0.0025**
	(2.09)		(2.09)
Recommendation revision × LQ indicator		0.0004	0.0004
		(0.30)	(0.29)
Lagged dependent variable	0.0143	0.0158	0.0141
	(1.04)	(1.15)	(1.01)
Firm size	-0.0043***	-0.0043***	-0.0043***
	(-4.25)	(-4.20)	(-4.26)
Leverage	0.0008	0.0007	0.0008
	(0.19)	(0.16)	(0.19)
Book-to-market	-0.00001	-0.00001	-0.00001
	(-0.30)	(-0.32)	(-0.31)
Number of analysts	0.0003*	0.0003*	0.0003*
	(1.81)	(1.78)	(1.81)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	21,381	21,381	21,381
Adj. R-squared	0.0420	0.0419	0.0420
One month long-short strategy returns (%)	0.85***	0.14	
	(3.11)	(0.49)	

Table 7: Forecast dispersion predicting return volatility

The dependent variable is the standard deviation of the firm's daily returns in the month following the annual earnings announcement month. Forecast dispersion is the standard deviation of forecast errors defined in Table 1 and based on analysts' closest forecasts issued during 60 days prior to the earnings announcement. The other independent variables are defined in Table 1 and measured prior to the announcement date. All regressions include the intercept. Robust standard errors are clustered by firm. *t*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
Forecast dispersion of all analysts	0.0731*		
	(1.64)		
Forecast dispersion of HQ analysts		0.1315**	
		(2.55)	
Forecast dispersion of LQ analysts			0.0340
			(0.79)
Lagged dependent variable	0.5179***	0.5141***	0.5206***
	(12.69)	(12.82)	(12.63)
Firm size	-0.0007	-0.0006	-0.0008*
	(-1.59)	(-1.40)	(-1.73)
Annual return	0.0002	0.0002	0.0002
	(0.53)	(0.56)	(0.46)
Leverage	-0.00065	-0.00068	-0.00052
	(-0.29)	(-0.31)	(-0.23)
Book-to-market	0.0015	0.0016	0.0015
	(1.25)	(1.41)	(1.16)
Number of analysts	-0.00004	-0.00004	-0.00003
	(-1.22)	(-1.40)	(-1.13)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	4,812	4,812	4,812
Adj. R-squared	0.691	0.693	0.691

Table 8: Post-earnings announcement drift and analysts' relative uncertainty

The table reports the cumulative drift for 1 to 11-month horizons following annual earnings announcements. Announcements are divided into two subsamples in which the standard deviation of the firm's HQ analysts' forecast errors is greater (the *high uncertainty* sample) or smaller (the *low uncertainty* sample) than the standard deviation of forecast errors for all analysts following the firm. Each stock is held in a calendar-time portfolio for the length of the horizon. The monthly value-weighted portfolio returns are regressed on the four Fama-French-Carhart factors to obtain the drift, which is the intercept of the regression (monthly alpha). A stock is assigned to the long or short portfolio depending on whether its earnings surprise is positive or negative, respectively, where earnings surprise is defined based on the consensus estimate. The high quality analysts are defined in Table 1. *, **, *** represent 10%, 5%, 1%, significance based on the regression *t*-statistics, respectively.

Drift horizon (months)	High uncertainty			Low uncertainty		
	Long	Short	Long-Short	Long	Short	Long-Short
1	1.09	-0.35	1.44*	-0.46	0.30	-0.76
2	1.36***	-0.12	1.48**	0.13	0.33	-0.20
3	1.87***	-0.58	2.46***	0.20	0.19	0.01
4	1.06***	0.04	1.02	-0.27	0.35	-0.62
5	0.90***	0.18	0.72	-0.05	-0.35	0.30
6	1.19***	-0.40	1.58***	-0.03	-0.20	0.17
7	0.97***	-0.35	1.42***	-0.16	0.07	-0.23
8	0.97***	-0.48*	1.45***	-0.18	-0.22	0.04
9	0.50***	-0.56**	1.06***	-0.14	-0.22	0.09
10	0.42***	-0.44**	0.85***	-0.08	-0.33	0.25
11	0.41***	-0.44**	0.85***	0.03	-0.4**	0.37*

