DISCUSSION PAPER SERIES

DP15057
(v. 2)

Do Analysts Learn from Each Other? Evidence from Analysts' Location Diversity

Ling Cen, Yuk Ying Chang and Sudipto Dasgupta

FINANCIAL ECONOMICS

ISSN 0265-8003

Do Analysts Learn from Each Other? Evidence from Analysts' Location Diversity

Ling Cen, Yuk Ying Chang and Sudipto Dasgupta

Discussion Paper DP15057
First Published 17 July 2020
This Revision 12 February 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Financial Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Ling Cen, Yuk Ying Chang and Sudipto Dasgupta

Do Analysts Learn from Each Other? Evidence from Analysts' Location Diversity

Abstract

We show that when the locations of analysts covering a firm are geographically more diverse, the individual forecasts of the analysts for that firm are less correlated. More geographical diversity of co-analyst locations leads to more accurate individual analyst forecasts. This suggests that analysts assign weights to co-analysts' forecasts when making their own forecasts, and the individual forecasts become more accurate due to a diversification effect. Moreover, in line with efficient weighted average forecasting, our results indicate that the weights assigned to peer forecasts vary with measures of the precision of the analyst's signal and those of the peers. Overall, our evidence suggests observational learning in the analyst setting. Our empirical design avoids typical pitfalls of outcome-on-outcome peer effects (Angrist, 2014) by showing that an analyst's expected absolute forecast error (proportional to standard deviation) is affected by the covariance of co-analyst's forecast errors (as captured by their locational diversity).

JEL Classification: G24, D83

Keywords: Information Diversity, learning, Herding, Analyst Forecasts

Ling Cen - ling.cen@cuhk.edu.hk
Chinese University of Hong Kong

Yuk Ying Chang - y.chang@massey.ac.nz
Massey University

Sudipto Dasgupta - s.dasgupta@cuhk.edu.hk
Chinese University of Hong Kong and CEPR

Acknowledgements
This paper was previously circulated under the title: "Do Analysts Learn from Each Other? – A New Approach to an Unresolved Question". We thank Kee-Hong Bae and Hongping Tan for providing us with the location data for analysts. We also thank Jarrad Harford, Gilles Hilary, Chayawat Ornthanalai, Rik Sen, and Frank Yu for helpful comments. Yuk Ying Chang acknowledges the Massey Business School for providing financial support.

Do Analysts Learn from Each Other? Evidence from Analysts' Location Diversity

LING CEN

Chinese University of Hong Kong

(ling.cen@cuhk.edu.hk)

YUK YING CHANG

Massey University

(y.chang@massey.ac.nz)

SUDIPTO DASGUPTA[≠]

Chinese University of Hong Kong, ABFER, CEPR and ECGI

(s.dasgupta@cuhk.edu.hk)

First Version: September 2019[#]

This version: December, 2021

Abstract. We show that when the locations of analysts covering a firm are geographically more diverse, the individual forecasts of the analysts for that firm are less correlated. More geographical diversity of co-analyst locations leads to more accurate individual analyst forecasts. This suggests that analysts assign weights to co-analysts' forecasts when making their own forecasts, and the individual forecasts become more accurate due to a diversification effect. Moreover, in line with efficient weighted average forecasting, our results indicate that the weights assigned to peer forecasts vary with measures of the precision of the analyst's signal and those of the peers. Overall, our evidence suggests observational learning in the analyst setting. Our empirical design avoids typical pitfalls of outcome-on-outcome peer effects (Angrist, 2014) by showing that an analyst's expected absolute forecast error (proportional to standard deviation) is affected by the *covariance* of co-analyst's forecast errors (as captured by their locational diversity).

Key Words: Information Diversity; Learning; Herding; Analyst Forecasts

JEL Classifications: G24; D83

≠ Corresponding author. Department of Finance, Chinese University of Hong Kong, CUHK Business School, Shatin, NT, Hong Kong.

This paper was previously circulated under the title: "Do Analysts Learn from Each Other? – A New Approach to an Unresolved Question". We thank Kee-Hong Bae and Hongping Tan for providing us with the location data for analysts. We also thank Jarrad Harford, Gilles Hilary, Chayawat Ornthanalai, Rik Sen, and Frank Yu for helpful comments. Yuk Ying Chang acknowledges the Massey Business School for providing financial support.

Do Analysts Learn From each Other? Evidence from Analysts' Location Diversity

Abstract

We show that when the locations of analysts covering a firm are geographically more diverse, the individual forecasts of the analysts for that firm are less correlated. More geographical diversity of co-analyst locations leads to more accurate individual analyst forecasts. This suggests that analysts assign weights to co-analysts' forecasts when making their own forecasts, and the individual forecasts become more accurate due to a diversification effect. Moreover, in line with efficient weighted average forecasting, our results indicate that the weights assigned to peer forecasts vary with measures of the precision of the analyst's signal and those of the peers. Overall, our evidence suggests observational learning in the analyst setting. Our empirical design avoids typical pitfalls of outcome-on-outcome peer effects (Angrist, 2014) by showing that an analyst's expected absolute forecast error (proportional to standard deviation) is affected by the *covariance* of co-analyst's forecast errors (as captured by their locational diversity).

Key Words: Information Diversity; Learning; Herding; Analyst Forecasts

JEL Classifications: G24; D83

Security analysts are important intermediaries in financial markets, and one of their major roles is to generate earnings forecasts for the companies they follow. The finance industry, media, and market participants pay considerable attention to the "consensus" forecast of earnings, and hence the informativeness of the forecasts matters for the functioning of financial markets. However, how informative the earnings forecasts are has been questioned. It has been suggested, for example, that because analyst forecasts are in the public domain, analysts' career concerns (Scharfstein and Stein, 1990; Trueman, 1994) and the possibility of free-riding (Hirshleifer and Teoh, 2003) create incentives for *uninformed herding*. Because such uninformed herding potentially reduces the idiosyncratic information that analyst forecasts can provide, the informativeness of the "consensus" estimate could be adversely affected. For example, Huang, Krishnan, Shon, and Zhou (2017) estimate that 63 percent of analysts exhibit uninformed herding behaviour, which they define as taking "… actions to drift toward the prevailing consensus, *regardless of the information contained in the consensus*", and that 16 percent exhibit anti-herding.[1]

On the other hand, since analysts can observe each other's forecasts, they could potentially learn from other analysts' forecasts, leading to better aggregation of information. In other contexts, such behaviour has been identified and is termed *informational* or *rational herding*, and occurs because of *observational learning* in a setting where individuals form posterior beliefs based on their own private signals and the actions of others (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992).[2] To the best of our knowledge, little evidence

[1] A large body of empirical evidence finds support for the career-concern related uninformed herding view. In an influential paper, Hong, Kubik, and Solomon (2000) find that inexperienced analysts are less likely to provide "bold" forecasts than experienced analysts, and are also more likely to lose their jobs after providing inaccurate or bold forecasts. Clement and Tse (2005) examine other analyst characteristics that could be related to career concern, such as prior forecast accuracy, brokerage size, forecast frequency, and the number of companies/industries that the analyst follows, and find similar evidence.

[2] For example, Zhang (2010) finds that patients waiting for transplant kidneys are more likely to turn down a kidney after observing other patients' rejection decisions. Zhang and Liu (2012) find evidence of observational learning among lenders in *Prosper.com* – the largest and oldest microloan market in the U.S. In crowdfunding

exists that observational learning occurs in the important setting of security analysts' earnings forecasts.[3] Given the bulk of evidence on uninformed herding, this may suggest that career concerns or similar incentives are so pervasive that they largely negate incentives for learning and information production by most analysts. If so, this is a serious concern for market efficiency.

Our paper makes two main contributions to the literature. First, we provide novel evidence that when making their own forecasts, analysts (1) incorporate *idiosyncratic* information from the other analysts' prior forecasts, and (2) assign weights to the prior forecast of other analysts that are systematically related to the potential informativeness of these forecasts. Overall, these results provide strong evidence of observational learning.

To evaluate whether analysts are influenced by the prior forecasts of other analysts, we rely heavily on the concept of analysts' *locational diversity* and how this affects the accuracy of both the consensus and individual analyst forecasts. Information advantage of local analysts (e.g., Malloy, 2005; Bae, Stulz, and Tan, 2008; O'Brien and Tan, 2015) suggests that analysts collect substantial amount of information from local sources in addition to corporate disclosures. However, local information sources also contain noise, and we hypothesize that this noise component is likely to be less correlated when the analysts are not based in the same region (i.e., analysts are geographically more diversely located). Thus, the more diversely located the analysts are, the more likely that the idiosyncratic noise components will be cancelled out (similar to a portfolio diversification effect) when the individual forecasts are aggregated in the consensus estimate. We elaborate below how we build on this idea to test for observational learning by individual analysts.

---

platforms, it has been widely documented that early contributions from individuals with information advantage on project quality attract subsequent contributions (see, for example, Kuppuswamy and Bayus, 2018).

[3] Kumar, Rantala, and Xu (2021) and Phua, Tham, and Wei (2021) are two contemporaneous papers that also examine analyst peer effects. We discuss these papers below.

The key insight is that, for an *individual* analyst, if the other analysts are based in more diverse locations, their forecast noises will be less correlated; however, this can improve the forecast accuracy of that individual analyst *only if he assigns a positive weight to the mean forecast of the other analysts* (or, more generally, uses a weighted average of these forecasts and his own signal, for his own forecast). Importantly, common shocks to future earnings that all analysts incorporate in their forecasts do not play any role here, since common components will not get diversified away even when the other analysts are geographically more diverse.

We validate this intuition in a model of analyst learning outlined in Section 3 of the paper. In this model, we show that when an analyst optimally assigns a positive weight to the prior forecasts of the other analysts, the analyst's mean (absolute) forecast deviation or forecast error (henceforth denoted *IEER*) is positively related to the covariance of the idiosyncratic noise in the forecasts of co-analysts (i.e., other analysts covering the same firm who have already generated their forecasts). The model also generates implications regarding how the strength of the association between the individual analyst's *IEER* and the covariance depends on the information content of the signals of the co-analysts and that of the analyst in question.

Our empirical constructs heavily exploit the idea that the geographical location of analysts is a major determinant of the precision of their informative signals as well the diversity of these signals. Our research thus contributes to the literature on how analysts' geographical location matters for their forecasting activity (Malloy, 2005; Bae, Stulz, and Tan, 2008; O'Brien and Tan, 2015; Jennings, Lee, and Matsumoto, 2017; Chen, Mayew, and Yan, 2018; Gerken and Painter, 2020), and how diversity in analysts' characteristics affect the accuracy of the consensus forecast (Merkley, Michaely, and Pacelli, 2020).

For our empirical setting, we measure analysts' locational diversity in terms of a Hirschman-Herfindahl Index (*HHI*) of analyst locations, which is an inverse measure of locational diversity. Consistent with the idea that more locational diversity implies less correlated noise in the analyst forecasts, we find that as the *HHI* decreases, (i) the average covariance between the noise components in forecasts of all analysts following a firm in a given year, and (ii) the absolute forecast error of the consensus forecast, both decrease. We then show that an individual analyst's *IEER* is positively related to the inverse diversity of the co-analysts following the same firm in the same year whose forecasts for the fiscal year are available to the analyst prior to his last forecast. For brevity, we refer to such analysts as co-analysts from now on. For this purpose, we construct an *HHI* based on the locations of these co-analysts, *OHHI* (which captures the covariance of the idiosyncratic signal noises of the other analysts). We then test our hypothesis regarding how the strength of the association between the *IEER* and the *OHHI* (the *OHHI*-sensitivity of *IEER*) is affected by various proxies for signal informativeness, as implied by the model of observational learning.

The problems of identifying observational learning from actions of peers are well recognized (Manski, 1993; Angrist, 2014). Our approach avoids the typical problems associated with outcome-on-outcome peer effects by relating our main dependent variable, a transformation of the standard deviation of an analyst's forecast error (the *IEER*), to the *covariance* of the forecast noises of co-analysts who have generated forecasts for the same firm in the same year as a particular analyst. The covariance is captured by our measure of the locational diversity of the co-analysts, *OHHI*.

