Analysts' Consensus Forecasts and Mispricing

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

January 17, 2017

Abstract:

We find that analysts' earnings forecasting performance is an analyst characteristic persistent over time and across firms. By dividing all analysts into the high and low quality (HQ and LQ) categories based on their forecast accuracy, we are able to generate a forecast measure superior to the consensus. The average forecast of the HQ analysts is more accurate than the consensus, but the market is unaware of this and reacts only to the surprise based on the consensus estimate. We find that the superiority of the HQ analysts comes from their better understanding of market-wide information. Recommendation changes of the HQ analysts predict future industry and market returns, while the consensus recommendation changes do not. Further, the VIX and the PEAD are higher following periods of greater uncertainty among the HQ analysts. Overall, the market inefficiently overweighs the information content of the consensus and does not utilize price-relevant information in the forecasts and recommendations of the HQ analysts.

*Michaely is from Cornell University and IDC (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).

Investors use analysts' consensus forecasts as the measure of market expectations of the firm's future earnings results. The perceived importance of consensus earnings estimates has greatly increased in the recent years, to the extent that even companies' investor relations departments tend to follow it on a continuous basis (Consensus earnings estimates report, 2013). However, a disadvantage of the consensus is that by construction, it gives equal weight to each analyst's estimate and hence ignores that analysts have different abilities due to 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. The fixation on a simple average of analysts' forecasts that disregards differences in analyst forecasting ability motivates a question whether one can obtain a forecast measure which is more informative than the consensus, and if so, whether the market is also aware of this and incorporates this into prices.

We find that analysts persistently differ in their forecasting ability, and, in general, investors would benefit from using the relatively more accurate forecasts of the high quality (HQ) analysts instead of the consensus. Further, the market does not incorporate the superior information in the HQ analysts' forecasts into prices in either the short run or the long run. The post-earnings announcement drift phenomenon takes place when HQ analysts are relatively uncertain about the firm's prospective earnings. The HQ analysts as a group have a predictive ability for the economy, industry, and market volatility in contrast to when all analysts are treated as equal.

We start with verifying a key necessary condition implicit in the principle of differentiating analysts in terms of their quality—that analysts' forecast accuracy is persistent over time. Because this study aims to compare the consensus, which is the average of all analysts' estimates, with other subsamples of analysts, the necessary condition requires that persistence in ranking exists within the entire sample of analysts rather than, for example, merely between the top 25% and bottom 25% of ranked analysts, which is often sufficient in other studies. We find that analysts who are categorized as the HQ ones in terms of their forecasting accuracy in the firm in a given year tend to be ranked as the HQ ones in the following year as well. Furthermore, this conclusion is not affected by the accuracy distribution cutoff value distinguishing between HQ and low quality (LQ) analysts. For simplicity, in the rest of the paper, we define HQ and LQ analysts as those, respectively, above and below the median in the accuracy ranking in the previous year. We find that annual earnings forecasts of HQ analysts are 5.7% more accurate compared to the full set of analysts following this firm. We also test whether analysts' persistent forecasting performance can be described as an analyst characteristic. We find that analysts who are HQ in one firm are also likely to be HQ in the other firms they follow.

The finding that analyst performance has a persistent component points to the question whether investors should always follow the HQ analysts and disregard the consensus which includes the forecasts by the LQ analysts. The rationale for the widespread following of the consensus lies in that averaging of analysts' estimates, whose number can reach forty for some firms, improves forecast accuracy by having a "portfolio effect" of reducing measurement errors in individual forecasts. In general, there is a tradeoff between the quality and quantity of forecasts: On the one hand, the consensus includes more estimates, which is beneficial for reducing the measurement error; on the other hand, analysts who tend to have more accurate forecasts have smaller measurement errors to begin with. Our analysis reveals that as the number of HQ analysts following the firm increases, the accuracy of the average of their estimates improves and exceeds that of the consensus.

The findings about the persistence of the HQ analysts' superior performance have implications to the efficiency of financial markets. If the market is aware that it is beneficial to focus on the estimates of HQ analysts rather than the consensus, it would react more vigorously to earnings surprises that are measured based on the average forecast of HQ analysts. However, we find the opposite—the earnings response coefficient to standardized unexpected earnings (SUE) of the consensus is much higher than that of SUE based on the average of the HQ analysts' estimates. This finding suggests that the high visibility and publicity of the consensus leads the market to put too much weight on the importance of the consensus, and at the same time, the market fails to incorporate the information embedded in the forecasts of HQ analysts. Further, one can exploit this inefficiency by taking a long (short) position in the stock if the HQ analysts' average forecast is greater (smaller) than the consensus as of one day before the announcement, which generates a 0.45% abnormal return over the announcement day and the following trading day.

We next investigate the source of the superior performance of the HQ analysts. We conjecture that the informational advantage of the HQ analysts can come from their superior ability to integrate the systematic component of information (information about the industry and overall economy) and/or be due to their relatively better access to macroeconomic data (Jennings, 1987). This explanation for better performance of a group of analysts also has similarity to the main theme in Hutton, Lee, and Shu (2012) that analysts are better than managers in forecasting earnings in firms relatively more exposed to macroeconomic factors. A testable implication of our hypothesis linking HQ analysts to economic factors is that HQ analysts provide stock recommendations that can predict industry and market returns, in contrast to the average recommendation of all analysts following the firm. The hypothesis also suggests that when HQ analysts are relatively uncertain, there is a high level of uncertainty in the economy, so that their uncertainty reflected in their forecasts should be associated with current and/or future market volatility.

We find empirical support for both predictions. By relying on the recommendation changes of HQ analysts, one achieves predictive ability for the market and industry returns in the following month. For the second prediction, to make the dispersion of the HQ analysts' forecasts comparable across firms and time, we normalize it by the dispersion of all analysts' forecasts and call this a dispersion ratio. When the dispersion ratio is greater than one HQ analysts are more uncertain than other analysts. Indeed, we find that this ratio is associated with a higher market volatility as measured by the percentage change in the VIX over the next month and year. In contrast, analysts' forecast dispersion at the consensus level, which commonly proxies for uncertainty about future earnings or, more generally, information uncertainty (e.g., Diether, Malloy, and Scherbina, 2002; Zhang 2006), does not have a relation to the VIX. Given that the VIX is often interpreted as the "fear index", investors should be worried about the economy and the stock market performance already when the HQ analysts become relatively uncertain compared to all other analysts.

Finally, we consider whether the dispersion ratio capturing high market-wide uncertainty and divergence of opinion among HQ analysts compared to that among all analysts is predictive of the post-earnings announcement drift (PEAD). The model in Abarbanell, Lanen, and Verrecchia (1995) suggests that when forecast dispersion is high, investors delay their complete response to earnings announcements, which could lead to a greater PEAD. We are unaware of existing studies confirming this prediction. The two closest related studies are Zhang (2006), which finds that analysts' forecast dispersion predicts the price drift following analysts' forecasts (the relation to the PEAD is not tested), and Hung, Li, and Wang (2015), which does not study analysts' forecast dispersion but finds that reduced uncertainty due to increased disclosure leads to a lower PEAD. The key innovation in our analysis is that we find the forecast dispersion of all analysts is not sufficient to measure the impact of forecasts on investor uncertainty or PEAD, while the degree by which uncertainty among the HQ analysts is greater than that of all analysts is.

We hypothesize that the PEAD is higher when the HQ analysts are relatively more uncertain, which would be indicated by the forecast dispersion of their forecasts exceeding that of all analysts. The standard PEAD strategy is that of buying (shorting) shares when SUE is positive (negative). We follow this strategy for the subsamples in which the forecast dispersion ratio is high and low and find a much greater PEAD if the HQ analysts are relatively uncertain than when they are less uncertain, 9.4% vs. 4.1% after 11 months, respectively. During most of this forecast horizon, the PEAD is not statistically different from zero in the sample when HQ analysts are relatively less uncertain, implying that the long-puzzling PEAD phenomenon arises only during the periods of high uncertainty among HQ analysts, i.e., during the periods of high market-wide uncertainty. Overall, these findings indicate that the superior information in the HQ analysts' forecasts not only predicts the immediate reaction to earnings announcements but also the long-term market response.

The conclusion that it is worthwhile to average estimates dates back to Makridakis and Winkler (1983) and Conroy and Harris (1987). Our main contribution in this regard lies in that we enhance the evidence on analyst forecast persistence (Stickel, 1992; Sinha, Brown, and Das, 1997) to the entire sample of analysts' estimates and utilize it to form a more accurate measure of the earnings news than the simple average. As such, our paper belongs to a recent literature that analysts' average estimates can be misleading (So, 2013; Buraschi, Piatti, and Whelan, 2016).

The paper also contributes to the inattention literature by showing that the market does not incorporate information that exists in the estimates of the more qualified analysts. The paper extends the literature that provides evidence that investors experience inattention to earnings announcements (Hirshleifer, Lim, and Teoh, 2009, 2011; Bailey, Kumar, and Ng, 2011; Chakrabarty and Moulton, 2012; Michaely, Rubin, and Vedrashko, 2016) by showing that they are also inattentive to the value embedded in individual estimates that are publicly available prior to earnings announcements.

Finally, the paper contributes to the large literature that is concerned with understanding the PEAD (Bernard and Thomas 1989; 1990) phenomenon. We extend recent studies that argue that uncertainty tends to amplify investors' valuation mistakes due to behavioral biases (Kumar, 2009), which in turn affects the PEAD phenomenon.

The paper proceeds as follows. In Section 2 we present the data and variables. Section 3 provides information on the persistence in analysts' quality. Section 4 provides a theoretical framework and empirical evidence for when it is unwise to use the consensus as the measure for expected earnings. Section 5 provides empirical evidence of the mispricing at the earnings announcement date. Section 6 studies the link of the superior ability of HQ analysts to market wide uncertainty and the PEAD. Section 7 concludes.