Table 9: HQ analysts predicting returns and volatility at the aggregate level

The dependent variables are value-weighted returns in 2-digit SIC industries (Panel A), value-weighted market returns (Panel B), and the VIX index return (Panel C) in the month following the month with the earnings announcement. Panels A and B uses recommendation revisions during the announcement month defined in Table 6, and Panel C uses the dispersion of analysts' forecasts defined in Table 7. The HQ and LQ analysts are defined in Table 1. The *mean recommendation revision* variables are averages across all analysts in a given industry (Panel A) and the entire market (Panel B). A monthly industry return is included in the weighted average if there is more than one recommendation change for a firm in the industry during the month. The control variables are the monthly *earnings-to-price ratio, dividend-to-price ratio, term spread, default spread, one-month T-bill rate, 30-year Treasury yield, the rate of inflation*, and economic policy uncertainty (EPU) change are described in Section 2. Panel A has the first seven controls, which are also interacted with industry fixed effects. *Dispersion* variables are firm-level value-weighted averages of analysts' forecast error dispersions with the firm's market capitalization as the weights. *Dispersion ratio for HQ (LQ) analysts* is the firm value-weighted average of the ratio of *Dispersion of HQ (LQ) analysts* to *Dispersion of all analysts*. The models use robust standard errors clustered by industry in Panel A, Newey-West standard errors with three lags in Panel B, and Huber-White robust standard errors in Panel C. In the long-minus-short portfolio returns for the industry specifications in Panel A, the long (short) position is in the industries for which the mean recommendation revision is positive (negative). The last lines in Panels A and B report the alphas from a market model regressions obtained as follows. In Panel A, industries whose mean recommendation revisions are positive (negative) are assigned to the long (short) portfolio each month, and the portfolio returns are value-weighted to produce a long-minus-short monthly return, which is then regressed on the market value-weighted return. In Panel B, the market return is multiplied by 1 (-1) if the mean recommendation revision this month is above (below) the median of mean recommendaton revisions during the previous 24 months. The last line in Panel C provides the alphas from regressing a VIX trading strategy return on the market value-weighted return, where the VIX strategy return is the VIX return during the month following the announcement month multiplied by 1 (-1) if the column's *Dispersion* or *Dispersion ratio* that month is above (below) the median of *Dispersion* or *Dispersion ratio* during the previous 24 months. *t*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Predicting Industry Returns

	(1)	(2)	(3)	(4)	(5)	(6)
All analysts' mean	0.002**			0.002*		
recommendation revision	(2.25)			(1.98)		
HQ analysts' mean		0.003**			0.003**	
recommendation revision		(2.64)			(2.38)	
LQ analysts' mean			-0.00001			0.0002
recommendation revision			(-0.01)			(0.17)
Lagged dependent variable	0.068***	0.068***	0.069***	-0.036***	-0.036***	-0.035***
	(7.71)	(7.73)	(7.74)	(-3.10)	(-3.08)	(-3.05)
Constant	0.009***	0.009***	0.009***	-0.081***	-0.082***	-0.081***
	(26.56)	(26.62)	(26.45)	(-165.68)	(-125.61)	(-158.58)
Industry fixed effects	No	No	No	Yes	Yes	Yes
Control variables interacted with industry fixed effects	No	No	No	Yes	Yes	Yes
Observations	16,331	16,331	16,331	16,331	16,331	16,331
Adj. R-squared	0.0048	0.0049	0.0046	0.0265	0.0266	0.0264
One month long-short predicted returns (%)				0.51**	0.66***	0.36
				(2.30)	(2.81)	(1.49)
Monthly alpha (%)				0.07	0.57**	-0.18
				(0.28)	(2.13)	(-0.58)

Panel B: Predicting Market Returns

	(1)	(2)	(3)	(4)	(5)	(6)
All analysts' mean	0.016*			0.012		
recommendation revision	(1.85)			(1.54)		
HQ analysts' mean		0.017**			0.014**	
recommendation revision		(2.25)			(1.96)	
LQ analysts' mean			0.005			0.003
recommendation revision			(0.91)			(0.53)
Lagged dependent variable	0.073	0.073	0.070	0.008	0.011	0.004
	(0.93)	(0.96)	(0.90)	(0.12)	(0.17)	(0.06)
Earnings-to-price ratio				0.737***	0.732***	0.757***
				(2.63)	(2.65)	(2.64)
Dividend-to-price ratio				-1.232*	-1.190*	-1.287*
				(-1.83)	(-1.78)	(-1.89)
Term spread				-0.438	-0.439	-0.441
				(-0.70)	(-0.70)	(-0.70)
Default spread				0.069	0.128	0.080
				(0.07)	(0.13)	(0.08)
One month t-bill rate				3.220	3.203	3.238
				(1.45)	(1.44)	(1.45)
Long term Treasury yield (30 year)				-0.002	-0.011	0.050
				(-0.00)	(-0.02)	(0.08)
Inflation				0.285	0.236	0.348
				(0.41)	(0.35)	(0.49)
Constant	0.009***	0.010***	0.008***	0.015	0.015	0.012
	(3.44)	(3.59)	(2.78)	(0.62)	(0.63)	(0.48)
Observations	265	265	265	258	258	258
Adj. R-squared (%)	0.0085	0.0175	-0.001	0.0463	0.0526	0.041
Monthly alpha (%)				0.43	0.51*	0.34
				(1.48)	(1.72)	(1.15)