We carefully design our tests to make sure we are identifying the effect of diversity on individual analyst forecast error. First, our regressions showing the effect of *OHHI* on *IERR* incorporate analyst and year (year×analyst) fixed effects, as well as firm and year (year×firm) fixed effects. Using analyst-year fixed effects largely rules out the concern that the results are driven by analysts who are clustered in a region (e.g., New York, where more than 40 per cent of the analysts in our sample are based),[4] or by analysts who are at a particular stage in their careers and they anchor their forecasts to the prior forecasts of co-analysts due to career-concerns. Moreover, using firm-year fixed effects implies that we exploit within-firm variation in *OHHI*, so that the results are unlikely to be driven by effects that change the firm's information environment.

Second, our results exploit *changes* in *OHHI* caused by analyst relocation, entry, and exit "events". The relocation of an analyst ("event analyst") causes, for all other analysts following the firms covered by this analyst, either an increase, decrease, or no change, in the *OHHI*. Since for the same relocation event, an analyst with overlapping coverage as the relocating analyst experiences heterogeneous changes in *OHHI* for the firms they both cover, it is unlikely that the change in *OHHI* is associated with factors that could affect forecast accuracy of the co-analyst in a particular way. A similar argument applies to analysts whose *OHHI* is affected by "new entrants" – that is, analysts who show up in I/B/E/S for the first time in our sample period – and those who exit (i.e., no longer show up in I/B/E/S). Our regression sample (henceforth referred to as the *Combined Sample*) combines the firm-years that experience changes in analyst location due to the relocation, entry, and exit events, but it excludes the forecasts of the triggering event analysts. On average, an analyst in the *Combined Sample* is exposed to 10 events in a given year. We show that both the within-analyst year and within-firm year variation of the change in *OHHI* is quite significant – the mean range (maximum minus minimum) of the change in *OHHI* being 70 percent and 55 percent on the sample mean *OHHI*, respectively. When we regress the change in *IERR* (from one year before

[4] For example, if several analysts covering a firm are based in New York, the New York-based analyst has a more diversified group of peer analysts (lower *OHHI*) than another analyst not based in New York and covering the same firm. If New York-based analysts happen be more accurate in their forecasts, we would be picking up a positive association between *OHHI* and forecast error.

the event year to one year after) on the change in *OHHI* of individual analysts over the corresponding period, we find a significantly positive association in the *Combined Sample*.

The *OHHI* for an analyst increases (decreases) when location decisions by event analysts covering the same firm result in more (less) correlation of the signal noises of the co-analysts. Thus, if the forecast error of that analyst increases (decreases), this suggests that the analyst assigns some weight to the forecasts of the peers following the same firm. There are two possible reasons why the analyst might do so. First, analysts might want to herd with other analysts and anchor their forecasts to the mean prior forecast of other analysts (uninformed herding). Second, they could also be using the mean forecasts of other analysts to come up with a more efficient weighted-average forecast that combines their own information and that represented in the mean of the other analysts' forecasts (observational learning).[5]

To investigate whether analysts learn from their peers – that is, assign weights to peer forecasts for informational reasons – we appeal to a literature that shows that analysts make more accurate forecasts when they are located closer to the firm's or peers' operations (Malloy, 2005; Bae, Stulz, and Tan, 2008; Jennings, Lee, and Matsumoto, 2017). We do two sets of tests of the learning hypothesis. First, for each analyst, we calculate the distance of the analyst's location from that of the firm for which the analyst is issuing a forecast. We argue that the longer this distance, the less precise is the analyst's own signal. Similarly, we compute the average distance of the analyst's location from that of all peer firms in the same 3-digit SIC industry as the firm being covered that this analyst also covers, and argue that the analyst's

[5] We note that the mean is not the most efficient aggregator of individual forecasts. As is well known, when there is common information and signals are not independent, the mean overweights the common information, and though unbiased, it is neither consistent nor minimum variance (see. Kim, Lim, and Shaw (2001) who explore this issue in the context of analyst forecasts). However, the mean is almost universally used as an aggregator of individual predictions, and in the particular setting of analyst forecasts, the consensus forecast is the most widely used and followed metric. Thus it could be expected that under observational learning, when it comes to making their own forecasts, the analysts would directly rely on the prevailing consensus as a way to incorporate the information content from other analysts' forecasts.

information quality is poorer if this distance is larger. Our model shows that the analyst puts more weight on the forecasts of the co-analysts when his information quality is poorer, and as a result, the sensitivity of the analyst's forecast error to the OHHI is higher. Our empirical results provide strong support for this prediction. Second, we also examine the consequences of better or worse information quality of the co-analysts, which would change the weight assigned by the analyst to their forecasts. We find that when the average distance of all co-analysts from the same 3-digit SIC industry peer firms they cover is greater, the sensitivity of the analyst's forecast error to *OHHI* decreases. On the other hand, the sensitivity is higher when the number of same 3-digit SIC peer firms covered by co-analysts is larger.

Our focus on observational learning from co-analysts' forecasts of the *same* firm also distinguishes our paper from two recent papers. Kumar, Rantala, and Xu (2021) examine "social learning" from the forecasts of co-analysts of other firms covered by the co-analysts and the analyst in question. In particular, they show that if the forecasts of these analysts are ex post over-optimistic, the analyst becomes less optimistic in his forecast of the focal firm. In addition, if these co-analysts make bold positive or negative forecasts for the firms they cover, the analyst's forecasts are more likely to be bold. Phua, Tham, and Wei (2021) examine within-brokerage peer learning. They find that analysts who occupy more central positions in the brokerage's network based on overlap of industrial coverage make more accurate forecasts.

Our study is also related to Merkley, Michaely, and Pacelli (2020) and Gerken and Painter (2020). Merkley, Michaely, and Pacelli (2020) show that cultural diversity of financial analysts improves quality of both consensus and individual forecasts. We argue that economic mechanisms through which cultural and geographical diversities affect analyst forecasts are different: cultural diversity generates different interpretations for the same information source, which may facilitate or impede learning. Therefore, Merkley, Michaely, and Pacelli (2020) argue that the relation between cultural diversity and forecast accuracy, theoretically unclear ex ante, is ultimately an empirical issue.[6] In contrast, geographical diversity implies different information sources and, therefore, our theoretical framework generates clear empirical predictions ex ante. In a paper contemporaneous to ours, Gerken and Painter (2020) utilize satellite data on the number of cars in the parking lots of U.S. retailers and analyst location information compiled from historical filings of form U4 to study whether analysts overweight local information. While both papers show that geographical diversity of analysts reduces error in consensus forecast, they emphasize different sources of influence on individual analyst forecasts. Contrary to existing literature mentioned above that suggests information advantage to local analysts, Gerken and Painter (2020) argue that analysts can overweight their local information signals. In contrast, we focus on the extent to which analysts incorporate other analysts' forecasts into account, and show that their forecasts errors are lower when the other analysts are geographically more diverse.[7]

The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 develops our hypotheses. Section 4 describes our data and variables. Section 5 reports and discusses empirical results. Section 6 states conclusions.

2. Related Literature

2.1 Uninformed herding, Learning, and Analyst forecasting Activity

The setting in which earnings forecasts occur is interesting because each analyst's forecast is in the public domain, and analysts revise their forecasts multiple times during a particular forecasting period, knowing the full set of forecasts that have already been made.

---

[6] For example, Merkley, Michaely, and Pacelli (2020) point out "ex ante, the relation between diversity and consensus forecast accuracy is unclear" and "to what extent cultural diversity improves output quality remains an empirical issue" in their introduction.

[7] Another difference between our papers is that the individual forecast error in Gerken and Painter (2020) is defined as the difference between the forecasted earnings and actual earnings, scaled by lagged price (which is essentially a measure of forecast optimism), while our individual forecast error is the absolute value of their measure.

The public nature of these forecasts affects the incentive for analysts to generate informative forecasts. On the positive side, to the extent that there is information content in other analysts' forecasts, analysts can learn from each other's forecasts (Surowiecki, 2004; Bloomfield and Hales, 2009; Clement et al., 2011). For example, if a particular analyst revises a prior forecast, that could alert the remaining analysts to the possibility of new information being available, resulting in more information production.

On the negative side, as has been extensively discussed in the literature, the public nature of the forecasts can also distort analysts' incentives. One consequence of this is "uninformed herding." Such herding behaviour could represent a form of free-riding (Hirshleifer and Teoh, 2003), which in turn would decrease the incentives of analysts to invest in information acquisition and make more efficient forecasts. Uninformed herding could also occur if career or reputational concerns cause analysts to supress their own information when this involves a significant deviation from the consensus (Scharfstein and Stein, 1990; Trueman, 1994). When an analyst's reputation for accuracy is not yet established, the cost of being "bold" but wrong might be particularly high. Instead, analysts might prefer to play it safe and herd to show that they "get it." Several authors find evidence consistent with such reputational herding. Hong et al. (2000) find that (1) experienced analysts are more likely to provide bold forecasts than inexperienced analysts, and (2) inexperienced analysts are more likely than experienced analysts to lose their jobs after providing inaccurate or bold forecasts. Clement and Tse (2005) extend Hong et al. (2000) by examining the role of other analyst characteristics that could mitigate career concern, such as prior forecast accuracy, brokerage size, forecast frequency, and the number of companies/industries that the analyst follows.

Some studies attempt to estimate the propensity of analysts to herd or "anti-herd", i.e., behave in a contrarian manner. Both Bernhardt, Campello, and Kutsoati (2006) and Huang et al. (2017) argue that uninformed herding is the tendency of analysts to bias their forecasts away from their posterior estimates of earnings towards the consensus – *irrespective of the informativeness of the consensus itself*. Anti-herding, on the other hand, corresponds to strategically differentiating the forecast from the consensus forecast. From this premise, the two studies design different test statistics to estimate the propensity of analysts to herd or anti-herd, and they come to sharply different conclusions. While Bernhardt et al. (2006) conclude that analysts anti-herd, Huang et al. (2017) estimate that 63 percent of analysts exhibit uninformed herding behavior, and 16 percent exhibit anti-herding.

Keskek, Tse, and Tucker (2014) argue that the reputation-herding theory (Sharfstein and Stein, 1990; Trueman, 1994) predicts that late forecasters during a specific "information discovery" or "information analysis" phase of information production are herders, so their forecasts will be less accurate than earlier ones. In contrast, the "trade-off theory" (Guttman, 2010) suggests that late forecasters are learners and thus should be more accurate. The authors find that earlier forecasters are of higher quality within each phase in terms of forecast accuracy improvement (relative to peer's outstanding forecasts), forecast boldness, and price impact – in other words, there is little evidence of "slow learning" by late forecasters.[8]

2.2 Local Information, Economic Decisions, and Analyst Activity

There is a growing literature in Economics and Finance on the importance of local information, particularly for analyst activity. Coval and Moskovitz (2001) find that investors earn abnormal returns when they tilt their holdings towards local stocks, suggesting the importance of local news.[9] Ivkovic and Weisbenner (2005) find that the average household receives higher returns from its local holdings relative to its nonlocal holdings, suggesting that investors can exploit local knowledge. Garcia and Norli (2012) find that firms that have

[8] Note that to interpret these results, one has to assume that later forecasters cannot exactly mimic the forecasts of earlier forecasters.
[9] Van Nieuwerburgh and Veldkamp (2009) provide a rationale for understanding why local information asymmetry persists even when investors can choose to learn what others know.

operations mainly in a few locations earn higher returns than geographically dispersed firms. Kang, Stice-Lawrence, and Wong (2019) use satellite data on car counts in the parking lots of U.S. retailers and find that local institutional investors trade in the same direction as this information. They also find that investors sometimes overweight local information. Chen (2017) finds that managers overreact to local economic news at the firm's headquarters; cutting investment in plants when the local economy at the firm headquarter is performing less well.

Turning to analyst activity, Malloy (2005) finds that forecast revisions by local analysts has stronger effect on stock prices. Jennings, Lee, and Matsumoto (2015) find that geographical distance of a firm's location from that of the analyst matters for the latter's forecast quality. In addition, they find that managers frequently refer to other firms in the same location in earnings conference calls, suggesting the relevance of local information for analyst activity. Chen, Mayew, and Yan (2018) find that analysts working in the same office and covering different firms in the same location are more accurate in their forecasts and generate stronger stock price response, suggesting that sharing local information via social interactions is useful. Bae, Stulz, and Tan (2008) and O'Brien and Tan (2015) confirm that geographical proximity generates information advantage of analysts in a cross-border and an IPO setting, respectively.