2. Data and variables

We use the sample of all annual EPS estimates and earnings announcements in I/B/E/S during the period from January 1992 to December 2015 that also have daily return data in CRSP. The starting year of 1992 is chosen because some of our analyses require analysts' recommendation data, which begins in 1993. Earnings estimates and actual earnings are adjusted for splits by using the daily cumulative adjustment factor from CRSP (Glushkov and Robinson, 2006).

2.1 Analyst rankings

Each year, we rank analysts based on two measures—the closest absolute forecast error and the value-weight absolute forecast error. The closet absolute forecast error is the absolute difference between an analyst's forecast estimate closest to the earnings announcement (but made at least one day prior to the announcement) and announced annual earnings, divided by the share price at the beginning of the calendar year.[1] The value-weight absolute forecast error takes all estimates of the analyst during the 300 days prior to the annual earnings announcement to produce a value-weight measure based on the number of days the estimate was outstanding as follows,

$$VWFE_t = \frac{FE_{300} \times d_1 + \sum_{j=2}^{n} (FE_j \times d_j)}{300} \quad (1)$$

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 estimate outstanding on the 300th day prior to the earnings announcement; d_1 is the number of days this estimate is outstanding (from the 300th day prior to the earnings announcement to the announcement day or to the day when an updated

[1] We focus on annual earnings rather than quarterly for two main reasons. First, most analysts provide annual forecasts but fewer analysts provide quarterly forecasts. Second, annual earnings announcements are typically more informative, associated with a conference call, and followed by a recommendation change, which suggests that analysts have a stronger incentive to be accurate in their estimates of annual figures.

estimated is released); *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 based on estimate *j;* and d_j is the number of days the estimate was outstanding.

Table 1 provides information on the sample. From the initial sample, we generate 861,349 firm-year-analyst rankings based on the closest forecast error. This number drops to 829,126 rankings once we require firms to have the Compustat data. Next, we realize that our ranking may suffer from small-sample bias when the number of analysts following the firm is small, so we exclude from the sample firms that have less than four analysts following, which further reduces are sample to 768,320 observations. Finally, we note that for most of the analyses in the paper (but not all) we require the analyst to appear in the data in two consecutive years, which drops these analyses sample by about 37% to 486,239 observations. While we are able to generate ranking for any analyst following the firm by using the closest forecast error measure, the value-weight measure requires us to have an analyst estimate at least 300 days prior to the annual earnings announcement day. This is a drawback of the value-weight sample and we have no ability of knowing whether the lack of estimate is due to a lack in the analyst's ability to forecast or simply a benign reason such as common practice in the industry or firm. As can be seen in Table 1, using the value-weighted absolute forecast error reduces the sample by approximately 60%, which reduces the power of our empirical analysis. Because the quantitative nature of our results is similar with both types of measures, we present the results using the closest forecast error throughout the paper.[2]

Finally, we note that an alternative ranking procedure would be to rank an analyst by averaging his forecast errors across all firms he follows in a given year rather than our approach

[2] We note that most of our empirical results also hold with the value-weight measure ranking, although the significance of several results is somewhat reduced. These results are available from the authors by request.

of ranking an analyst in each firm in a given year. There are several advantages of the former approach ranking analysts over all firms. An average analyst follows 15 firms in a given year, which means we could avoid small sample bias arising if a firm is followed by too few analysts and, perhaps, achieve a higher level of persistence in analyst ranking. We would also avoid losing the observations of the first year when an analyst begins covering a firm because we would rely on the analyst's ranking in the previous year in other firms. However, the approach ranking an analyst across all firms also has its pitfalls. First, ranking an analyst across firms in a given year can be misleading if analysts' ability to predict earnings is mainly firm-specific rather than industry-specific or economy-wide. Second, with the year-level ranking, we will end up with some firms followed mostly by HQ or LQ analysts, and even by just one analyst quality type in some firms. This will undermine our study's objective of analyzing whether it is worthwhile to follow the average estimate of HQ analysts or the consensus estimate, for which the presence of both HQ and LQ analysts in a firm is required. The empirical results would also be less general because they become dependent on the analyst quality distribution across firms, i.e., on firm characteristics, for example, if the most talented analysts tend to follow the most complex and largest firms. In contrast, the firm-year analysis by design makes the fractions of HQ and LQ analysts equal in each firm in year t-1 when they are ranked and approximately equal in the analyzed year, and this proportion is similar across all firms. This makes firms comparable in our analysis regardless of their characteristics.

2.2 Analysts' forecasts and other controls

In firm-level regressions, we control for five firm characteristics—size, annual return, book-to-market, number of analysts, and leverage. Size is the market value of the firm's equity at the end of the quarter prior to the announcement quarter based on Compustat. The annual return is the annual return of the firm's equity during the 12 months prior to earnings announcement month. The Book-to-market ratio is stockholder equity minus the preferred stock plus deferred taxes divided by the market value of equity. The number of analysts is the number of analysts who made forecasts of the year's earnings; the data for this variable are obtained from the IBES Summary file. Leverage is the ratio of total book debt to book assets. In the regression models, we also control for analyst characteristics. An analyst's overall tenure is the number of years since the analyst first appeared in the I/B/E/S file. An analyst's tenure in the company is the number of years since the analyst began covering this company in the I/B/E/S file. The size of an analyst's brokerage house is the number of analysts in the brokerage house. An analyst's firm coverage is the number of firms covered by the analyst.

3. Persistence in analysts forecasting ability

3.1. Defining high and low quality analysts

This section describes how we partition analysts to either high or low quality based on their absolute forecast error and then analyzes whether this classification of analysts persists in the following year. To this end, we sort analysts in a given firm-year based on their absolute forecast error. 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. We note that 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 $p^2 + (1-p)^2$ percent of all analysts.

Figure 1 plots the fraction of analysts that retain their rankings in consecutive years and the expected fraction of analysts assuming no performance correlation across years. We find that with almost all cut-off values, the actual fraction of persistent forecasting performance is above the expected fraction. 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 0.829. The exception is the strict definition of HQ analysts as the best 5%. We explain this with a high competition among the top forecasters, which causes frequent changes in relative rankings at the very top.[3] Nevertheless, the overall important finding is that for almost all of the performance cutoff values, there is a sizeable persistent component, so that it does not matter which exact cutoff value we choose to partition high and LQ analysts.

In the remainder of the paper, we define HQ and LQ analysts based on whether their absolute forecast error for the firm-year is below or above the median absolute forecast error for the firm-year, respectively. We choose the median as the cutoff because of its advantage that the numbers of analysts in the high and low quality groups are exactly equal in year t-1 and relatively close in year t. This mitigates a concern that differences between the average estimates of HQ and LQ analysts could be generated by too few analysts in either a high or low quality group, potentially causing small sample bias in that group.

3.2. Mean absolute forecast error

[3] Stickel (1992) reports that analysts on the All-American Research Team are dropped from the list when they have relatively inaccurate forecasts.

Figure 2 shows the mean absolute forecast error during the 300 days prior to the earnings announcement. We observe that the reduction in the mean forecast error accelerates around quarterly earnings announcements at 90, 180, and 270 day marks. Each year, all analysts are ranked based on their forecast accuracy in the particular firm, and the figure shows forecast errors next year for three different subsamples of analysts based on their ranking. For the comparability of the subsamples, we require that firm-years have at least one HQ analyst who provided an estimate at least 300 days prior to the earnings announcement and also made a forecast in year *t*. Hence, the sample of firms and earnings announcements is the same for the three plots.

The solid line is for the entire sample of analyst estimates. It is clear that the mean absolute forecast error is the largest for this sample. One can also observe that during the 300 days, the mean absolute forecast error of the entire sample of analysts declines from 2.6% to around 1.2% one day before the earning announcement. Excluding estimates of analysts who did not cover the firm in the previous year improves the forecasts and reduces the absolute forecast error throughout the 300 days. For example, absolute forecast error is 1.17% prior to earnings announcements based on estimates of analysts that are ranked at t-1. This means that estimates of new analysts who cover the firm for the first year are associated with an increase in the average absolute forecast error. Finally, the forecasts of the HQ analysts have an absolute forecast error of 1.14% prior to earnings announcements. These differences between the samples of analysts are economically meaningful. Just before the earnings announcement date, the estimate based on HQ analysts improves the accuracy by 5.35% compared to the entire sample ($\frac{1.14\%}{1.20\%} - 1$), greater than the average accuracy improvement over a 20 day period of 2.52%.

Table 2 analyzes how the classification of analysts to low or high quality is associated with various analysts' characteristics and how persistent it is over time. In Panel A, we find that HQ analysts tend to have more experience both in the company and in the financial industry. We also find that HQ analysts work in larger brokerage firms. We do not find evidence that analysts' quality is related to the number of firms they cover. Panel B conducts regression analysis that shows that the quality level of an analyst persists. In specifications (1) and (2), probit regression results are provided where the dependent variable is an indicator that equals one if the analyst is of HQ and zero otherwise. In specifications (3) and (4), the dependent variable is absolute forecast error, a continuous variable, which allows us to control for firm fixed effects in the model. In specifications (1) and (3), we control with firm characteristics, and in specification (2) and (4) we control with both firm and analyst characteristics. The results unambiguously show that the high quality indicator is highly significant in all specifications, i.e., analysts' rankings are persistent in consecutive years. High (low) quality analysts tend to retain their ranking category and have a low (high) absolute forecast error in the following year.