Panel C: Predicting VIX Index Return

	(1)	(2)	(3)	(4)	(5)
Dispersion - all analysts	1.197				
	(0.22)				
Dispersion - HQ analysts		4.339			
		(0.59)			
Dispersion - LQ analysts			1.475		
			(0.24)		
Dispersion ratio - HQ analysts				0.097*	
				(1.82)	
Dispersion ratio - LQ analysts					-0.018
					(-0.51)
Lagged dependent variable	-0.134	-0.136	-0.135	-0.124	-0.130
	(-1.46)	(-1.50)	(-1.46)	(-1.35)	(-1.42)
EPU change	0.042	0.043	0.042	0.048	0.041
	(0.76)	(0.79)	(0.77)	(0.87)	(0.74)
Value-weighted market return	0.827*	0.840**	0.832**	0.830*	0.818*
	(1.97)	(2.00)	(1.98)	(1.97)	(1.95)
Earnings-to-price ratio	-2.609	-2.649	-2.610	-2.480	-2.619
	(-1.47)	(-1.49)	(-1.47)	(-1.41)	(-1.48)
Dividend-to-price ratio	5.189	4.857	5.105	6.072	5.581
	(1.12)	(1.05)	(1.09)	(1.37)	(1.24)
Term spread	-2.095	-2.211	-2.053	-2.179	-2.150
	(-0.61)	(-0.64)	(-0.60)	(-0.63)	(-0.63)
Default spread	-4.382	-4.290	-4.442	-4.258	-4.556
	(-1.20)	(-1.18)	(-1.23)	(-1.20)	(-1.25)
One month TB rate	-5.495	-5.967	-5.719	-6.425	-5.366
	(-0.43)	(-0.47)	(-0.45)	(-0.50)	(-0.41)
Long term Treasury yield (30 year)	2.144	2.344	2.118	2.213	2.077
	(0.64)	(0.70)	(0.63)	(0.65)	(0.62)
Inflation	-2.495	-2.624	-2.556	-2.277	-2.358
	(-0.62)	(-0.65)	(-0.63)	(-0.57)	(-0.59)
Constant	0.092	0.092	0.094	-0.007	0.114
	(0.57)	(0.57)	(0.58)	(-0.04)	(0.68)
Observations	197	197	197	197	197
Adj. R-squared (%)	0.0127	0.0154	0.0128	0.0295	0.0134
Obs. in Long portfolio	115	124	74	141	64
Obs. in Short portfolio	97	88	138	71	148
Monthly alpha (%)	1.85	2.04	-0.57	3.32***	-0.56
	(1.34)	(1.60)	(-0.40)	(2.68)	(-0.41)

Table 10: Alternative definitions for HQ analysts and replication of results

Column (1) replicates Table 3 for the number of HQ analysts following the firm so that their average forecast has a lower absolute SUE than the consensus forecast. Column (2) replicates the test for the difference between the coefficients on SUE in columns (4) and (5) of Table 4 where the minimum number of the HQ analysts following the firm is provided in column (1) here. Column (3) corresponds to the trading strategy return in column (1) of Table 5. Column (4) replicates the coefficients on the cross-terms and one month long-short strategy returns in columns (2) and (3) of Table 6. Column (5) replicates the coefficients on the forecast dispersion in columns (2) and (3) of Table 7. Column (6) replicates the coefficients on recommendation revisions in columns (5) and (6) of Table 9, Panel A. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Table 3 Number of HQ analysts required for ABS(SUE HQ) < ABS(SUE consensus)	Table 4 Market reaction to SUE when the number of HQ is as in column (1)	Table 5 Trading strategy return on announcement day	Table 6 Recommendation revisions predicting next month return	Table 7 Predicting next month return volatility	Table 9 Predicting next month industry return
	(1)	(2)	(3)	(4)	(5)	(6)
HQ: top 70%	3 or more	Reaction to consensus greater by 0.023 (significant *)	0.18% (significant **)	HQ: significant*** LQ: not significant Trading: 1.43%***	HQ: significant*** LQ: not significant	HQ: significant*** LQ: not significant
HQ: top 30%	6 or more	Reaction to consensus greater by 0.069 (not significant)	0.20% (significant*)	HQ: significant** LQ: not significant Trading: 1.20%***	HQ: significant* LQ: not significant	HQ: significant** LQ: not significant
Value-weighted measure, HQ: top 50%	4 or more	Reaction to consensus greater by 0.001 (not significant)	0.10% (not significant)	HQ: significant*** LQ: not significant Trading: 0.69%**	HQ: significant*** LQ: not significant	HQ: significant*** LQ: not significant