3. Hypothesis Development

3.1 Analysts' Geographical Dispersion and Forecast Errors

To develop our hypotheses, we consider a simple model of analyst learning. Here, we present the main ingredients of that model, discuss the association of the model parameters with our empirical constructs, and develop the main hypotheses. Details of the model and derivations are provided in Appendix B.

In the model, analyst learning from the prior forecasts of other analysts takes a particularly simple form: an analyst arrives at his "optimal" forecast via a suitably chosen weighted average of his forecast, based on his own information set (excluding the forecasts of the other analysts) and the mean of the prior forecasts of the other analysts. We note that a weighted average or mean is not, in general, an efficient aggregator of individual forecasts (Kim et al., 1998). Nonetheless, the practise of combining individual predictions into the mean is ubiquitous – not only is the consensus mean forecast that combines *all* the analyst forecasts an extremely common and well-publicized metric, many other forecasts such as central bank forecasts of macroeconomic variables are also mean forecasts. Accordingly, we assume that analysts also use the mean when aggregating the forecasts of other analysts. However, they determine how much weight to assign to this mean relative to their own forecasts.

3.2 Forecast Environment and the Mean Absolute Deviation (MAD)

We now describe the forecast environment. Let x_i denote an individual analyst i's forecast of the earnings-to-price ratio. We assume that

$$x_{it} = f + \epsilon_{it} \qquad (3.1)$$

where f denotes the actual earnings-to-price ratio, and ϵ_{it} denotes all other sources of the analyst-specific forecast noise. We assume that ϵ_{it} is normally distributed with mean 0 and variance σ^2. Although we do not explicitly model it, the analyst could be assumed to be a Bayesian forecaster who has a prior forecast on *f* and observes a signal which is the earnings plus a noise term.[10] In what follows, to avoid confusion with an analyst's absolute forecast error (defined below), we refer to the error ϵ_{it} as *forecast noise* or *signal noise*.

[10] A Bayesian analyst has a prior on the true earning f (assumed to have a Normal distribution with mean f and precision h) and observes $z_i = f + \xi_{it}$, where ξ_{it} is an analyst-specific noise term with precision s_i. Let $f = \mathrm{f} + \eta$. The Bayesian forecast conditional on z_i is $x_{it}=(h/(h+s_i))\mathrm{f} + (s_i/(h+s_i))z_i = (h/(h+s_i))(f-\eta)+(s_i/(h+s_i))(f+\xi_{it},)=f+(s_i/(h+s_i))* \xi_{it}+(h/(h+s_i))(- \eta)=f+\zeta_{it}$, where ζ_{it} has mean zero and precision $h+s_i$. See Barron, Kim, Lim, and Stevens (1998) for an exposition of the Bayesian framework.

The individual (absolute) forecast error is given by $|x_{it} - f|$, that is, the individual absolute deviation. Einhorn, Hogarth, and Klempner (1977) derive the expression for the mean absolute deviation (*MAD*) for the case of normally distributed errors. This is given by:

$$MAD = 2\sigma\phi(0) \tag{3.2}$$

where ϕ denotes the density of the standard normal variable. Clearly

$$\frac{\partial(MAD)}{\partial\sigma} > 0. \tag{3.3}$$

Thus, the *MAD* is increasing in the standard deviation of the forecast noise.

3.3 Analysts' Location Diversity and the Consensus Forecast Error

Extension to the *MAD* of the consensus forecast is straightforward. The consensus is the mean of *N* forecasts x_{it}, given by $X_t = f + \frac{\sum_{i=1}^{N} \epsilon_{it}}{N}$. The standard deviation is given by $\sigma_N = \sqrt{\frac{1}{N}[\sigma^2 + (N-1)\sigma_{nm}]}$, where σ_{nm} denotes the covariance between the noise in the signals for any pair of analysts *n* and *m*. The expression for the *MAD* of the consensus forecast follows from (3.2) by replacing σ with σ_N.

Our first hypothesis relates to the effect of greater location diversity of analysts on the *MAD*. Throughout, our measure of location diversity is the Herfindahl-Hirschman Index (*HHI*) of analyst locations, which is an *inverse measure* of diversity. Suppose there are *Q* possible locations (we leave details for Section 4). The *HHI* is defined as

$$HHI_j = \sum_{q=1}^{Q} \left(\frac{Number\ of\ analysts\ in\ location\ q}{Total\ number\ of\ analysts\ covering\ firm\ j}\right)^2 \tag{3.4}$$

where the denominator is the total number of analysts following firm j in year t, and the numerator is the number of analysts in a particular location q.

Hypothesis 1 (H1): (i) The average covariance between analysts' forecast noise is increasing in the *HHI* of analyst locations. (ii) The *MAD* of the consensus forecast is increasing in the *HHI* of analyst locations.

Hypothesis H1(i), if true, is a direct test of the assumption that we will maintain throughout, namely, that the idiosyncratic noise in analyst signals will be less correlated if the analyst locations are more geographically dispersed, as proxied by the inverse of the *HHI*. Note that from Equation (3.3) and the expression for σ_N, Hypothesis 1(ii) immediately follows via the effect of location diversity on σ_{nm} as postulated in (i).

We next turn to individual analyst's forecast errors (*IEER*). A key assumption underlying our framework is that analyst i, in making his last forecast, assigns a weight (1-α_i) to the mean of the latest forecasts of all other analysts, and a weight α_i to his own signal. In Appendix B, we show that the analyst can improve forecast accuracy by choosing an appropriate value of α_i. The focus of much of our subsequent discussion is the effect of analyst geographic location on the weight 1-α_i assigned to the mean forecast of the other analysts. Here, we provide intuitive arguments supporting the development of our hypotheses. Appendix B provides formal derivations.

Since analysts assign positive weights on the mean forecasts of other analysts to improve the accuracy of their forecasts, they incorporate the idiosyncratic components of the signals of other analysts in their own forecasts. The diversification principle therefore implies that the less correlated the individual idiosyncratic components of the other analysts' signals are, the lower will be analyst i's expected forecast error. For our empirical tests, we construct an *HHI* based on the geographical location of the *other* analysts (co-analysts) whose latest forecast for the same firm in the same year is available to the analyst *i*, and label this as *OHHI*. The *OHHI* is therefore an analyst-specific inverse measure of the geographical location diversity of the co-analysts for each firm-year, and as long as the analyst assigns a positive weight on the mean forecast of the co-analysts, we show in the model that it will be positively related to analyst *i*'s mean absolute forecast error:

Hypothesis 2 (H2): The sensitivity of the absolute forecast error of an individual analyst (*IEER*) with respect to the *OHHI* – the locational *HHI* based on the locations of the co-analysts following the firm with prior forecasts – is positive.

In other words, if we regress an individual analyst's forecast error on that analyst's *OHHI*, as long as $0<\alpha_i<1$, the regression coefficient will be positive.

In Appendix B, we derive an explicit solution for α_i and show that $0< \alpha_i <1$. However, it is important to note that Hypothesis 2, if true, can hold even under uniformed herding, and not only when analysts are engaging in informed herding or observational learning, as is assumed in the model and the derivation of α_i. This is because under uninformed herding, the analyst's forecast is also a weighted average of his own posterior of the earnings and the average of the forecasts of the other analysts.[11] However, as we show in Appendix B, the observational learning model has several unique implications regarding how α_i and hence the sensitivity of the *IEER* to *OHHI* depends on the quality of analyst *i*'s signal and that of the co-analysts.

Since our next two hypotheses build on these results, we summarize these intuitive results here. We show that lower precision of the analyst's own signal causes the analyst to place higher weight on the remaining analysts' mean forecast. Moreover, an analyst assigns a

[11] See, for example, Huang et al. (2017).

higher weight to the mean forecast of the co-analysts if these co-analysts have, ceteris paribus, more precise signals (i.e., better information quality).[12]

We next discuss tests of the learning hypothesis. Following Malloy (2005) and Bae, Stulz, and Tan (2008), we argue that an analyst i's information about a particular firm j 's earnings is likely to be more precise when (i) the analyst *i* is located closer to the firm *j*, and (ii) the average distance of the analyst i's location from that of all other firms in the same 3-digit SIC industry as firm *j* that they cover ("peer" firms) is shorter. We identify higher information quality of co-analysts of analyst *i* covering firm *j* in terms of (i) their average proximity to peer firms of the firm *j* (i.e., firms in the same 3-digit SIC industry as the firm *j* that they also cover), and (ii) the number of such peer firms covered by all the co-analysts of analyst *i* covering firm *j* as a group.

Hypothesis 3 (H3): Analysts will assign a higher weight to the forecasts of other analysts (lower α_i) and the sensitivity of *IERR* to *OHHI* will be higher, if, ceteris paribus,

(i) the information quality of the signals of other analysts following the firm, measured as discussed above, is higher. The sensitivity of *IERR* to *OHHI* will be unambiguously larger in this case.

(ii) the analyst's own information quality, measured as discussed above, is poorer. The sensitivity of *IERR* to *OHHI* will be larger under plausible conditions in this case.

4. Data and Variables

[12] As the model makes clear, these comparative static results "work through" the effects of the underlying parameters on α_i, and the effect of α_i on the sensitivity of the *IEER* to *OHHI*. We focus on the *OHHI* sensitivity of *IEER* rather than the direct effect on α_i (which captures the sensitivity of an individual analyst's forecast to the average of other analysts' forecasts, i.e. the "Consensus" forecast) for two reasons. First, whether analyst learning affects analyst forecast *error* is of independent interest. Second, the error term in a regression of an individual analyst's forecast on the Consensus will be correlated with the Consensus, leading to biased estimates.

We obtain realized earnings and analyst earnings forecasts from the I/B/E/S Unadjusted Detail History files. Analyst location was published in Nelson's Directory of Investment Research, available up to 2008. Our main analysis is limited to the period 1994-2010 (under the assumption that our analyst location information in 2008 is roughly accurate over a span of two additional years).[13]

We primarily consider the last annual earnings per share (EPS) forecast of each analyst for each firm fiscal year. We use the I/B/E/S actual earnings instead of the Compustat earnings because the I/B/E/S has a policy of reporting actual earnings numbers that are consistent with forecasts, i.e., it excludes the same items from actual EPS numbers that the majority of analysts exclude from their forecasts (Christensen, 2007). The sources of accounting and financial data are Compustat and CRSP, respectively.

4.1 Main Dependent Variables

The main dependent variable is the absolute forecast error, where the forecast is either a consensus forecast of all analysts covering a firm or an individual forecast of an analyst. The consensus forecast error is computed as follows: for each firm year, we first calculate the difference between the mean of the earnings forecasts of all analysts and the corresponding actual earnings being forecasted, scaled by the share price as of the end of the previous fiscal year. To obtain our primary dependent variable (*ERR*), we take the absolute value of the calculated difference. For the analyses of individual forecasts of the analysts, our dependent variable is *IERR*, calculated in a similar way as *ERR*, as the absolute difference between the

[13] We thank Kee-Hong Bae and Hongping Tan for providing us with the location data for analysts, which include the analyst's identity number, the identity number of the affiliated equity research firm, and the time information. The coverage is limited before 1994. Nelson Publishing Inc. stopped producing its Directory of Investment Research after 2008. Using Nelson's Directories, we also manually check and exclude observations for which there is insufficient information to clearly identify the location of the analyst.

individual forecast value of each analyst and the actual value being forecasted, scaled by the prior fiscal year-end share price.

Another dependent variable is *RHO*, which measures the degree to which analysts' forecasts covary relative to the overall level of uncertainty in a firm year. It is the Fisher transformation (Fisher, 1915, 1921) of the variable ρ, estimated following Barron et al. (1998):

$$RHO = \frac{1}{2} ln\left(\frac{1+\rho}{1-\rho}\right), \text{where } \rho = \frac{SE-\frac{D}{N}}{\left(1-\frac{1}{N}\right)D+SE} \quad (4.1)$$

where *SE* is the squared difference between the mean of the last annual EPS forecasts of all analysts and the actual value of EPS being forecasted (both scaled by the market share price as of the end of the previous fiscal year), in a firm year; *D* is the variance of these last individual forecasts (scaled by the market share price); and *N* is the number of analysts in that firm year. When analysts are more geographically diverse, there is lower common uncertainty in relation to the total uncertainty, whereby both ρ and *RHO* are smaller. Hence, analyst geographical diversity and *RHO* are expected to have a negative relationship.