We next conduct additional, cross-firm tests to determine whether forecasting performance is an analyst characteristic, in which case an analyst who performs well in one firm should perform well in other firms he or she follows. In Table 3, we conduct two tests of this hypothesis. We define an analyst's performance in other firms as that of high (low) quality if the analyst belongs to the high (low) quality category in the majority of the other firms the analyst follows during the year (i.e., excluding this firm)[4]. Panel A calculates the percentage of firms in which an analyst has the same quality (high or low) in other firms as in a given firm. We find that HQ analysts in a given firm are also ranked as the HQ ones in the other firms 54.5% of the

[4] 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 is excluded from this analysis.

time, and LQ analysts in a given firm also tend to be LQ in the other firms the same year. Panel B tests whether ranking as a HQ on in other firms last year can predict an analyst's forecasting performance, the HQ analyst status, in a given firm this year. We estimate two probit models for the HQ analyst indicator predicted by the HQ indicator in the previous year, which is equal to one if the analyst is of HQ in the majority of the other firms followed by the analyst that year. We control for the lagged HQ analyst indicator based on the persistency relation in Table 2. We find that analysts who were of HQ in the majority of other firms last year are 5.1% more likely to be HQ this year. This persistency effect is in addition to the performance persistency effect in this firm, whose coefficient is also positive and significant. The cross-firm findings in Table 3 suggest that analysts' forecasting performance is a persistent analyst characteristic.

4. A tradeoff between the number of analysts and the quality of analysts

Our finding that a HQ analyst consistently makes better earnings forecasts than a LQ analyst at the individual level leads us to a question whether investors should heed the *average* of the HQ analysts' forecasts rather than the consensus forecast, which averages both HQ and LQ analysts' forecasts. In this section, we consider how this decision is affected by the relative number of HQ and LQ analysts in the firm. Intuitively, the greater the number of forecasts the smaller the measurement error, but forecasts by HQ analysts have on average a smaller difference from the actual earnings to begin with. This leads to a tradeoff between the total number of analysts and the relative number of HQ analysts within the pool of analysts. A model illustrating this relation is in the Appendix. The theoretical framework tells us that as the fraction of HQ analysts increases, it is more likely that investors obtain a more accurate forecast than the consensus by following only the average estimate of HQ analysts. We verify this intuition in this section. Our objective is to empirically determine the smallest number of HQ analysts that allows one to heed only HQ analysts instead of the consensus.

We begin with the summary statistics of the distributions of high, low, and unranked (those who did not follow the firm in the previous year) analysts in Table 4. It shows how the relative number of the three analyst types depends on the number of HQ analysts that existed in the firm just prior to the earnings announcement. We find that as the number of HQ analysts increases, their fraction increases as well, while the fraction of LQ analysts stays approximately the same, and the fraction of unranked analysts decreases. We do not have an affirmative explanation for this result except for the intuitive relation that, in equilibrium, a large following by HQ analysts in a firm makes other analysts (LQ and unranked newcomers) less inclined to stay and compete with them or start to follow the firm. The positive relation between the number and fraction of HQ analysts allows us to conduct our analysis by considering the number of HQ analysts as the variable of interest when deciding whether to have an earnings estimate based on the average of all analysts' estimates (i.e., the consensus) as opposed to averaging only the HQ analysts' estimates.

Figure 3 plots how the number of HQ analysts affects the percentage of announcements for which the average forecast of HQ analysts is more accurate than the consensus. It is clear that as the number of HQ analysts increases, the fraction of announcements for which HQ analysts do better than the consensus increases, and the relation is almost a straight line. This confirms the prediction of the theoretical model. Further, we observe that when the number of HQ analysts is five or more, it is on average better to ignore the estimates of LQ analysts, i.e., the consensus estimate, and, instead, rely on a measure based solely on the HQ analysts' forecasts. Table 5 complements the analysis in Figure 3 with statistical tests comparing the absolute SUE of consensus with the absolute SUE of HQ analysts. We find that when the number of HQ analysts in the firm is less than four, investors are better off heeding the consensus estimate, but as the number of HQ analysts increases, investors should switch their attention to the average of HQ analysts' estimates. When there are eight or more HQ analysts, the accuracy improvement of the HQ analysts over the consensus becomes statistically significant.

5. Is the market aware of high quality analysts?

5.1 Comparing earnings response coefficients

The previous section demonstrates that using earnings forecasts of HQ analysts, one can generate an earnings forecast superior to the consensus forecast. We next test whether the market is aware of this possibility. To that end, we analyze the immediate market reaction to earnings announcements. In general, researchers tend to use the difference between actual earnings and the consensus as the proxy for the unexpected component of earnings news. However, it is not certain whether the market is surprised and reacts to a deviation from the consensus or a deviation from a different earnings estimate, at least to some extent. This empirical question we address here. Because we find that the average forecast of the HQ analysts is preferable to the consensus when the number of HQ analysts is large, we formulate a hypothesis that the market is aware of this relation and pays attention only to HQ analysts under such circumstances. If that is the case, the earnings response coefficient to *SUE of HQ analysts* should be greater than that of *SUE of consensus* when the number of HQ analysts is greater than four.

Table 6 reports regression results in which the dependent variable is the buy-and-hold abnormal return based on the four-factor model for the earnings announcement day and the following trading day. The competing coefficients of interest are on the independent variables

SUE of consensus and *SUE of HQ analysts*. The results clearly show that it is *SUE of consensus* to which the market tends to react to. The coefficient of *SUE of consensus* is 0.60 and is highly significant in the full sample, while *SUE of HQ analysts* is not significantly different than zero. Based on the results for the number of analysts in Table 5, specifications (3)-(4) ((5)-(6)) consider the subsamples of earnings announcements that have at least four (eight) HQ analysts, respectively, i.e., they include only earnings announcements for which *SUE of HQ analysts* is on average equally accurate (more accurate) than *SUE of consensus*. In all the regressions in this subsamples, the market still unambiguously reacts to *SUE of consensus* and not to *SUE of HQ analysts*. The coefficient of *SUE of consensus* is significant at 1%, while *SUE of HQ analysts* is insignificant.

We note that these two variables are highly correlated according to the last row of the table, for example, in the full sample of earnings announcements, the correlation is 0.93. This high correlation may cause a statistical problem because it increases the variance of the two coefficients. We therefore report the variance inflation factor (VIF), whose values above 10 but below 30 in columns (3)-(6) suggest that the high correlation between the two SUEs has a first-order effect on the standard deviation of the two coefficients, but it is not too severe (Belsley, Kuh, and Welsch 1980) to question the complete dominance of the SUE of consensus over the SUE of HQ analysts.

5.2. Trading based on the HQ analysts' forecasts

Because forecasts by the HQ analysts are relatively accurate, we suggest they can be used as a proxy for the actual earnings figures—one can replace actual earnings with the HQ analysts' forecast in the SUE formula. When this out-of-sample proxy for SUE is positive (negative), one can expect the market to react on average positively (negatively) to the announcement. A simple trading strategy for the immediate announcement reaction would be to buy the stock when this new surprise measure (the signal) is positive and shorting it when it is negative. Importantly, because the HQ analysts' forecast is available to investors at least one day before the actual announcement, such ability to predict announcement returns would indicate market inefficiency.

We conduct the empirical tests of this hypothesis in Table 7. For robustness, we consider two definitions of the out-of-sample surprise measure: the HQ analysts' forecast minus the LQ analysts' forecast and based on the HQ analysts' forecast minus the consensus. The first Buy Indicator is equal to one if the out-of-sample surprise measure is positive and is zero otherwise. We also use a stronger signal, Buy Indicator 2, which is one or zero if the out-of-sample surprise measure is, respectively, in the top or bottom quartile of its distribution in the previous year (ensuring that it stays out of sample). We regress the two-day BHAR on each of these Buy indicators and find their coefficients to be positive and significant. We conduct these tests in the full sample and the sample with 8 or more HQ analysts, in which they are more accurate than the consensus on average. Consistent with the hypothesis, the predictive ability of the buy indicator is twice as strong in the latter sample. The last line of the table reports the two-day announcement abnormal returns of the long minus short trading strategy based on the Buy indicators. The returns for all but one signal and sample combinations are statistically significant and, for example, reach 0.45% in the sample with 8 or more HQ analysts.

The overall conclusion from Table 6 is that the market reacts to *SUE of consensus* even when HQ analysts' forecasts are more informative than the consensus. Hence, the market seems to overreact to deviations from the consensus compared to deviations from the average estimate of HQ analysts. Another way to state this conclusion is that the market overreacts to LQ analysts and underreacts to HQ analysts.

6. Predicting market and industry returns

We next study the sources of the superior ability of HQ analysts. We conjecture that the informational advantage of the HQ analysts can come from their superior ability to process macroeconomic data or superior access to such data. The underlying logic for this conjecture is twofold. The underlying logic for this conjecture comes from the literature that it is a harder task for analysts to value a company in absolute terms than in relative terms, in which case analysts seem to have predictive ability in valuing companies compared to each other (e.g., Boni and Womack 2006; Da and Schaumburg, 2011). This conjecture is reinforced given our finding that analyst quality is persistent: the macro literature also finds that there is persistence in the ability to predict macro changes. For example, Ang, Bekaert, and Wei (2007) find that linear combinations of forecasts with weights computed based on past performance are better than simple averaging of forecasts. Finally, we note that our conjecture is consistent with the findings of Hutton, Lee and Shu (2012) that analysts' superiority in forecasting compared to management is mostly related to their understanding of the macro economy.

This leads to three testable implications for HQ analysts' forecasts, one related to the first moment and the other two related to the second moment of market performance. First, HQ analysts provide stock recommendations that can predict industry and market returns, in contrast to recommendations by the LQ analysts and the consensus. Second, HQ analysts' forecast dispersion indicates their uncertainty about the future state of the economy and should be associated with current and/or future market volatility. The third hypothesis is that the PEAD is greater during the periods of high uncertainty, which are determined by a greater dispersion of HQ analysts' forecasts.

We begin with testing whether our finding of market inattentiveness to the more accurate analysts can be useful for predicting the performance of the market and industry. To this end, we use our classification of analysts into high and low quality to construct the measures of average recommendation change. Typically, the month of the annual earnings news is associated with the greatest number of recommendation changes issued by analysts who cover the specific firm. This practice allows us to link most recommendation changes with either HQ or LQ analysts (we include only analysts who are ranked in the same month in which the recommendation is changed). We produce monthly average measures of recommendation changes based on the entire sample of recommendation changes, based only on HQ analysts' recommendation changes, and based only on LQ analysts' recommendation changes. In Table 8, Panel A the analysis is done at the market level, i.e., we calculate each of these measures for each calendar month. In Panel B, the analysis is at the two digit SIC code level, so that we calculate each of these measures for each month-industry.

The dependent variable in Table 8, Panel A is the monthly value-weighted market return. A recommendation is an integer 1-5, where 1 is strong buy, 5 is strong sell and 3 is hold. A recommendation change is the negative of the current recommendation of an analyst minus the previous recommendation of the analyst, so that a positive (negative) recommendation change is an upgrade (downgrade). *Mean rec. change* is the mean of all recommendation changes of all analysts who provided a recommendation change during the month in which the firm's earnings are announced. As an alternative measure, we also create the *Percentage rec. change down* which is the percentage of recommendations that are downgrades among all recommendation changes made by analysts who provided a recommendation change during the month in which the firm's earnings are announced. By definition, the two measures are negatively correlated. *High quality (Low quality) mean or percentage recommendation change* are analogous variables that are based only on recommendation changes of high (low) quality analysts. All independent variables are measured in the month prior to the dependent variable's month.

The regression results in Panel A reveal that the only variables that are statistically significant in explaining next month's market return are the variables that are based on recommendation changes of the HQ analysts. The coefficient of *High quality mean rec. change* is significant at the 10% level, while the coefficient of *Percentage rec. change down, high quality* is significant at the 5% level. In contrast, we see that the coefficients of the LQ analysts are not correlated with the next month's market return. These results suggest that HQ analysts' recommendations are informative about the future state of the market. Note that the null is that all recommendation changes are internalized by the market in the same month they are released, and hence no recommendation change should be predictive of the market return.

In Table 8, Panel B, we repeat the analysis of Panel A at the industry level. The results are somewhat different but provide similar interpretation. Only the recommendation change measures in specifications (2) and (6) are significant. In specification (2), the coefficient is significant at the 5% level, suggesting that HQ analysts' recommendations are not fully internalized by the market because they are predictive of the industry return in the following month. Interestingly, the downgrade measure of HQ analysts is short of being significant (*t-statistic* of -1.49). Nevertheless, in specification (6), we see that the low quality measure is positively correlated with industry return which implies that downgrades of the LQ analysts are positively correlated to industry returns, implying that the opposite of their recommendations is predictive of industry returns.

The second test of the relation between HQ analysts forecasting ability and overall market performance is based on the dispersion of analysts' forecasts. Our finding for HQ analysts' recommendations suggests that HQ analysts utilize both the macroeconomic and firm-specific information, while LQ analysts do not pay much attention to the market-wide information. This means that a relatively high forecast dispersion for the HQ analysts implies they are uncertain about the firm and market-wide factors, while a high dispersion for the LQ analysts tells only that they are uncertain about the firm.

Because the magnitude of forecast dispersion is firm-specific, it needs to be normalized to make it comparable across firms and years and, thus, usable as a measure of market uncertainty. We define a dispersion ratio, which is equal to the standard deviation of the HQ analysts' estimates divided by the standard deviation of all analysts' estimates (i.e., the consensus). A greater dispersion ratio then means HQ analysts are relatively more uncertain. We hypothesize that this uncertainty about the market can be related to high market volatility, which we measure by changes in VIX.

Table 9 reports regressions of the percentage changes in the monthly and annual average VIX on the dispersion of all analysts' (consensus) forecasts in Panel A and on the contemporaneous and lagged dispersion ratios in Panel B. We find that a higher dispersion ratio is associated with an increase in the VIX during the current period and predicts a greater VIX change in the following month and year. In contrast, the dispersion of the consensus is not associated with the VIX either contemporaneously or in the predictive fashion. The predictive relation we find implies that the HQ analysts see ahead of the market.

Finally, we test another prediction based on the dispersion ratio that a high dispersion ratio, implying high information uncertainty, is associated with a greater PEAD. We illustrate the findings in Figure 4 and report statistical significance results in Table 10. We calculate the PEAD using the calendar-time approach, in which the intercept is the monthly alpha. Recent studies show that the PEAD is largely driven by relatively illiquid stocks (Sadka, 2006; Ng, Rusticus, and Verdi, 2008) and limits to arbitrage (Chung and Hrazdil, 2011), which together with the approach to calendar-time analysis in Novy-Marx and Velikov (2016), motivates using value-weighted portfolio returns. Therefore, we report the findings for the value-weighted calendar-time portfolios and also find they are very similar if we use equal-weighted portfolios in untabulated results. We consider drift horizons from 1 to 11 months to avoid an overlap with the next annual earnings announcement. 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. To make our PEAD results comparable with the standard PEAD measurement in the literature, we use the consensus earnings surprise when we assign an announcement to the long (short) portfolio if earnings surprise is positive (negative).

Figure 4 reports the returns of the long minus short strategy for the sample of announcements with the high uncertainty, defined as the dispersion ratio greater than one, full sample, and the low uncertainty sample, in which the ratio is less than one. The high uncertainty PEAD is clearly above the full-sample PEAD, and the low uncertainty PEAD is below the full sample PEAD. Table 10 reports the statistical significance of, separately, long, short, and long minus short strategies for the two subsamples of high and low uncertainty. We find that the low uncertainty PEAD is not significant in any of these strategies except at the 11-month horizon (and even then, it is significant at the 10% level only). The high uncertainty PEAD is highly significant at all horizons, particularly, after good news. Overall, the findings for the economy and VIX predictability and the PEAD based on HQ analysts' information output are consistent with our findings for the immediate reaction—that the market is unaware of HQ analysts and that the PEAD is produced primarily during the periods of high information uncertainty at the market level.

7. Conclusion

We find that when firms have a large number of HQ analysts the consensus forecast has a higher absolute forecast error than that of the average estimate of these HQ analysts. The reason for this improved forecast ability lies in the tradeoff that exists between having many analysts' estimates, which reduces measurement error, and having HQ analysts' estimates, which on average have a smaller absolute forecast error. Because analysts' forecasting ability tends to persist, disregarding LQ analysts estimates can be beneficial in many cases. However, the market does not seem to know about these shortcomings of the consensus because the market reacts to deviations from the consensus and not to deviations from the average estimate of HQ analysts' estimates. A simple trading strategy for the immediate market reaction based on HQ analysts' estimates underlines that the market is inefficient in ignoring them.

We note that our approach suggesting to ignore the estimates of LQ analysts in order to improve on the consensus is only one way of generating an alternative to the consensus forecast. Other ideas may include a weighting scheme in which one gives more weight to the estimates by HQ analysts than the weight to LQ analysts or uses observable analysts' characteristics, such as analysts' tenure and brokerage firms, to improve the classification of HQ and LQ analysts.

Nevertheless, the advantage of our approach is in its simplicity of creating a rather straightforward single yardstick that is tractable and can easily be compared to the consensus.[5]

Our study also provides insights about which abilities make the HQ analysts superior and find that macroeconomic factors play an important role in their forecasts. The relation we find between greater forecast dispersion among HQ analysts, representing market-wide uncertainty, and the PEAD can have two different explanations based on the literature. A risk factor explanation following Abarbanell, Lanen, and Verrecchia (1995) suggests that investors discount both good and bad earnings surprises when uncertainty is high, which results in a PEAD. This finding can also be consistent with a behavioral bias explanation in the literature because higher uncertainty is related to a higher PEAD after good news and lower (more negative) PEAD after bad news (Zhang, 2006). Overall, the evidence for the immediate and long-term market response to earnings news suggests that the market is not justified in focusing only at the consensus and needs to utilize the information output of HQ analysts.

[5] A parsimonious simple model often performs better out-of-sample, as it is less subject to over-fitting of the data to a specific model. For example, it is well known in the literature that measurement errors are a major concern as they tend to reduce out-of-sample performance, e.g., DeMiguel et al. (2009) find that mean-variance optimization analysis underperforms the 1/N rule (equal investment across N assets) and conclude that the gain from optimal diversification is more than offset by estimation error.

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. $N(0, \sigma_B)$, and $\sigma_G < \sigma_B$. Analysts' forecasts are equal to their signals.

To obtain a more accurate forecast, 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

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

This simplifies to the following condition:

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

Hence, as the number of *G* analysts increases or as the number of *B* analysts decreases, the inequality is more likely to hold, so that we would prefer to consider the signals of only *G*-type analysts. The left-hand-side monotonically declines with n_G. Because the signal variances are unobserved, the model's testable predictions are based on the relative numbers of high and low quality analysts in the firm. Hence, this framework shows that as the fraction of high quality analysts increases, the condition for considering only high quality analysts' estimates is more likely to be satisfied, which would make 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.

Ang, A., Bekaert, G. and Wei, M., 2007. Do macro variables, assets markets, or surveys forecast inflation better? Journal of Monetary Economics 54, 1163-1212.

Bailey, W., Kumar, A. and D. Ng, 2011. Behavioral biases of mutual fund investors. Journal of Financial Economics 102, 1-27.

Belsley, David. A., Edwin. Kuh, and Roy. E. Welsch. 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley and Sons.