4.2 Main Explanatory Variables

The key explanatory variable is the location diversity of analysts, constructed for each firm year. Figure A3 in Appendix C displays the US geographical distribution of analysts in 2006. Not surprisingly, a large number of analysts are located in New York City. Figure A4 shows the US distribution of headquarters of the listed firms in our sample in 2006, which is more dispersed than the analyst location distribution shown in Figure A3. Our location sample covers the period between 1994 and 2010, for which the location data has good coverage. Our main geographical classification is based on U.S. Metropolitan Statistical Areas (MSAs) and cities that are not located in a U.S. MSA ("non-MSA cities").

4.2.1 Construction of group location diversity measure

For each firm year, we pool all analysts following the same firm in the same year together and then divide them into different location groups based on the U.S. MSAs and non-MSA cities in which they are located. We then calculate the Herfindahl-Hirschman Index for each firm year ($HHI_{j,t}$) as an inverse measure of geographical location diversity:

$$HHI_{j,t} = \sum_{all\ q} \left(\frac{Number\ of\ analysts\ in\ location\ q_{j,t}}{Total\ number\ of\ analysts_{j,t}} \right)^2 , \tag{4.2}$$

where the subscripts *j* and *t* index firm and year, respectively. $Total\ number\ of\ analysts_{j,t}$ is the total number of analysts producing forecasts for firm *j* in year *t*. *Number of analysts in location* $q_{j,t}$ is the number of analysts producing forecasts for firm *j* in year *t* in the q^{th} location. The largest possible value of the *HHI* is 1.0, when all analysts are based in the same MSA or a non-MSA city. The lower the *HHI*, the more diverse the geographical location of analysts. Hence, the *HHI* is an inverse measure of diversity.

4.2.2 Construction of "Other Location Diversity" measure (OHHI)

Particularly useful for our analysis is an analyst-firm-year specific HHI, which we term *OHHI*, which for analyst *i* and firm *j* is the HHI based on the locations of prior co-analysts (i.e. those co-analysts who have forecasts published before the publication of analyst *i*'s last forecast in the same firm fiscal year under consideration),

$$OHHI_{i,j,t} = \sum_{all\ q} \left(\frac{Number\ of\ prior\ co\text{-}analysts\ in\ location\ q_{j,t}}{Total\ number\ of\ prior\ co\text{-}analysts_{j,t}} \right)^2 \tag{4.3}$$

$OHHI_{i,j,t}$ is the other location-Herfindahl-Hirschman Index for analyst *i* of firm *j* in year *t*. We exclude analyst *i* while computing this measure and co-analysts whose only forecast for the fiscal year is published after analyst *i*'s last forecast. $Total\ number\ of\ prior\ co - analysts_{j,t}$ is the total number of the prior co-analysts producing annual EPS forecasts for firm *j* in year *t*. $Number\ of\ analysts\ in\ location\ q^{\neq i}_{j,t}$ is the number of the prior co-analysts in the q^{th} location where q is one of the U.S. MSAs or non-MSA cities. Similar to *HHI*, the lower the *OHHI*, the more diverse the geographical location of the prior co-analysts.

4.2.3 Proxies for information quality

Based on a strand of literature that suggests that geographical proximity is associated with better information quality (e.g., Malloy, 2005; Bae, Stulz, and Tan, 2008; and Jennings et al., 2017), we adopt two measures of an analyst's geographical distance from firms. The first measure is the distance in kilometres between the analyst's location and the business location of the firm being covered ("BUS DIS"). The other measure is the weighted average distance between the analyst's location and 50 states, where the weights reflect the degree of relevance of the states for the operation of the firm being covered ["GEO DIS"] (García and Norli, 2012).[14] Based on BUS DIS and GEO DIS, we construct the following 3x2 distance measures to capture information quality of analysts: (1) the analyst's geographical distance from the firm in question (2) the average of the analyst's geographical distance from the peer firms, in the same 3-digit SIC industry as that of the firm under consideration, covered by the analyst, and (3) the average of the individual average distances of all the prior co-analysts from peer firms (in the same 3-digit industry as that of the firm under consideration) that they cover.

The literature also shows that the degree of analyst coverage and industry experience explain forecasting performance (e.g., Bradley, Gokkaya, and Liu (2016); Merkley, Michaely, and Pacelli (2017)). Consistently, we expect that prior co-analysts' coverage of peer firms, in the same industry, improves information quality of their forecasts for the firm under consideration. We thus compute the number of peer firms, in the same 3-digit SIC industry as

[14] We obtain the latitude and longitude of firms' business locations from (https://www3.nd.edu/~mcdonald/) on 24 February 2019 and relevance of states for firms' operation, up to year 2008, from Diego Garcia's webpage (https://sites.google.com/site/financieru/resources/software).

that of the firm being forecast, covered by prior co-analysts, as a proxy for quality of signals of these prior co-analysts.

We define other variables in Appendix A. All variables except count, time, and dummy variables are winsorized at 1% and 99% to minimize the impact of outliers and data errors.

4.3 Relocation, Entry, and Exit Samples

Our main tests are based on a sample which we call the *Combined Sample*. The sample is constructed by identifying firm-years in which a firm is affected by one of the following "events" induced by a triggering analyst: (i) an analyst following the firm changes his location (ii) a new analyst with no prior history in I/B/E/S starts covering a firm, or (iii) an analyst covering the firm ceases to have any record of coverage in I/B/E/S. Such relocation, entry, or exit events are not specific to any particular firm in the portfolio of firms covered by the triggering "event analyst", and result in different changes to the *OHHI* of the co-analysts of such an event analyst in the firms that he covers. As we discuss below, a typical analyst in a given year in the *Combined Sample* is exposed to multiple such events associated with firms he covers (the mean is 10 and the median is 6). There are also multiple events per firm-year (mean of 2.3 and median of 2). As a result, there is significant within-analyst-year as well as within-firm-year variation in the *change* in the *OHHI*. Our empirical specification examines the association between the change in *OHHI* (the independent variable of interest) and the change in the analyst's forecast error (*IEER*) from period *t-1* to period *t+1* for an analyst exposed to at least one event in period *t*. By exploiting within-analyst-year variation in changes or shocks to *OHHI* that are associated with events exogenous to a particular analyst, our approach therefore avoids standard endogeneity concerns that the choice of an analyst's location and hence the *OHHI* could be correlated with information quality or location characteristics (for example, an analyst located farther from the cluster of co-analysts would have high *OHHI* and may also be informationally disadvantaged, which could correlate with the analyst's forecast error). Moreover, the induced within-firm-year variation in *OHHI* also absorbs firm-year specific factors that could affect forecast error.

More specifically, the *Relocation Sample*, based on relocations of analysts from one MSA to another MSA, is constructed as follows. Suppose analyst *i* is recorded in MSA A in year *t* and in MSA B in year *t*+1. We identify those firms which analyst *i* covers in year *t*-1 as event firms. The relocation event induces heterogenous changes in the *OHHI* of analysts covering an event firm from year *t*-1 to year *t*+1. There are 1,278 relocations in our sample. We observe a fairly even distribution of relocations between 1997 and 2006, except 2003 in which approximately 25% of relocations take place. About 56% of relocations are associated with job turnovers when analysts move from one brokerage house to another brokerage house.

The *Entry and Exit Sample* consists of the entry and exit events of analysts either entering or leaving the I/B/E/S database. This sample is constructed as follows. For an entry, suppose analyst *i* first appears in the database in year *t* (event year). We identify those firms which analyst *i* covers in year *t* as event firms. For an exit, suppose year *s* is the last year in which analyst *i* appears in the database (event year). We identify those firms which analyst *i* covers in year *s*-1 as event firms. The entry and exit events induce heterogeneous changes in the *OHHI* of analysts covering an event firm from year *s*-1 to year *s*+1. There are 4,421 entry and exit events. The majority of them fairly spread between 1998 and 2008, except 2001 and 2002 in which there are more events, accounting for approximately 10% and 11% of these events, respectively. This pattern is consistent with that documented in Derrien and Kecskés (2013). The *Combined Sample* combines the *Relocation* and the *Entry and Exit* Samples.

4.4 Summary Statistics and Correlations

Table 1A reports the summary statistics of *OHHI* and Δ*OHHI* (changes in OHHI) based on the *Combined Sample*. The prior co-analysts' locations are fairly diversely distributed, with mean and median *OHHI* of 0.30 and 0.26. Noticeably, in relation to the average *OHHI*, the variation of *ΔOHHI* within firm-year and within analyst year is quite substantial. The within-firm-year mean and median of the range (i.e., maximum minus minimum) of *ΔOHHI* are 0.16 and 0.10, while the within-analyst-year mean and median of the range are 0.21 and 0.15. These represent substantial variation when compared with the mean (median) value of the *OHHI* of 0.295 (0.264). A typical sample analyst is exposed to multiple events in a given year, with a mean of 10 and median of 6. The mean and median of the corresponding statistics for an average firm-year are 2.3 and 2, respectively.

Table 1B shows other summary statistics. Forecast errors for the Combined Sample are generally smaller than those in the full sample, possibly because the analysts in the former sample have more experience. The median and mean error of the consensus forecasts are 0.29% and 1.16% in the full sample (0.21% and 0.88% in the Combined Sample), respectively. Meanwhile, the statistics of the absolute errors of the last forecasts of all analysts have a median of 0.21% and mean of 0.89% in the full sample (0.15% and 0.59% in the Combined Sample).[15] The reason why the statistics for individual forecasts are smaller than those for consensus forecasts is that there are a larger number of individual forecasts for firm years with larger coverage, and forecasts for firms with larger coverage tend to be more accurate. On the other hand, in both the full sample and the *Combined Sample*, the average difference between the mean of individual forecast absolute errors and the mean consensus absolute error for the same firm year ("mean(IERR) – ERR") is positive, consistent with Jensen's Inequality.

[15] Our forecast error statistics are comparable to summary statistics reported in the literature. For example, Clement et al. (2011) report a median of the mean consensus absolute errors of 0.2%. Walther and Willis (2013) show a mean of the individual forecast absolute errors of 0.95%.

Other statistics are also similar to those reported in the literature. The statistics of experience variables are comparable to those in Bradley, Gokkaya, and Liu (2017). The statistics of coverage variables are in line with Clarke, Khorana, and Patel (2007), Soltes (2014), and Guan, Wong, and Zhang (2015).

5. Results

5.1 Location HHI and Common Uncertainty

One of the basic premises of our analysis is that analysts' signals are less correlated if their locations are more geographically dispersed (Hypothesis H1(i)). This is plausible because if they are more dispersed, analysts are likely to collect their information from different sources, and signals are more likely to be idiosyncratic. Analysts also are more likely to share/exchange information or incorporate each other's views expressed in earnings conference calls if they are from the same region or familiar with each other, which could cause their signals to be more correlated.

In Table A1 in Appendix D, we show that geographical location of analysts matters. Specifically, analyst forecasts are sensitive to news about the state in which the analyst is located, to the extent that this state is relevant for the firm's operations (following Garcia and Norli (2012)). Results reported in the first two columns show that the total number of earnings forecasts an analyst makes for all the stocks he covers in a given week is positively related to the number of news stories related to the state in the previous week. Results reported in the last four columns show that an analyst's forecast is less optimistic relative to those of other analysts covering the same firm for the same fiscal year if the number of risk- or uncertainty-related news stories for his state, relative to the total number of such stories in all 50 states in the previous week, is higher. These results suggest that location matters for analysts' forecasting behavior, and so analysts from the same location are likely to have more correlated forecasts.