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

Bernard, V., and Thomas, J., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305-340.

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

Buraschi, A., Piatti, I., and P. Whelan, 2016, The cross section of subjective bond risk premia, Imperial College working paper.

Chakrabarty, B. and P. Moulton, P., 2012. Earnings announcements and attention constraints: The role of market design. Journal of Accounting and Economics 53, 612–634.

Chung, Dennis Y., and Karel Hrazdil, 2011. Market efficiency and the post- earnings announcement drift. Contemporary Accounting Research 28, 926-956.

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

Clement, M.B. and Tse, S.Y., 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters?. The Accounting Review, 78(1), 227-249.

Clement, M.B. and Tse, S.Y., 2005. Financial analyst characteristics and herding behavior in forecasting. The Journal of finance, 60(1), 307-341.

Conroy, R. and Harris, R., 1987. Consensus forecasts of corporate earnings: Analysts' forecasts and time series methods. Management Science, 33(6), 725-738.

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

Da, Z., and Schaumburg, E. , 2011. Relative valuation and analyst target price forecasts. Journal of Financial Markets 14, 161-192.

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.

Duru, A. and Reeb, D.M., 2002. International diversification and analysts' forecast accuracy and bias. The Accounting Review, 77(2), 415-433.

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.

Hirshleifer, D., Lim, S. and S.H Teoh, 2011. Limited investor attention and stock market misreactions to accounting information. Review of Asset Pricing Studies 1, 35-73.

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., 2015. Post-earnings-announcement drift in global markets: Evidence from an information shock. Review of Financial Studies, 28 (4), 1242-1283.

Hutton, A., Lee, L., and Shu, S., 2012. Do managers know better? The relative accuracy of management and analysts forecasts. Journal of Accounting Research 50, 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.

Jennings, R., 1987. Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. Journal of Accounting Research, 25, 90-110.

Kumar, A., 2009. Hard-to-value stocks, behavioral biases, and informed trading. Journal of Financial and Quantitative Analysis 44. 1375-1401.

Lakonishok, J. and Lee, I., 2001. Are insider trades informative?. Review of Financial studies, 14(1), 79-111.

Livnat, Joshua, and Richard R. Mendenhall. "Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts." Journal of Accounting Research 44, no. 1 (2006): 177-205.

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.

Makridakis, S. and Winkler, R.L., 1983. Averages of forecasts: Some empirical results. Management Science, 29(9), 987-996.

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

Michaely, R., Rubin, A., and A. Vedrashko, 2016. Further evidence on the strategic timing of earnigns news: Joint analysis of weekdays and time of day, Journal of Accounting and Economics 62, 24-45.

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.

Ng, J., Rusticus, T. O., and R. S. Verdi, 2008. Implications of transaction costs for the post–earnings announcement drift. Journal of Accounting Research 46, 661-696.

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

Piotroski, J.D. and Roulstone, D.T., 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. The Accounting Review,79(4), 1119-1151.

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.

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

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

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

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

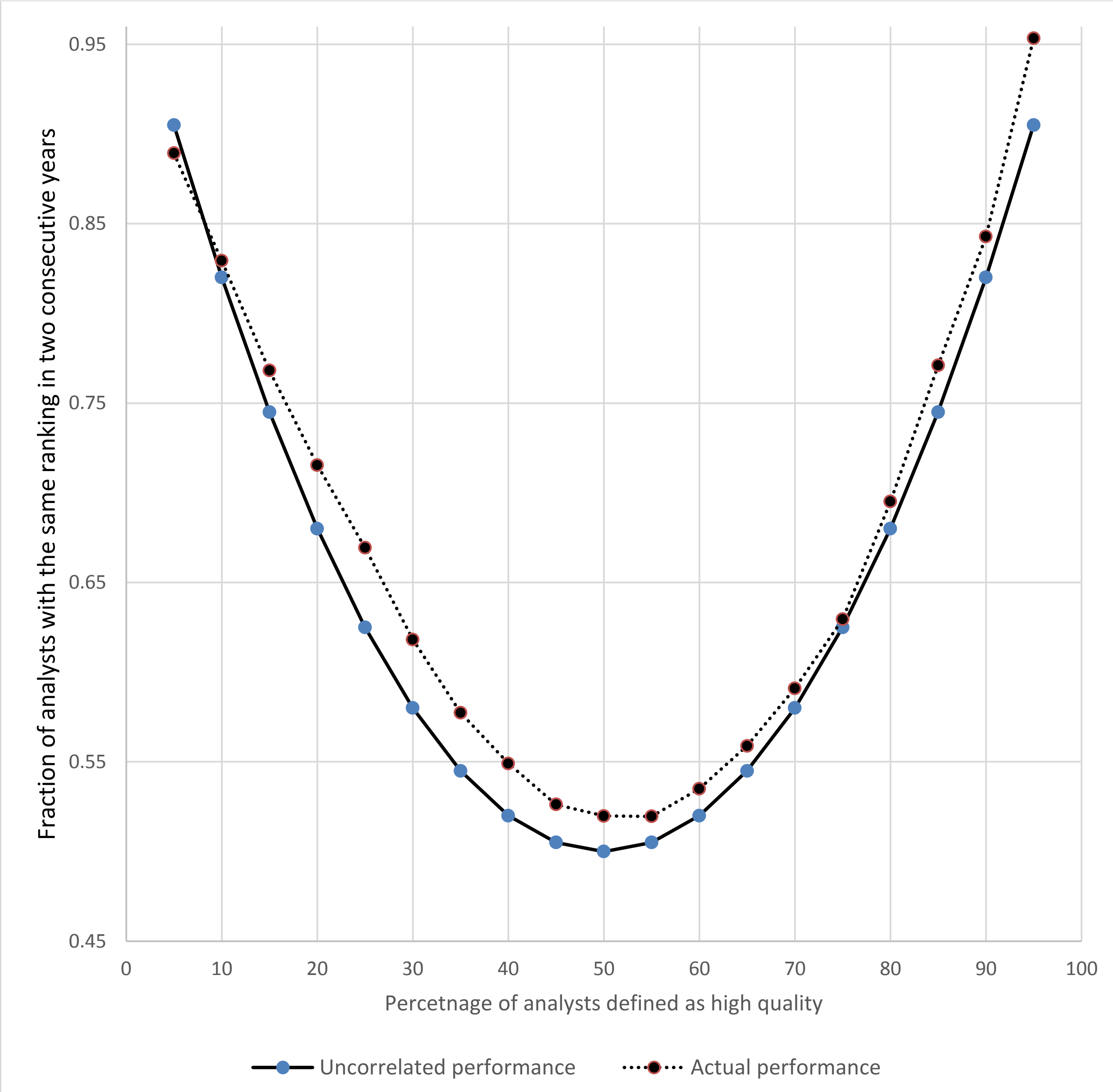

Figure 1: Persistence in analysts' forecasting performance. The figure depicts how the fraction of analysts retaining their ranking of either high or low forecasting quality in two consecutive years depends on the cutoff percentile in the definition of high quality analysts. High quality analysts are those whose closest absolute forecast errors are below the cutoff forecast error percentile for the firm 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 includes only firms that are followed by ten or more analysts. The uncorrelated performance line shows what analysts' forecasting performance would be if it were uncorrelated between consecutive years.