To test H1(i), we follow Barron et al. (1998) and, for each firm year, obtain an estimate of the ratio of the pairwise covariance of the analysts' forecast errors (C) scaled by the overall uncertainly (V - the average of each analyst's expected variance of earnings, conditional on his information).[16] Following Equation (14) in Barron et al. (1998), we estimate $\rho = \frac{C}{V}$, which the authors call *Consensus*, the ratio of common uncertainty to overall uncertainty. Since ρ is similar to a correlation coefficient, we take a Fisher transformation (Fisher, 1915, 1921) of ρ and regress the transformed measure, *RHO*, on *HHI*, with a number of firm-year level control variables, and firm and year fixed effects. The results are reported in the first four columns of Table 2. Our dependent variables are based on the last forecasts of all analysts in the firm's fiscal year. The regressions reported in the first two columns are in levels for the full sample, while in the third and fourth columns, all variables are in first difference. In Column (2), we also control for the Fisher-transformed *Consensus* based on the *first* forecast of each analyst in the firm's fiscal year, to control for other factors that could simultaneously affect both the *HHI* and the *Consensus*. The regression reported in the third column is for the full sample, while that in the fourth column is for the *Combined Sample*. In all four columns, the transformed ρ is positively and significantly related to the locational HHI (an inverse measure of analyst locational dispersion), consistent with H1(i). The economic magnitude of this relationship is also meaningful. For example, the coefficient of *HHI* in Column (1) suggests that one standard deviation reduction in *HHI* (0.234 in Table 1B) is associated with a decrease of 0.066 in *RHO*, which is around 10% of the mean of *RHO* (shown in Table 1B).

In the last four columns of Table 2, we report results of a related exercise. Here, we directly test whether location dispersion affects the mean absolute error of the consensus forecast (Hypothesis H1(ii)). Since the absolute error does not take negative values, we can

[16] In Appendix B, we show that under plausible conditions, greater diversity translates to lower error covariance in our model setting even in the presence of learning.

mainly focus on a logarithmic transformation of the absolute error (*ERR*); however, we also report results for regressions in first difference. Control variables are the same as in the first four columns. Columns (5) and (6) report the results in levels, while for the results reported in Columns (7) and (8), all variables are in first difference. Results reported in Column (6) also control for the consensus error based on the first forecasts of all analysts in the same firm fiscal year. Consistent with Hypothesis H1(ii), we find that the *HHI* is positively and significantly related to the consensus error.

5.2 Diversity and Individual Analyst Forecast Error

Having shown the relationship between firm-level location diversity measure and (i) the pairwise correlation between individual analyst errors, as well as (ii) the mean consensus forecast error, we next turn to our main query: whether analysts take the prior forecasts of other analysts into account when making their own forecasts. As discussed in Section 2, there could be two types of reasons why the prior consensus forecast influences an analyst's forecast. First, there could be career concern or reputation-related reasons for paying attention to and not deviating too much from the consensus forecast of prior analysts; second, if there is information content in the forecasts of prior analysts, there could be information-related reasons for taking cognizance of these prior forecasts. However, let alone establishing possible motives, even establishing that analysts influence each other is challenging (Cohn and Juergens, 2014).

The discussion leading to Hypothesis H2 says that if an analyst assigns some positive weight to the mean prior forecast of the co-analysts (equivalently, if his forecast is a weighted average of his own signal and the prior forecasts of co-analysts), then his forecast error will be affected by the standard deviation of the prior forecast noises of these co-analysts. If greater location diversity of the co-analysts is associated with lower covariance between their forecast errors, then location diversity of co-analysts will be negatively related to an analyst's mean absolute forecast error or *IEER*. Recalling that the *OHHI* is an inverse measure of locational diversity of the co-analysts, the forecast error will thus be positively related to *OHHI*. To test this hypothesis, we regress Δ*IEER*, the change in the absolute error of the last forecast of an individual analyst, on the corresponding Δ*OHHI* with analyst×year and year×firm fixed effects. By incorporating analyst-year fixed effects, we ensure that our results are not driven by a clustering of analysts with particular traits (relevant for forecast accuracy) into certain geographical regions or due to a particular analyst's career concerns.[17] Moreover, incorporating firm-year fixed effects further absorbs factors related to a firm's information environment that could affect forecast errors.

Table 3 presents the results. The first column presents results in levels for *the Combined Sample*, while Columns (2)-(7) present results in difference for the Relocation Sample, the Entry/Exit Sample and the Combined Sample. The variable of interest is OHHI in Column (1) and Δ*OHHI* in Columns (2)-(7). Other control variables (in differences from year *t*-1 to year *t*+1 in Columns (2)-(7)) are drawn from the literature. These comprise the time between the release date of analyst's forecast and the data date of the earnings being forecast, the monthly cumulative stock returns from the data date of the last annual earnings to the date of the analyst's forecast, the variance of the monthly stock returns for the past 12 months preceding the date of the analyst's forecast, an indicator of a negative forecast, and the time since the analyst's first forecast for the firm in question. For the results in difference, we also report a specification in which we include the change in the average distance of the locations of the co-analysts from an analyst's location (Δ*LN*(*CODIS*)), and interact this variable with Δ*OHHI*. This is done to control for changes in the correlation between an analyst's own forecast error and

[17] For example, many brokerages are headquartered in the state of New York. This implies that any random firm is likely to have a higher proportion of analysts based in New York covering it, and so these New York-based analysts will have lower *OHHI* compared with analysts based elsewhere. If New York-based analysts are more accurate in their forecasts, we would find a positive association between *OHHI* and *IERR*.

those of the co-analysts due to the relocation, entry or exit events, which can affect the analyst's forecast error as long as some weight is assigned to the mean forecasts of the co-analysts.

In all regressions, standard errors are clustered at the analyst×year level. The coefficient of *OHHI* and Δ*OHHI* in Columns (1) and (4), respectively, for the *Combined Sample* are significant at the 1% level. The coefficient of Δ*OHHI* remains similarly positive and significant for the relocation and entry/exit subsamples. The other control variables have signs and significance that are generally consistent with those reported in the literature. The change in average distance from co-analysts has a significant negative sign, which is consistent with the possibility that the average correlation between the noise in the analyst's signal and those of the co-analysts lowers the analysts forecast error (*IEER*).[18]

5.3 Observational Learning

We next test several implications that are unique to the learning hypothesis, as mentioned in Hypothesis H3. First, we test Hypothesis H3(i), which posits that an analyst is likely to assign a higher weight to co-analysts' prior forecasts when the information quality of the signals of these latter analysts is better. We construct two measures of information quality based on the average geographical distance of these analysts from peer firms, in the same 3-digit SIC industry as the firm under consideration, that they cover, and another measure based on the number of such peer firms that they cover. The measures of geographic distance are based both on the distance from the peer firm's headquarter ("BUS DIS"), as well as the weighted average distance of the analyst's location from all 50 states, with the weight being the relevance of the state for the peer firm's business ("GEO DIS"), as described in section 4.2.3. A larger average distance of the co-analysts suggests poorer information quality of the co-analysts, and thus, as predicted by Hypothesis 3(i), a smaller weight on the *OHHI*.

[18] This can be readily seen from Equation (3.2) and Equation (A3) in Appendix B.

Consistently, in the first four columns of Table 4, we find that the interaction of Δ*OHHI* and the average distance of the co-analysts from the peer firms they cover (*X**Δ*OHHI*) is negative and significant (at the 5 percent level in Columns (1), (3), and (4), and at the 10 percent level in Column (2)). Next, we argue that if analysts cover more similar firms, their forecasts as a group will be more informative, and hence the analyst will assign a higher weight to their forecasts, resulting in a higher sensitivity of individual forecast errors to *OHHI*. In Columns (5) and (6), we find that the interaction of *ΔOHHI* and the total number of peer firms covered by all the prior co-analysts (in hundreds) are positive and significant at the 1 percent level.

In Table 5, we turn our attention to the information quality of an individual analyst. Information quality is measured in two ways: the distance of the analyst from the firm being covered, and the average distance of the analyst from all same 3-digit SIC peer firms that he also covers. When these distance measures are higher, we posit that the analyst's information quality is poorer. Hypothesis 3(ii) suggests that the interactions of these information quality measures with Δ*OHHI* should be positive, since the analyst would assign more weight to the forecasts of co-analysts when his own information quality is poorer. This is what we find – the interactions of the distance measures with Δ*OHHI* are positive and significant at the 1 percent level in all eight columns of Table 5.[19]

6. Conclusion

Using US data for the 1994-2010 period, we document that when analysts following the same firm are more geographically diverse, their forecasts are less correlated and their consensus forecasts are more accurate. We further show that when the other analysts covering the same firm are more geographically spread, an analyst's individual forecast is also more

[19] While the coefficient of Δ*OHHI* itself is negative, it can be verified that the net effect of Δ*OHHI* (the sum of the coefficients of the first two terms in each column) becomes positive at slightly above 100 kilometres and reaches the estimate from Table 3 at around 1,400 kilometres (see Appendix E).

accurate. These results are consistent with the notion that higher analyst diversity is associated with more complete "cancelling out" of idiosyncratic components of their individual forecasts, whereby both consensus and individual forecasts have smaller errors. This is similar to the diversification effect in the portfolio theory.

The results suggest that analysts assign positive weights to the mean forecasts of other analysts, which is consistent with both uninformed herding and observational learning, whereby an analyst's forecast is a weighted average of the mean prior forecast of other analysts and his own signal. We find strong evidence for the latter type of behavior. We document that when an analyst is more distant from sources of information, the analyst puts more weight on the forecasts of co-analysts; similarly, when the co-analysts are closer to sources of information or cover more stocks from the same industry as a particular firm, an analyst assigns more weight to their forecasts in making his own forecast. Overall, our results suggest that analysts incorporate peer analyst forecasts into their own forecasts to generate more informationally efficient forecasts.

References

Angrist, J. (2014). The perils of peer effects. *Labor Economics*, 30, 98-108.

Bae, K. H., Stulz, R. M., & Tan, H. (2008). Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. *Journal of Financial Economics*, 88(3), 581-606.

Banerjee, A. V. (1992). A simple model of herd behavior, *The Quarterly Journal of Economics*, 107(3), 797-817.

Barron, O. E., Kim, O., Lim, S. C., & Stevens, D. E. (1998). Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review*, 73(4), 421-433.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics*, 80(3), 657-675.

Bikhchandani, S. Hirshleifer, D. & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of Political Economy*, 100(5), 992-1026.

Bloomfield, R., & Hales, J. (2009). An experimental investigation of the positive and negative effects of mutual observation. *The Accounting Review*, 84(2), 331-354.

Bradley, D., Gokkaya, S., & Liu, X. (2017). Before an analyst becomes an analyst: Does industry experience matter? *Journal of Finance*, 72(2), 751-792.

Chen, B. (2017). Seeing is Believing: The Impact of Local Economic Conditions on Firm Expectations, Employment and Investment. Working paper, Department of Economics, Harvard University.

Chen, Q., & Jiang, W. (2006). Analysts' weighting of private and public information. *Review of Financial Studies*, 19(1), 319–355.

Chen, Q., Mayew, W., and Yan, H. (2018). Do Social Interactions Communicate or Garble Information? Evidence from Equity Analysts, Working paper, Fuqua School of Business, Duke University.

Christensen, T. E., (2007). Discussion of "letting the `tail wag the dog': the debate over GAAP versus street earnings revisited". *Contemporary Accounting Research*, 24(3), 741-762.

Clarke, J., Khorana, A., Patel, A., & Rau, P. R. (2007). The impact of all-star analyst job changes on their coverage choices and investment banking deal flow. *Journal of Financial Economics*, 84(3), 713-737.

Clement, M. B., Hales, J., & Xue, Y. (2011). Understanding analysts' use of stock returns and other analysts' revisions when forecasting earnings. *Journal of Accounting and Economics*, 51(3), 279-299.

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

Cohn, J. B., & Juergens, J. L. (2014). How much do analysts influence each other's forecasts? *Quarterly Journal of Finance*, 4(3), 1-35.

Coval, J.D., Moskowitz, T.J. (2001).The geography of investment: informed trading and asset prices. *Journal of Political Economy*, 109, 811–841.

Derrien, F., & Kecskés, A. (2013). The real effects of financial shocks: Evidence from exogenous changes in analyst coverage. *Journal of Finance*, 68(4), 1407-1440.

Einhorn, H., Hogarth R., & Klempner, E. (1977). Quality of group judgment. *Psychological Bulletin*, 84(1), 158-172.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika*, 10(4), 507–521.

Fisher, R. A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3–32.
Garcia, D., & Norli, Ø. (2012). Geographic dispersion and stock returns. *Journal of Financial Economics*, 106(3), 547-565.
Gerken, W., and Painter, M. (2019). Are Crowded Crowds Still Wise? Evidence from Financial Analysts' Geographic Diversity, Working paper, University of Kentucky.
Guan, Y., Wong, M. F., & Zhang, Y. (2015). Analyst following along the supply chain. *Review of Accounting Studies*, 20(1), 210-241.
Guttman, I. (2010). The timing of analysts' earnings forecasts. *The Accounting Review*, 85(2), 513-545.
Herfindahl, O. C. (1950). *Concentration in the Steel Industry*. Doctoral dissertation, Columbia University.
Hirschman, A. O. (1964). The paternity of an index. *American Economic Review*, 54(5), 761-762.
Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1-3), 337-386.
Hong, H., Kubik, J. D., & Solomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics*, 31(1), 121-144.
Huang, R., Krishnan, M., Shon, J., & Zhou, P. (2017). Who herds? Who doesn't? Estimates of analysts' herding propensity in forecasting earnings. *Contemporary Accounting Research*, 34(1), 374-399.
Ivkovic´, Z., Weisbenner, S. (2005). Local does as local is: information content of the geography of individual investors' common stock investments. *Journal of Finance*, 60, 267–306.
Jennings, J., Lee, J., & Matsumoto, D. (2017). The effect of industry co-location on analysts' information acquisition costs. *The Accounting Review*, 92(6), 103-127.
Kang, J.K., Stice-Lawrence, L., & Wong, F. (2019). The Firm Next Door: Using Satellite Data to Tease Out Information Acquisition Costs, Working paper, Marshall School of Business, University of Southern California.
Keskek, S., Tse, S., & Tucker, J. (2014). Analyst information production and the timing of annual earnings forecasts. *Review of Accounting Studies*, 19(4),1504–1531.
Kim, O., Lim, S., & Shaw, K. (1998). The inefficiency of the mean analyst forecast as a summary forecast of earnings. *Journal of Accounting Research*, 39(2), 329-335.
Kuppuswamy, V. & Bayus, B. L. (2018). Crowdfunding creative ideas: The dynamics of project backers in kicksarter, Working Paper, University of North Carolina.
Kumar, A., Rantala, V., and Xu, R., (2021). Social learning and analyst behaviour. Forthcoming, *Journal of Financial Economics*.
Malloy, C. J. (2005). The geography of equity analysis. *Journal of Finance*, 60(2), 719-755.
Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60 (3), 531–542.
Merkley, K., Michaely, R., & Pacelli, J. (2017). Does the scope of the sell‐side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports. *The Journal of Finance*, 72(3), 1285-1334.
Merkley, K. J., Michaely, R., & Pacelli, J. (2020). Cultural diversity on Wall Street: Evidence from sell-side analysts' forecasts. *Journal of Accounting and Economics*, Forthcoming.
O'Brien, P. C., & Tan, H. (2015). Geographic proximity and analyst coverage decisions: Evidence from IPOs. *Journal of Accounting and Economics*, *59*(1), 41-59.
Phua, K., Tham, M., and Wei, C. (2021), Peer effects in equity research. Forthcoming, *Journal of Financial and Quantitative Analysis*.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *American Economic Review*, 80(3), 465-479.

Soltes, E. (2014). Private interaction between firm management and sell-side analysts. *Journal of Accounting Research*, 52(1), 245-272.

Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business. *Economies, Societies and Nations*, New York, NY, US: Doubleday & Co.

Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies*, 7(1), 97-124.

Van Nieuwerburgh, S.,Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64, 1187–1215.

Walther, B. R., & Willis, R. H. (2013). Do investor expectations affect sell-side analysts' forecast bias and forecast accuracy? *Review of Accounting Studies*, 18(1), 207-227.

Zhang, J. (2010). The sound of silence: Observational learning in the U.S. kidney market. *Marketing Science*, 29(2), 315-335.

Zhang, J. & Liu, P. (2012), Rational herding in microloan markets. *Management Science*, 58(5), 892-912.

Table 1A. OHHI (Prior Co-Analysts' Location Diversity) and Event Exposure Summary Statistics

The table reports summary statistics of OHHI, ∆OHHI (change in OHHI) and the number of events (i.e., relocations, entries and exits of analysts), based on the *Combined Sample*, consisting of those analyst-firm-years affected by these events. The event analysts are excluded. *OHHI* is the Herfindahl-Hirschman Index of the geographical locations, based on the U.S. Metropolitan Statistical Areas (MSAs) and cities that are not in a U.S. MSA., of all prior co-analysts who have annual EPS forecasts published before the publication of analyst *i*'s last annual EPS forecast in the same firm fiscal year. To be included in the sample, we require that the data for constructing all the variables used for an observation are available. When there are multiple forecasts with multiple activations on the same date in the I/B/E/S database, we take the forecast of the latest activation date.

	Mean	Median	S.D.
OHHI	0.295	0.264	0.151
∆OHHI	0.003	0.000	0.150
∆OHHI > 0	0.103	0.067	0.118
∆OHHI < 0	-0.098	-0.068	0.105
Range of ∆OHHI within firm year	0.161	0.103	0.189
S.D. of ∆OHHI within firm year	0.069	0.042	0.080
Range of ∆OHHI within analyst year	0.205	0.154	0.220
S.D. of ∆OHHI within analyst year	0.106	0.083	0.087
Number of relocations, entries and exits per firm year	2.309	2.000	1.982
Number of relocations, entries and exits per analyst year	9.892	6.000	10.941

Table 1B. Other Summary Statistics

The table includes summary statistics of the main dependent and independent variables we use in this paper. Appendix A provides the definitions of the variables. The sample period is from 1994 to 2010. To be included in the sample, we require that the data for constructing all the variables used for an observation are available. When there are multiple forecasts with multiple activations on the same date in the I/B/E/S database, we take the forecast of the latest activation date. The *Combined Sample*, consists of those analyst-firm-years affected by Relocation, Entry, or Exit. * indicates a firm-level variable. All other variables (without *) are individual-level variables.

	Full Sample			Combined Sample		
VARIABLES	Mean	Median	S.D.	Mean	Median	S.D.
Panel A: Main Variables						
*ERR (consensus)	1.163	0.291	2.895	0.898	0.216	2.571
IERR (individual)	0.889	0.210	2.593	0.590	0.154	1.464
*Jensen's Inequality: $[\text{mean}(IERR)_{j,t} - ERR_{j,t}] \geq 0$	0.166	0.009	0.773	0.152	0.018	0.631
*RHO	0.659	0.390	0.917	0.530	0.323	0.713
*HHI	0.453	0.388	0.234	0.382	0.344	0.175
Panel B: Distance Variables						
BUS DIS, in km	1698	1553	865	1650	1193	1441
GEO DIS, in km	1615	2637	763	1603	1535	1005
BUS IDIS, in km	1718	1580	896	1697	1547	1064
GEO IDIS, in km	1624	1641	770	1615	1631	832
CODIS, in km	1155	896	1014	1141	893	947
Panel C: Control Variables in Tables 2-5						
FIRM_EXP, in years	3.230	1.996	3.592	5.178	4.115	3.811
LN(HORIZON)	4.173	4.205	0.866	4.184	4.234	0.877
RETTODATE	0.021	0.040	0.353	0.023	0.042	0.351
SIGMA	0.111	0.095	0.063	0.108	0.094	0.059
LOSS	0.089	0.000	0.285	0.072	0.000	0.258
Panel D: Additional Control Variables in Table 2						
*COVERAGE	9	7	9	9	7	9
NFIRM	18	16	12	18	17	10
NSIC3	6	5	4	6	5	4
FORECAST_EXP, in years	7.167	5.911	5.573	9.067	7.984	5.408
*LN(SIZE)	7.760	7.700	1.737	7.584	7.457	1.649
*LN(BM)	-0.955	-0.885	0.682	-0.899	-0.837	0.680
*RET	0.014	0.013	0.036	0.012	0.011	0.036
*PROFIT	0.129	0.144	0.220	0.121	0.137	0.208
*VOL_ROE	0.076	0.007	0.321	0.059	0.007	0.220
*EPS_SKEW	0.000	0.000	0.016	0.000	0.001	0.018
SP500	0.401	0.000	0.490	0.566	1.000	0.496

Table 2. Analyst Location Diversity (HHI), Ratio of Average Covariance to Total Uncertainty (RHO) and Absolute Error of Mean Consensus Forecasts (LN(ERR))

The regression sample in Columns (1)-(3) and (5)-(7) is the full sample, while that in Columns (4) & (8) is the *Combined Sample*. In Columns (1) & (2), the dependent variable is *RHO* based on a Fisher transformation (Fisher 1915, 1921). *RHO* measures the degree to which analysts' beliefs covary relative to the overall level of uncertainty in a firm year. The pre-transformed *RHO* is estimated following Barron et al. (1998) as $(SE - D/N)/[(1 - 1/N)\ D + SE]$, where *SE* is the squared difference between the mean of the last annual EPS forecasts of all analysts, who make the forecasts for firm *j* in year *t*, and the actual EPS being forecasted (both scaled by the market share price of the firm as of the end of the previous fiscal year); *D* is the variance of the individual scaled forecast of the analysts; *N* is the number of the analysts (Equation (16) in Barron et al. 1998). In Columns (5) & (6), the dependent variable is the natural logarithm of 1 plus the absolute error of the mean of the last forecasts of all analysts ("mean consensus"), *LN(ERR)*, for firm *j* in fiscal year *t.* In Columns (3), (4), (7) & (8), all variables are in first difference. *HHI* is the Herfindahl-Hirschman Index of all analysts, who make the forecasts for firm *j* in fiscal year *t*, for their geographical locations, based on the U.S. Metropolitan Statistical Areas (MSAs) and cities where the city is not in a U.S. MSA. *COVERAGE* is the number of the analysts in a firm fiscal year. *FIRST ERR* [*FIRST RHO*] is the same as *ERR* [*RHO*], except that it is based on the first forecasts of the analysts. For brevity, the results of other control variables are not reported. Estimated coefficients and the robust standard errors (in parentheses) are reported. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

Sample	Full	Full	Full	Combined	Full	Full	Full	Combined
	Level	Level	First Difference	First Difference	Level	Level	First Difference	First Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RHO	RHO	RHO	RHO	LN(ERR)	LN(ERR)	LN(ERR)	LN(ERR)
HHI	0.282***	0.191***	0.251***	0.229**	0.127***	0.061***	0.080**	0.092*
	(0.044)	(0.036)	(0.064)	(0.107)	(0.023)	(0.012)	(0.029)	(0.053)
LN(FIRST ERR)						0.545***		
						(0.019)		
FIRST RHO		0.337***						
		(0.017)						
COVERAGE	-0.018***	-0.015***	-0.026***	-0.018***	0.003***	-0.001	0.004**	0.004*
	(0.002)	(0.002)	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
Year & Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year & Firm S.E. Cluster	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,722	34,714	28,505	11,919	33,773	33,773	27,865	11,962
Within R^2	0.018	0.131	0.014	0.017	0.215	0.583	0.100	0.100

Table 3. Prior Co-Analysts' Location Diversity (OHHI) and Absolute Error of Individual Forecasts (IEER)

The sample is *Combined Sample*. The relocating, entering, and exiting analysts are excluded. The dependent variable is based on the absolute error of analyst *i*'s last annual EPS forecast for firm *j* in year *t* (*IERR*). In Column (1), the dependent variable is the natural logarithm of 1 plus *IERR*. In Columns (2) - (7), the dependent variable is the difference in *IERR*. *OHHI* is the Herfindahl-Hirschman Index of the geographical locations, based on the U.S. Metropolitan Statistical Areas (MSAs) and cities that are not in a U.S. MSA., of all prior co-analysts (i.e., those other analysts who have annual EPS forecasts published before the publication of analyst *i*'s last annual EPS forecast in the same firm fiscal year). *LN*(*CODIS*) is the natural logarithm of 1 plus the average distance (in km) of analyst *i*'s city from the cities of the prior co-analysts. The other variables are defined in Appendix A. Explanatory variables are at the level (indicated in in square brackets) in Column (1) and in difference (Δ) in Columns (2) – (7). Estimated coefficients and the robust standard errors (in parentheses) are reported. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