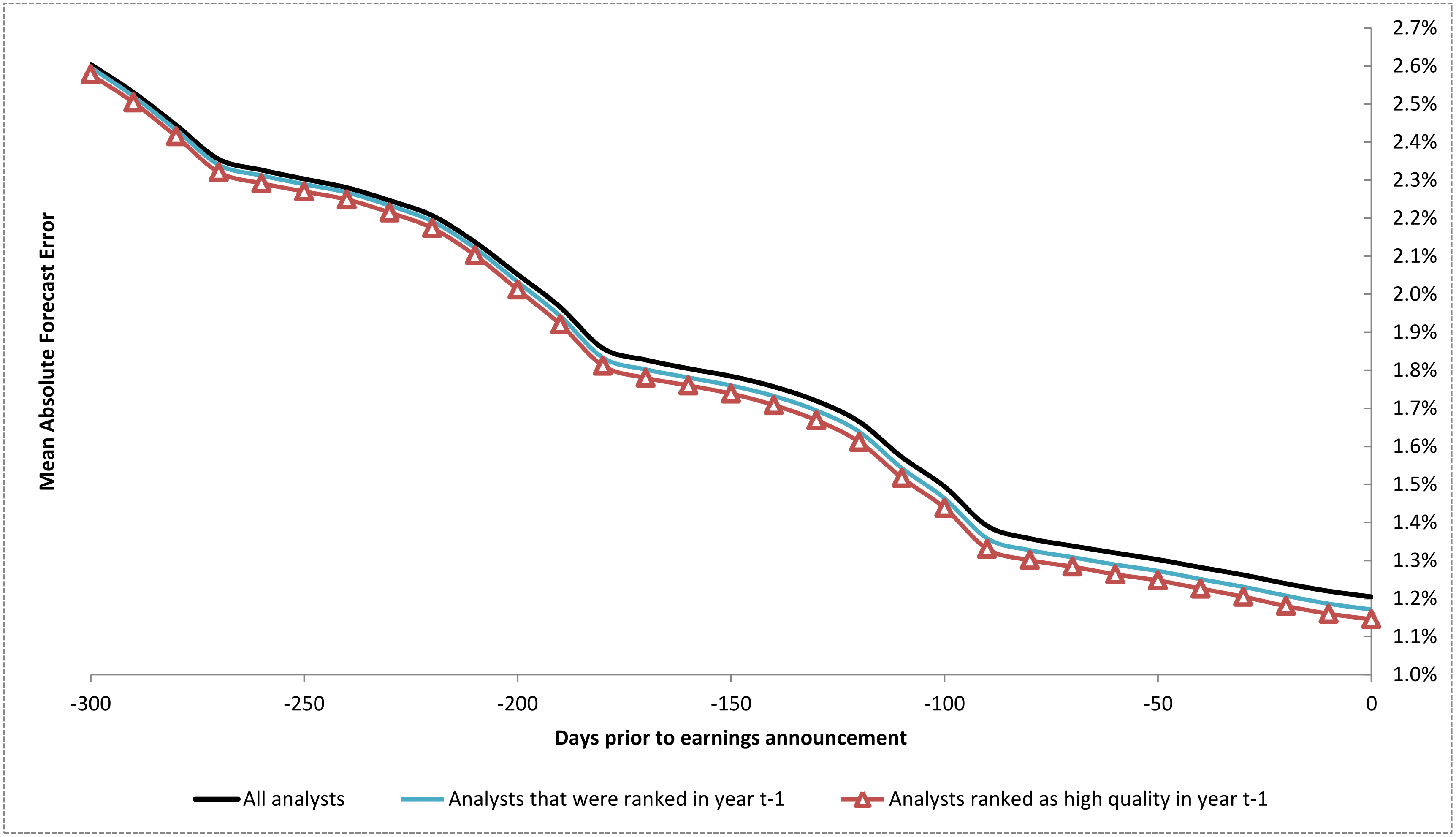

Figure 2: The absolute forecast error of analysts' estimates during the period starting 300 days before the earnings announcement day. The sample includes only the firms that have at least one analyst ranked as high quality in year t-1 who continued following the firm in year t. High quality analysts are those whose forecast errors are below the median forecast error for the firm in year t-1, and the depicted forecast errors are in year t. Absolute forecast errors are measured for individual analysts and averaged at the firm-announcement level and then across firms and announcements for the particular day during 300 days before to the announcement day.

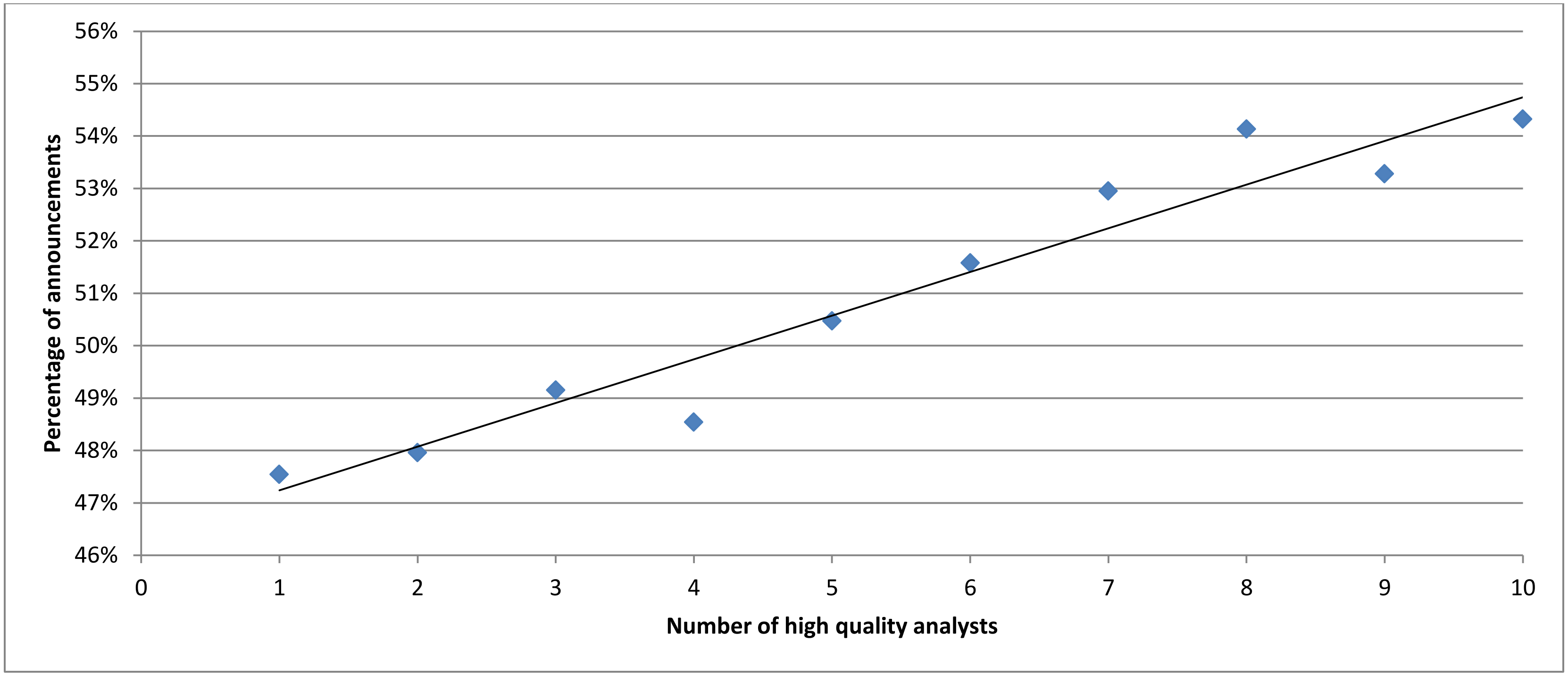

Figure 3: Percentage of announcements for which the average forecast of high quality analysts is more accurate than the consensus forecast. The figure depicts the relation between the number of high quality analysts that provide forecasts and the fraction of cases in which the absolute earnings surprise of the average estimate of high quality analysts is smaller than the absolute earnings surprise of all analysts. The straight line is a fitted regression line for these two variables. High quality analysts are those whose forecast errors are below the median forecast error for the firm's earnings announcement in year t-1.

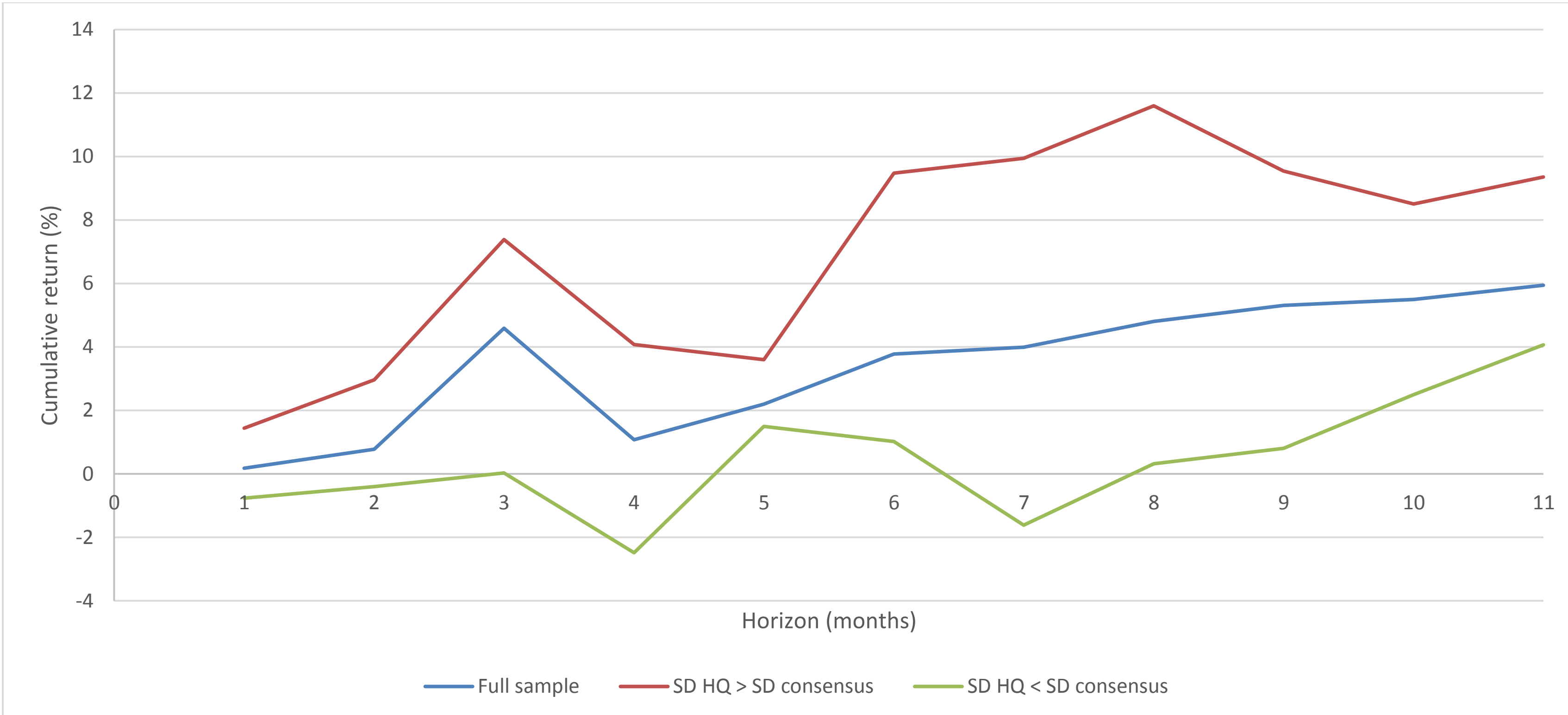

Figure 4. Cumulative post-announcement drifts depending on analysts' uncertainty

The figure shows the cumulative drift for 1 to 11-month horizons following earnings announcements. The horizontal axis is the drift horizon, which is the number of months a stock is held in the calendar-time portfolio. The monthly value-weighted portfolio returns are regressed on the four Fama-French-Carhart factors to obtain the cumulative drift, which is the intercept of the regression (monthly alpha) multiplied by the portfolio's horizon. The vertical axis is the return of the long-minus-short strategy, which is long is earnings surprise is positive and short if earnings surprise is negative, where earnings surprise is defined based on the consensus estimate. The graphs are for the full sample and two subsamples in which the standard deviation of high quality analysts' forecasts (SD HQ) is greater or smaller than that of the consensus (SD consensus). High quality analysts are those whose forecast errors are below the median forecast error for the firm's earnings announcement in year t-1.

Table 1: Sample description

The table summarizes the EPS one-year-ahead (FPI=1) estimates in the I/B/E/S detailed file during the reporting years 1992-2015. Each line is a subsample of the line above.

Sample	Number of firm-year-analysts	Number of firm-analysts	Number of analysts	Number of firms
All analysts on I/B/E/S	861,349	273,328	18,608	13,530
All analysts + Compustat	829,126	263,651	18,459	13,277
Firms have at least 4 analysts	768,320	245,652	17,930	9,606
Analysts appearing in consecutive years in the same firm	486,239	165,534	13,482	7,831
Analysts appearing in consecutive years in the same firm and have value-weighted measures in both years	196,972	80,617	8,793	5,156

Table 2: Persistence in analyst quality and analyst characteristics

Panel A provides analysts' characteristics. High (low) quality analysts are those whose closest absolute forecast errors are below (above) the median closest absolute forecast error for the firm's earnings announcement in year t-1. The closest absolute forecast error is defined in Figure 1. *Tenure* is the number of years since the analyst first appeared in the I/B/E/S file. *Tenure in company* 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 that are in the analyst's brokerage house. *Coverage* is the number of firms covered by the analyst. Panel B reports models in which where the dependent variables are an indicator that equals one if the analyst was high quality and zero otherwise (columns 1 and 2) and the analyst's absolute forecast error (columns 3 and 4). *Firm size* is the firm's market value of equity equal to the stock price at the beginning of the calendar year times the number of shares outstanding. *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 total assets. *Book-to-market* is the book value of common equity divided by the market value of equity. *Number of analysts* is the number of analysts following the firm. All independent variables are measured in year t-1. The coefficients in specifications (1) and (2) are marginal probability effects. All specifications include the intercept. Robust standard errors are clustered by firm. *z*-statistics and *t*-statistics are in parentheses in the first two and last two columns, respectively. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A. Analyst characteristics

N=768,320	High quality analysts	Low quality analysts	Difference in means (t-statistic)
Tenure	7.06	7.00	0.06*** (4.68)
Tenure in company	2.97	2.90	0.07*** (8.86)
Brokerage house size	65.94	63.31	2.63*** (18.68)
Coverage	17.58	17.54	0.05 (1.51)

Panel B. Predicting analysts' forecasting performance

	High quality analyst indicator		Absolute forecast error	
	(1)	(2)	(3)	(4)
High quality analyst	0.0414***	0.0407***	-0.0007***	-0.0007***
indicator at t-1	(25.54)	(25.12)	(-16.07)	(-15.84)
Firm size	0.0034***	0.0011***	-0.0061***	-0.0061***
	(10.52)	(3.14)	(-25.12)	(-25.15)
Annual return	-0.0003	0.0015	-0.0008***	-0.0009***
	(-0.50)	(0.89)	(-5.70)	(-5.69)
Leverage	0.0006	0.0007	0.0058***	0.0058***
	(0.37)	(0.18)	(6.64)	(6.61)
Book-to-market	0.0001	0.0001	0.00002	0.00002
	(1.40)	(0.59)	(1.48)	(1.48)
Number of analysts	0.0007***	0.0019***	0.0002***	0.0002***
	(12.40)	(14.52)	(10.80)	(10.71)
Tenure		0.0018***		-0.00004***
		(4.39)		(-7.14)
Tenure in company		0.0058***		0.00001
		(9.60)		(1.49)
Brokerage house size		0.0004***		-0.000001
		(10.61)		(-0.23)
Coverage		-0.0008***		0.00003***
		(-12.32)		(8.38)
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects			Yes	Yes
Observations	486,239	486,239	486,239	486,239
Adj. R-squared			0.344	0.345

Table 3: Relation between analyst quality across firms

The table reports how analysts' forecasting quality in one firm is related to their quality in other firms, with Panel A showing the contemporaneous and Panel B showing the predictive relations. High and low quality analysts and the high quality analyst indicator are defined in Table 2. *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. Panel B reports probit regressions predicting the high quality analyst indicator based on analysts' performance in the other firms in the previous year. All independent variables are from the previous year, and the reported coefficients are marginal probability effects. The control variables are defined in Table 2. All specifications include the intercept. 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. High and low quality analysts' forecasting performance in other firms

		This firm		
		High quality	Low quality	t-statistic: Difference from 50%
Other firms	High quality	54.5%	45.5%	52.2***
	Low quality	42.5%	57.5%	90.1***

Panel B. Probit predicting the high quality analyst status

	High quality in other firms	High quality analyst in year t-1	Tenure	Tenure in company	Brokerage house size	Coverage	Number of observations
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 4: Distribution of analysts of different types

The table reports the average number of high quality, low quality, and unranked analysts in the earning announcement year. High and low quality analysts are defined in Table 2. Unranked analysts are those who did not provide annual forecasts for the firm in the previous year.

High Quality Analysts		Low Quality Analysts		Unranked Analysts	
Number	Fraction	Number	Fraction	Number	Fraction
1	0.265	1.659	0.399	1.908	0.336
2	0.371	2.083	0.346	2.123	0.283
3	0.405	2.695	0.327	2.560	0.268
4	0.403	3.672	0.349	2.969	0.249
5	0.419	4.315	0.340	3.365	0.240
6	0.432	4.945	0.336	3.682	0.232
7	0.432	5.782	0.342	4.092	0.226
8	0.439	6.449	0.342	4.367	0.218
9	0.448	7.011	0.337	4.700	0.214
10 or more	0.461	9.201	0.337	5.715	0.202

Table 5: The number of high quality analysts and the improvement in forecast accuracy

The table compares the accuracy of the average forecast of the high quality analysts and that of the consensus. SUE of Consensus (SUE of high quality analysts) is the difference between the actual earnings and the average forecast provided by all analysts (average forecast provided by high quality analysts) normalized by stock price at the beginning of the year. High quality analysts are defined in Table 2. *Accuracy Improvement* is the percentage reduction from the absolute SUE of the consensus to the absolute SUE of the high quality analysts. *t*-statistics is the difference in means between the absolute SUE of consensus and that of the SUE of high quality analysts.

Number of high quality analysts	Number of earnings announcements	Absolute SUE of Consensus	Absolute SUE of high quality analysts	Accuracy Improvement	t-statistics Absolute SUE difference
1	8,001	0.00980	0.01076	-9.78%	8.69***
2	10,512	0.00790	0.00816	-3.29%	4.73***
3	7,625	0.00656	0.00667	-1.56%	2.38**
4	4,814	0.00587	0.00585	0.36%	-0.44
5	3,690	0.00481	0.00480	0.32%	-0.36
6	2,884	0.00459	0.00455	0.91%	-0.79
7	2.107	0.00416	0.00411	1.34%	-1.15
8	1.751	0.00441	0.00428	2.88%	-2.43**
9	1,388	0.00380	0.00370	2.67%	-1.96**
10 or more	3,854	0.00346	0.00337	2.59%	-3.44***

Table 6: Immediate Reaction to Earnings News

The table reports the earnings response coefficients for two measures of earnings surprise. The first one is based on all analysts' forecasts (SUE of consensus), and the second one is based only on the forecasts of high quality analysts. High quality analysts are defined in Table 2. The dependent variable is the buy and hold abnormal return (based on the four factor model) during the earnings announcement day and the following trading day. SUE of consensus is the forecast error of the average of all estimates provided by all analysts. SUE of high quality analysts is the forecast error of the average of estimates provided by the high quality analysts. All other variables are defined in Table 2. Specifications (1) and (2) use the entire sample of earnings announcements, and specifications (3)-(4) and (5)-(6) use the sample of earnings announcements by firms that have at least four and eight high quality analysts, respectively. Correlations and the variance inflation factor (VIF) are provided for the two SUE measures for each of the samples in the last rows. The intercept and year fixed effects are included in all regressions. Robust standard errors are clustered by firm. *t*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Full Sample		4 or more high quality analysts		8 or more high quality analysts	
	(1)	(2)	(3)	(4)	(5)	(6)
SUE of consensus	0.60253***	0.60887***	0.53810*	0.55386*	1.44798***	1.44178***
	(6.54)	(6.61)	(1.88)	(1.94)	(2.57)	(2.56)
SUE of high quality analysts	-0.00840	-0.00902	0.08363	0.06951	-0.66611	-0.66672
	(-0.10)	(-0.11)	(0.32)	(0.27)	(-1.15)	(-1.15)
Size		-0.00000		-0.00000		-0.00000
		(-1.48)		(-1.46)		(-1.26)
Annual Return		-0.00065		0.00061		0.00378**
		(-1.01)		(0.63)		(1.99)
Leverage		0.00570***		0.00238		0.00296
		(3.62)		(1.01)		(0.71)
Book to market		0.00002		-0.00019***		-0.00127
		(0.37)		(-9.08)		(-1.05)
Number of analysts		0.00000		-0.00005		-0.00005
		(0.06)		(-0.63)		(-0.33)
Observations	44,709	44,709	20,221	20,221	6,968	6,968
Adjusted R-squared	0.0128	0.0138	0.00922	0.00930	0.0140	0.0146
Correlations of SUE of consensus and SUE of high quality analysts	0.931		0.971		0.981	
VIF						
SUE of Consensus	7.49	7.50	17.41	17.45	26.29	26.32
SUE of high quality analysts	7.49	7.49	17.40	17.44	26.31	26.34