Sample:	Combined	Combined	Combined	Relocation	Relocation	Entry & Exit	Entry & Exit
Regression:	Level	Difference	Difference	Difference	Difference	Difference	Difference
	[LN(IERR)]	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
[OHHI] ΔOHHI	0.187***	0.556***	0.535***	0.545***	0.557***	0.730***	0.698***
	(0.016)	(0.047)	(0.049)	(0.101)	(0.107)	(0.075)	(0.079)
ΔOHHI*ΔLN(CODIS)			-0.015		0.002		0.007
			(0.021)		(0.043)		(0.027)
ΔLN(CODIS)			-0.013**		-0.028**		-0.024***
			(0.005)		(0.012)		(0.008)
[LN(FIRM_EXP)] ΔLN(FIRM_EXP)	0.011***	-0.004*	-0.004	0.004	0.004	0.001	0.002
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
[LN(HORIZON)] ΔLN(HORIZON)	0.086***	0.130***	0.131***	0.112***	0.115***	0.134***	0.133***
	(0.002)	(0.004)	(0.004)	(0.006)	(0.007)	(0.004)	(0.005)
[RETTODATE] ΔRETTODATE	0.100***	0.205***	0.198***	0.140***	0.136***	0.257***	0.246***
	(0.005)	(0.013)	(0.014)	(0.035)	(0.036)	(0.022)	(0.023)
[SIGMA] ΔSIGMA	0.674***	1.619***	1.567***	2.067***	2.208***	1.623***	1.472***
	(0.054)	(0.153)	(0.157)	(0.362)	(0.387)	(0.255)	(0.241)
[LOSS] ΔLOSS	0.101***	0.198***	0.168***	-0.001	0.013	0.143*	0.092
	(0.015)	(0.044)	(0.045)	(0.103)	(0.106)	(0.080)	(0.083)
CONSTANT	-0.270***	0.228***	0.224***	0.183***	0.186***	0.242***	0.237***
	(0.019)	(0.002)	(0.003)	(0.007)	(0.007)	(0.003)	(0.004)
Year×Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst S.E. Clusters	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	124,854	127,484	113,813	31,578	28,134	92,719	82,771
Adjusted R^2	0.672	0.600	0.600	0.620	0.621	0.656	0.659

Table 4. OHHI and IEER: Signal Quality of Prior Co-Analysts

The sample is the *Combined Sample*. The dependent variable (Δ*IERR*) is the change (from year *t-1* to year *t+1*) in the absolute error of the last forecast of analyst *i* covering firm *j*, where firm *j* is a firm affected by a relocation, entry or exit event in year *t*. The relocating, entering, and exiting analysts are excluded. Δ indicates a variable in difference (year *t+1* value minus year *t-1* value). *OHHI* is the Herfindahl-Hirschman Index of the geographical locations, based on the U.S. Metropolitan Statistical Areas (MSAs) and cities that are not in a U.S. MSA., of all prior co-analysts (i.e., those generating their forecasts for firm *j* in the same fiscal year prior to firm *i*'s last forecast). In Columns (1)-(4), the independent variable *X* is the average distance of all prior co-analysts of analyst *i* from all peer firms of firm *j* that they cover. Peer firms are other firms in the same 3-digit SIC industry as firm *j*. In Columns (1)-(2), we first compute the individual average distance (in km) between each prior co-analyst and the business location of all peer firms covered by that analyst. We then take the average of the individual average distance of all the prior co-analysts. *X* ["BUS IDIS"] is the logarithm of 1 plus the average of the individual average distance. In Columns (3)-(4), each co-analyst's distance from a peer firm of firm *j* that he covers is a weighted average distance from all 50 states, for which the weight is the degree of relevance of the state for the business of each peer firm covered by that analyst. The remaining steps for computing *X* ["GEO IDIS"] are the same as those for calculating *X*["BUS IDIS"] in Columns (1)-(2). In Columns (5)-(6), *X* ["ICOVERAGE"] is the total number of other peer firms (in hundreds) of firm *j* in the same SIC3 industry, covered by all the prior co-analysts of analyst *i*. *LN*(*CODIS*) is the natural logarithm of 1 plus the average distance (km) of analyst *i*'s city from the cities of the prior co-analysts. Estimated coefficients and the robust standard errors (in parentheses) are reported. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

X:	Peer firms' BUS IDIS	Peer firms' BUS IDIS	Peer firms' GEO IDIS	Peer firms' GEO IDIS	Peer firms' ICOVERAGE	Peer firms' ICOVERAGE
	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR
	(1)	(2)	(3)	(4)	(5)	(6)
X*ΔOHHI	-0.110**	-0.102*	-0.121**	-0.106**	0.159***	0.151***
	(0.053)	(0.055)	(0.051)	(0.054)	(0.047)	(0.049)
ΔOHHI	1.186***	1.079***	1.226***	1.086***	0.358***	0.358***
	(0.355)	(0.373)	(0.345)	(0.361)	(0.060)	(0.063)
X	-0.059***	-0.061***	-0.059***	-0.059***	-0.239***	-0.233***
	(0.009)	(0.010)	(0.009)	(0.009)	(0.016)	(0.017)
ΔOHHI*ΔLN(CODIS)		0.064		0.029		-0.010
		(0.120)		(0.110)		(0.027)
ΔOHHI*ΔLN(CODIS)*X		-0.030		-0.023		0.023
		(0.025)		(0.023)		(0.042)
ΔLN(CODIS)*X		-0.002		0.003		-0.010
		(0.006)		(0.005)		(0.007)
ΔLN(CODIS)		-0.014		-0.039		-0.004
		(0.039)		(0.033)		(0.007)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
X*Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year×Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst S.E. Clusters	Yes	Yes	Yes	Yes	Yes	Yes
Observation	88,386	87,612	78,493	77,754	127,484	113,813
Adjusted R^2	0.653	0.652	0.615	0.615	0.602	0.602

Table 5. OHHI and IEER: Signal Quality of Analyst

The sample is the *Combined Sample*. The dependent variable (Δ*IERR*) is the change (from year *t-1* to year *t+1*) in the absolute error of the last forecast of analyst *i* covering firm *j*, where firm *j* is a firm affected by a relocation, entry or exit event in year *t*. The relocating, entering, and exiting analysts are excluded. Δ indicates a variable in difference (year *t+1* value minus year *t-1* value). *OHHI* is the Herfindahl-Hirschman Index of the geographical locations, based on the U.S. Metropolitan Statistical Areas (MSAs) and cities that are not in a U.S. MSA., of all prior co-analysts (i.e., those generating their forecasts for firm *j* in the same fiscal year prior to firm *i*'s last forecast). *LN*(*D*) is the logarithm of 1 plus the distance (in km) between the analyst's city and the firm's business location ["BUS DIS"] in Columns (1) & (2), and the weighted average of the distance between the analyst's city and 50 states ["GEO DIS"] in Columns (3) & (4), for which the weight is the degree of state's relevance for the firm's business. In Columns (5)-(6) and (7)-(8), we replace the location of the firm being forecast in Columns (1)-(2) and (3)-(4) by the locations of the other peer firms, in the same 3-digit SIC industry as that of the firm under consideration, covered by the analyst and compute the average distance of the analyst's city from these peer firms. *LN*(*CODIS*) is the natural logarithm of 1 plus the average distance (km) of analyst *i*'s city from the cities of the prior co-analysts. Estimated coefficients and the robust standard errors (in parentheses) are reported. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

Distance of analyst's city from:	Firm's BUS DIS	Firm's BUS DIS	Firm's GEO DIS	Firm's GEO DIS	Peer Firms' BUS IDIS	Peer Firms' BUS IDIS	Peer Firms' GEO IDIS	Peer Firms' GEO IDIS
	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR	ΔIERR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LN(D)*ΔOHHI	0.221***	0.191***	0.363***	0.503***	0.196***	0.162***	0.268***	0.395***
	(0.062)	(0.062)	(0.092)	(0.106)	(0.056)	(0.056)	(0.077)	(0.099)
ΔOHHI	-1.031**	-0.825*	-2.118***	-3.144***	-0.860**	-0.626	-1.421**	-2.362***
	(0.441)	(0.441)	(0.665)	(0.768)	(0.396)	(0.398)	(0.557)	(0.717)
LN(D)	0.053***	0.052***	0.034	0.044	0.039**	0.040**	-0.008	-0.007
	(0.019)	(0.019)	(0.021)	(0.029)	(0.016)	(0.016)	(0.021)	(0.026)
ΔOHHI*ΔLN(CODIS)		0.001		0.026		-0.029		0.214
		(0.279)		(0.453)		(0.227)		(0.405)
ΔOHHI*ΔLN(CODIS)*LN(D)		-0.001		-0.004		0.003		-0.030
		(0.040)		(0.063)		(0.033)		(0.056)
ΔLN(CODIS)*LN(D)		-0.031***		-0.036**		-0.027***		-0.014
		(0.010)		(0.015)		(0.008)		(0.015)
ΔLN(CODIS)		0.211***		0.261**		0.185***		0.096
		(0.069)		(0.105)		(0.056)		(0.110)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LN(D)*Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Analyst S.E. Clusters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	107,439	106,874	89,008	88,016	105,487	104,936	86,119	85,184
Adjusted R^2	0.608	0.607	0.545	0.544	0.609	0.608	0.544	0.542

Appendix A. Variable Definitions

This appendix provides the variable definitions in alphabetic order. Subscripts *i*, *j*, and *t* (*s*) index analysts, firms, time, respectively. * indicates a control variable without tabulated results.

BIAS RELATIVE TO MEAN(MEDIAN) CONSENSUS$_{i,j,t}$: the difference between analyst *i*'s last annual EPS forecast and the mean(median) of last annual EPS forecasts of all analysts who make EPS forecast for firm *j* in year *t*, scaled by the share price of the firm as of the end of the previous fiscal year. (Data Sources: I/B/E/S and CRSP)

BUS DIS (DISTANCE FROM FIRM'S BUSINESS LOCATION)$_{i,j,t}$: the distance between analyst *i*'s city and the business location of firm *j*. (Data Source: Nelson's Directories and Bill McDonald's website https://www3.nd.edu/~mcdonald/)

BUS IDIS (DISTANCE FROM PEER FIRMS' BUSINESS LOCATION)$_{i,j,t}$: the distance between analyst *i*'s city and the average business location of other peer firms covered by the analyst in the same 3-digit SIC industry as firm *j* being forecast. (Data Source: Nelson's Directories and Bill McDonald's website https://www3.nd.edu/~mcdonald/)

CODIS$_{i,j,t}$: the average distance of analyst *i*'s city from the cities in which all prior co-analysts are located. Prior co-analysts are those other analysts who have annual EPS forecasts published before publication of analyst *i*'s last annual EPS forecast in the same firm fiscal year (firm *j* year *t*).

COVERAGE$_{j,t}$: the number of analysts covering firm *j* for fiscal year *t*. (Data Source: I/B/E/S)

ERR$_{j,t}$: $= \frac{|F-A|}{P} \times 100$, where *F* is the last mean annual EPS (earnings per share) forecast of all analysts for firm *j* in year *t*. *A* is the actual value of the EPS being forecast. *P* is the market share price of firm *j* as of the end of the last fiscal year. (Data Sources: I/B/E/S and CRSP)

**EPS_SKEW*$_{j,t}$: the difference between the mean and median of EPS, scaled by the market share price for the end of fiscal year *t*-1 for firm *j* for the period between fiscal year *t*-4 and fiscal year *t* + 4, excluding fiscal year *t*. (Data Source: I/B/E/S)

FIRST ERR$_{j,t}$: $= \frac{|F-A|}{P} \times 100$, where *F* is the first mean annual EPS (earnings per share) forecast of all analysts for firm *j* in year *t*. *A* is the actual value of the EPS being forecast. *P* is the market share price of firm *j* as of the end of the last fiscal year. (Data Sources: I/B/E/S and CRSP)

FIRST RHO$_{j,t}$: $= \frac{1}{2} ln\left(\frac{1+\rho}{1-\rho}\right)$, where $\rho = \frac{SE-\frac{D}{N}}{\left(1-\frac{1}{N}\right)D+SE}$ where *SE* is the squared difference between the mean of the first annual EPS forecasts of all analysts, who produce forecasts for firm *j* in year *t*, and the actual value of EPS being forecasted, scaled by the market share price of the firm as of the end of the previous fiscal year. *D* is the variance of the individual scaled forecasts of the analysts. *N* is the number of the analysts. (Data Sources: I/B/E/S and CRSP)

FNUM$_{i,t}$: analyst *i*'s weekly number of annual EPS forecasts. (Source: I/B/E/S)