Table 7: Abnormal return on earnings announcement day

The dependent variable is the buy and hold abnormal return (based on the four factor model) during the earnings announcement day and the following trading day. High and low quality analysts are defined in Table 2. *Buy Indicator 1* and *Buy Indicator 2* are based on the *Signal* variable, which is equal to (high quality average forecast minus low quality average forecast) in columns (1), (2), (5), and (6), and (high quality average forecast minus consensus forecast) in the remaining columns. *Buy Indicator 1* equals one if *Signal* is positive and zero if it is negative. *Buy Indicator 2* equals one if *Signal* is positive and greater than the 25th percentile of positive *Signals* in year t-1, and zero if *Signal* is negative and smaller than the 75th percentile of negative *Signals* in year t-1. All regressions include the intercept and year fixed effects, and standard errors are clustered at the firm level. *t*-statistics are provided in parentheses. The bottom part of the table provides returns of the trading strategy that is long if *Buy Indicator* is 1 and short if *Buy Indicator* is 0. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Full Sample				8 or more high quality analysts			
	Signal: High quality average minus low quality average		Signal: High quality average minus consensus		Signal: High quality average minus low quality average		Signal: High quality average minus consensus	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Buy indicator 1	0.00190**		0.00147*		0.00353**		0.00261	
	(2.49)		(1.93)		(2.15)		(1.62)	
Buy indicator 2		0.00062*		0.00072**		0.00167**		0.00197**
		(1.96)		(2.29)		(2.15)		(2.45)
Size	-0.000001*	-0.000001*	-0.000001*	-0.000001*	-0.000001	-0.000001*	-0.00000	-0.00000
	(-1.71)	(-1.76)	(-1.70)	(-1.84)	(-1.55)	(-1.64)	(-1.52)	(-1.50)
Annual return	-0.00011	-0.00038	-0.00012	-0.00066	0.00418**	0.00403	0.00418**	0.00594**
	(-0.18)	(-0.53)	(-0.19)	(-0.88)	(2.17)	(1.61)	(2.15)	(2.33)
Leverage	0.00430***	0.00627***	0.00431***	0.00522***	0.00154	0.00107	0.00159	0.00046
	(2.69)	(3.38)	(2.70)	(2.81)	(0.36)	(0.19)	(0.38)	(0.08)
Book to market	-0.00000	0.00000	-0.00000	0.00000	-0.00167	-0.00088	-0.00166	-0.00076
	(-0.04)	(0.03)	(-0.04)	(0.04)	(-1.21)	(-0.69)	(-1.20)	(-0.61)
Number analysts	0.00007	0.00003	0.00007	0.00003	0.00001	0.00004	0.00001	0.00000
	(1.25)	(0.38)	(1.28)	(0.51)	(0.08)	(0.23)	(0.08)	(0.01)
Observations	44,709	32,339	44,709	32,850	6968	4134	6968	4049
R-squared	0.002	0.002	0.001	0.002	0.006	0.007	0.006	0.009

	Return of two day trading strategy (%)							
Long-short return	0.19**	0.16*	0.14*	0.18**	0.34**	0.36*	0.24	0.45**

Table 8: Predictive Market and Industry Returns

The dependent variables are the firm's monthly value-weighted return and the industry return (based on two-digit SIC code) in Panels A and Panel B, respectively. The analysis considers all recommendation changes made on the same month in which the analysts are ranked (based on observed actual annual earnings). A recommendation is an integer 1-5, where 1 is strong buy, 5 is strong sell, and 3 is hold. A recommendation change is the negative of the current recommendation of an analyst minus the previous recommendation of analyst, so that a positive (negative) recommendation change is an upgrade (downgrade). High and low quality analysts are defined in Table 2. *Mean rec. change* is the mean of all recommendation changes of all analysts who provided a recommendation change during the month in which the firm's earnings is announced. *Percentage rec. change down* is the percentage of recommendations that are downgrades out of all recommendation changes made by analysts who provided a recommendation change during the month in which the firm's earnings is announced. High (low) quality mean rec. change is the mean of recommendation changes of high (low) quality analysts who provided a recommendation change during the month in which the firm's earnings is announced. Percentage rec. change down, high (low) quality is the percentage of recommendations that are downgrades out of all recommendation of high (low) ranked analysts who provided a recommendation change during the month in which the firm's earnings is announced. All independent variables are measured in the month prior to the dependent variable's month. The models use Newey-West standard errors with three lags. *t*-statistics are provided in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Predicting Market Returns

	(1)	(2)	(3)	(4)	(5)	(6)
Market return (t-1)	0.07426	0.07124	0.07181	0.07679	0.07676	0.07093
	(0.93)	(0.91)	(0.90)	(0.98)	(0.99)	(0.91)
Mean rec. change	0.01362					
	(1.41)					
High quality mean rec. change		0.01563*				
		(1.86)				
Low quality mean rec. change			0.00494			
			(0.67)			
Percentage rec. change down				-0.04227		
				(-1.62)		
Percentage rec. change down, high quality					-0.05154**	
					(-2.04)	
Percentage rec. change down, low quality						-0.01117
						(-0.59)
Constant	0.00918***	0.00959***	0.00814***	0.02286**	0.02626***	0.01164
	(3.53)	(3.60)	(2.96)	(2.55)	(3.09)	(1.60)
Obs. (number of months)	266	266	266	266	266	266
Adjusted R-squared	0.0125	0.0181	0.0074	0.0165	0.0252	0.0074

Table 8 (continued)

Panel B: Predicting Industry Returns

	(1)	(2)	(3)	(4)	(5)	(6)
Industry return (t-1)	0.0481***	0.0477***	0.0485***	0.0485***	0.0482***	0.0480***
	(2.78)	(2.75)	(2.79)	(2.79)	(2.78)	(2.77)
Mean rec. change	0.0011					
	(1.26)					
High quality mean rec. change		0.0021**				
		(2.39)				
Low quality mean rec. change			-0.0003			
			(-0.36)			
Percentage rec. change down				-0.0018		
				(-0.72)		
Percentage rec. change down, high quality					-0.0036	
					(-1.49)	
Percentage rec. change down, low quality						0.0041*
						(1.89)
Constant	0.0087***	0.0088***	0.0085***	0.0092***	0.0096***	0.00740***
	(10.52)	(10.66)	(10.31)	(7.92)	(9.34)	(7.09)
Observations	6,792	6,792	6,792	6,792	6,792	6,792
Adjusted R-squared	0.00229	0.00289	0.00207	0.00214	0.00243	0.00257

Table 9: Dispersion, dispersion ratio, and VIX

The dependent variable in the regressions is *RVIX*, which is the monthly percentage change of the VIX index in specifications (1)-(4) or over the year (the average of end of month VIX divided by the average end of month VIX in the previous year) in specifications (5)-(8). *Dispersion* is the firm market value-weighted standard deviation of all analysts' forecast errors in a given month or year, and *Dispersion ratio* is the firm market value-weighted average of the ratio (standard deviation of high quality analysts' forecast errors)/(standard deviation of all analysts' forecast errors). Forecast error is the 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. High quality analysts are defined in Table 2. Specifications (1)-(4) are for the first six months of each calendar year. *t*-statistics are provided in parentheses and calculated using the Newey-West standard errors with three lags in the annual regression and the robust estimator of Huber and White otherwise. *, **, *** represent 10%, 5%, 1%, significance, respectively.

Panel A: Forecast dispersion and VIX

	Monthly				Annual			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-0.010	-0.016	-0.007	-0.015	-0.153	-0.081	-0.146	-0.012
	(-0.55)	(-0.88)	(-0.40)	(-0.79)	(-0.75)	(-0.76)	(-1.10)	(0.20)
Dispersion	1.349		0.101		66.851		62.938	
	(0.27)		(0.02)		(0.77)		(1.31)	
Dispersion at t-1		4.340		3.610		41.339		11.270
		(0.65)		(0.53)		(0.75)		(0.36)
RVIX at t-1			-0.111	-0.109			0.398***	0.391***
			(-1.06)	(-1.06)			(2.08)	(5.98)
Adj. R-squared	-0.007	-0.004	-0.004	-0.001	0.030	-0.012	0.162	0.093
Number of obs.	133	133	132	132	23	23	23	23

Panel B: Dispersion ratio and VIX

	Monthly				Annual			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-0.065**	-0.054**	-0.065**	-0.056**	0.067	-0.296**	-0.504	-0.222
	(-2.48)	(-2.31)	(-2.45)	(-2.34)	(0.11)	(-2.17)	(-1.58)	(-1.57)
Dispersion ratio	0.083**		0.082**		-0.056		0.604	
	(2.58)		(2.55)		(0.09)		(0.91)	
Dispersion ratio at t-1		0.067**		0.070**		0.383***		0.290*
		(2.04)		(2.11)		(3.64)		(1.83)
RVIX at t-1			-0.120	-0.130			0.467***	0.369***
			(-1.18)	(-1.23)			(3.52)	(4.66)
Adj. R-squared	0.032	0.020	0.034	0.026	-0.047	0.0367	0.1144	0.1396
Number of obs.	133	133	132	132	23	23	23	23

Table 10: Post-earnings announcement drift and analysts' relative uncertainty

The table reports the cumulative drift for 1 to 11-month horizons following earnings announcements. Announcements are divided into two subsamples in which the standard deviation of high quality analysts' forecast errors (SD of high quality) is greater or smaller than that of the consensus (SD of consensus). 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. High quality analysts are defined in Table 2. *, **, *** represent 10%, 5%, 1%, significance, respectively, based on *t*-statistics.

Drift horizon (months)	SD of high quality > SD of consensus			SD of high quality < SD of consensus		
	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*