GEO DIS (DISTANCE FROM FIRM'S WEIGHTED LOCATIONS)$_{i,j,t}$: the distance between analyst *i*'s city and the weighted average of relevant states of firm *j*, where the weight is the degree of relevance of the state for the firm. (Data source: Nelson's Directories and Diego Garcia's webpage (https://sites.google.com/site/financieru/resources/software))

GEO IDIS (DISTANCE FROM PEER FIRMS' WEIGHTED LOCATIONS)$_{i,j,t}$: the distance between analyst *i*'s city and the average of the weighted average of relevant states of other peer firms covered by the analyst in the same 3-digit SIC industry as firm *j*, where the weight is the degree of relevance of the state for the peer firm. (Data source: Nelson's Directories and Diego Garcia's webpage (https://sites.google.com/site/financieru/resources/software))

$HHI_{j,t} := \sum_{all\ q} \left(\frac{Number\ of\ analysts\ in\ location\ q_{j,t}}{Total\ number\ of\ analysts_{j,t}}\right)^2$, where $Total\ number\ of\ analysts_{j,t}$ is the total number of analysts producing annual EPS forecasts for firm *j* in year *t*. *Number of analysts in location* $q_{j,t}$ is the number of analysts producing the annual EPS forecasts in the q^{th} location where *q* is one of the U.S. Metropolitan Statistical Areas (MSAs) or a city which is not in a U.S. MSA. (Data Sources: I/B/E/S, Nelson's Directories)

ICOVERAGE$_{i,j,t}$: the total number of other peer firms (in hundred), in the same 3-digit SIC industry as that of firm *j*, covered by all prior co-analysts who have annual EPS forecasts published before publication of analyst *i*'s last annual EPS forecast in the same firm fiscal year (firm *j* year *t*). (Data Source: I/B/E/S)

IERR$_{i,j,t}$: the absolute error of individual forecasts, estimated as the absolute difference between analyst *i*'s last annual EPS forecast (for firm *j* in year *t*) and the actual value of EPS being forecast, scaled by the market share price of firm *j* as of the end of the last fiscal year. (Data Sources: I/B/E/S and CRSP)

**LN(BM*$_{j,s-1}$*)*: the natural logarithm of firm *j*'s book value of equity divided by its market capitalization at the end of fiscal year *s*-1, i.e., LN(BM) = ln(ceq/(csho × prcc_f)), used as a control variable in year *t*. (Data Source: Compustat)

$LN(FIRM\_EXP_{i,j,t})$: the natural logarithm of 1 plus the number of days since the release day of the first annual EPS forecast for firm *j* made by analyst *i*. (Data Source: I/B/E/S)

**$LN(FORECAST\_EXP_{i,t})$:* the natural logarithm of 1 plus the number of days since the release day of the first annual EPS forecast made by analyst *i.* (Data Source: I/B/E/S)

$LN(HORIZON_{i,j,t})$: the natural logarithm of 1 plus the number of days between the release date of analyst *i*'s earnings forecast for firm *j* and the data date of the realized earnings being forecast. (Data Source: I/B/E/S)

**$LN(SIZE_{j,s-1})$:* the natural logarithm of firm *j*'s market capitalization at the end of fiscal year *s*-1, i.e., LN(SIZE) = ln(csho × prcc_f), used as a control variable in year *t*. (Data Source: Compustat)

$LOSS_{i,j,t}$: an indicator that equals 1 if the forecast (for firm *j* in year *t*) made by analyst *i* is negative and equals 0 otherwise. (Data Source: I/B/E/S)

$NEWSNUM_{i,s-1}$: the number of news stories about the state in which analyst *i* is located in the previous week. (Data Source: Bloomberg)

**$NFIRM_{i,t}$:* the number of firms for which analyst *i* generates annual EPS forecasts in fiscal year *t.* (Data Source: I/B/E/S)

**$NSIC3_{i,t}$:* the number of 3-digit SIC industries for which analyst *i* makes annual EPS forecasts in fiscal year *t.* (Data Source: I/B/E/S)

$OHHI_{i,j,t} := \sum_{all\ q} \left(\frac{Number\ of\ prior\ co-analysts\ in\ location\ q_{j,t}^{\neq i}}{Total\ number\ of\ prior\ co-analysts_{j,t}^{\neq i}} \right)^2$, where $Total\ number\ of\ prior\ co-analysts_{j,t}^{\neq i}$ is the total number of other non-*i* analyst who have annual EPS forecasts published before publication of analyst *i*'s last annual EPS forecast in the same firm fiscal year (i.e., firm *j* year *t*). $Number\ of\ analysts\ in\ location\ q_{j,t}^{\neq i}$ is the number of prior co-analysts in the *q*th location where *q* is one of the U.S. Metropolitan Statistical Areas (MSAs) or a city which is not in a U.S. MSA. (Data Sources: I/B/E/S, Nelson's Directories)

**$PROFIT_{j,s-1}$:* the operating income of firm *j* for fiscal year *s*-1 over the book value of equity as of the end of fiscal year *t*-2, i.e., PROFIT = ib/lagged ceq, used as a control variable in year *t*. (Data Source: Compustat)

**$RET_{j,s-1}$:* the average monthly stock returns for firm *j* for the past 12 months in relation to the date of the last actual annual earnings in year *s*-1, used as a control variable in year *t.* (Data Sources: I/B/E/S and CRSP)

$RETTODATE_{i,j,t}$: the cumulative stock returns (using monthly data) for firm *j* between the data date of the last annual earnings and the date on which the earnings forecast made by analyst *i* is released. (Data Source: I/B/E/S)

$RHO_{j,t} := \frac{1}{2} ln \left(\frac{1+\rho}{1-\rho} \right)$, where $\rho = \frac{SE - \frac{D}{N}}{\left(1-\frac{1}{N}\right) D + SE}$ where *SE* is the squared difference between the mean of the last annual EPS forecasts of all analysts, who make forecasts for firm *j* in year *t*, and the actual value of EPS being forecasted, scaled by the market share price of the firm as of the end of the previous fiscal year. *D* is the variance of the individual scaled forecast of the analysts. *N* is the number of analysts. (Data Sources: I/B/E/S and CRSP)

$RISK\_NEWSNUM_{i,j,s-1}$ (in 10,000s): the number of news stories about the state in which analyst *i* is located that mention terms associated with risk/uncertainty in the previous week, scaled by the total number of risk/uncertainty news stories for all 50 states; the variable takes a value of 0 when the state of analyst *i*'s location is not a relevant state for firm *j* (being forecast).

$RISK\_NEWSNUM\_GROWTH_{i,j,s-1}$ is the growth of *RISK_NEWSNUM* to the previous week from the average of the immediate prior 4 weeks, and takes a value of 0 when the state of analyst *i*'s location is not a relevant state for firm *j* (being forecast).

$SIGMA_{i,j,t}$: the variance of the raw monthly stock returns for firm *j* for the past 12 months in relation to the month in which the EPS forecast (for firm *j* made by analyst *i*) is released. (Data Sources: I/B/E/S and CRSP)

**$SP500_{i,j,t}$:* an indicator that equals 1 if firm *j* is in the S&P 500 index when the forecast, made by analyst *i*, is released and equals 0 otherwise. (Data Source: CRSP)

**$VOL\_ROE_{j,s-1}$:* the variance of the residuals from an AR(1) model for firm *j*'s annual ROE using the past 10-fiscal-year series, used as a control variable in year *t*. ROE is calculated as the ratio of earnings to the beginning book value of equity, i.e., ROE = ib/lagged ceq. (Data Source: Compustat)

Appendix B: **A Model of Analyst Learning**

We build on the model introduced in Section 3 to incorporate individual analyst learning from the forecasts of other analysts.

Individual Analyst Forecast Error and Location Diversity

Let the forecasts of all analysts observed by analyst *i* before making his own forecast be given by:

$$x_{wt} = f + \epsilon_{wt}^{-}\,,\ \ w \neq i. \qquad \text{(A1)}$$

Analyst *i* then observes his own signal $f + \epsilon_{it}$, and posts his own forecast. The covariance of ϵ_{it} with ϵ_{wt}^{-}, $w \neq i$, $Cov(\epsilon_{wt}^{-}, \epsilon_{it})$, is denoted by σ_{iw}'. For w, $k \neq i$, assume that $Cov(\epsilon_{wt}^{-}, \epsilon_{kt}^{-}) = c$.

We assume $\sigma_{iw}' \leq c$. In our setting, the covariance structure of errors directly reflects the diversity of analyst locations. We can interpret *c* to reflect the *average* diversity of the location of analysts whose forecasts are available to analyst *i* when he makes his last forecast in a firm-year. Greater diversity corresponds to lower *c*.

The variance of ϵ_{it} is assumed to be σ_i^2, and we assume that the variance of ϵ_{wt}^{-} is the same for all *w*, and is denoted by *s*. The inverse of these parameters captures, respectively, the informativeness, or precision, of analyst *i*'s signal and that of the other analyst following the firm. The distinction is useful when we consider the effect of a change in the quality of analyst *i*'s own signal or that of the average quality of the signal of the other analysts as a group on the weight analyst *i* assigns to the mean forecast of other analysts.

We assume that analyst *i*, in making the last forecast, assigns a weight $(1-\alpha_i)$ to the mean forecast of the previous round of all other analysts, and a weight α_i to his own signal (or his own Bayesian update of forecast, given his signal, that is, what his forecast would be if he did not observe other analysts' forecasts) in the last round. We show that the analyst can improve forecast accuracy by choosing an appropriate value of α_i.

Then, an analyst's forecast is:

$$\hat{x}_{it} = f + \alpha_i \epsilon_{it} + (1-\alpha_i)\frac{\sum_{w \neq i} \bar{\epsilon}_{wt}}{N-1} \quad \text{(A2)}$$

It is straightforward to show that the standard deviation of $\hat{x}_{it}$ is given by $\hat{\sigma}_i$, where

$$\hat{\sigma}_i^2 = \sigma_f^2 + \left(\sigma_i^2 \alpha_i^2 + s\frac{(1-\alpha_i)^2}{N-1}\right) + (1-\alpha_i)\left(2\alpha_i \sigma'_{iw} + c(1-\alpha_i)\frac{N-2}{N-1}\right) \quad \text{(A3)}$$

The *MAD* of analyst *i*'s last forecast of the round (henceforth *IERR* – individual error) is given by Equation (3.2), with $\sigma = \hat{\sigma}_i$. The *IERR* is increasing in the covariance *c*. If diversity lowers the covariance (Hypothesis H1), an analyst's forecast accuracy will improve with greater diversity of the other analysts. Differentiating the expression for MAD in Equation (3.2) with respect to *c*, and noting that, for the individual analyst, σ in that expression are replaced by $\hat{\sigma}_i$, respectively, we have:[20]

$$\frac{\partial(IEER)}{\partial c} = \frac{\partial(IEER)}{\partial \hat{\sigma}_i}\frac{\partial \hat{\sigma}_i}{\partial c} = \frac{\sqrt{2}}{\sqrt{\pi}}\frac{1}{2\hat{\sigma}_i}(1-\alpha_i)^2\frac{N-2}{N-1} > 0 \; if \; 1 > \alpha_i. \quad \text{(A4)}$$

This leads to **Hypothesis 2**.

Hypothesis 2 was derived under the assumption of a fixed α_i. This is appropriate if the choice of the weight is determined by factors outside this framework, such as career concern, reputation, or incentives to curry favour from management. However, if α_i is chosen optimally to minimize expression (A3), by virtue of the Envelope Theorem, Hypothesis 2 remains valid.

[20] It could be argued that the parameter c corresponding to location diversity should be related to a "primitive" covariance between the signal noises of co-analysts before learning has occurred. Below, we show that the covariances of the errors are also positively related to such a primitive *c* after learning has occurred.

Thus, both *uninformed herding* and *observational learning* as modelled here are consistent with a positive association between analyst's location diversity and individual forecast error.

To distinguish implications of learning that are distinct from those of uninformed herding, we consider the optimal choice of α_i under learning. We assume that the objective of an analyst is to minimize the mean absolute forecast error. From Equations (3.2), it immediately follows that this is achieved by choosing α_i to minimize the expression in Equation (A3). The solution is:

$$\alpha = \frac{1}{1+\frac{(N-1)(\sigma_i^2-\sigma'_{iw})}{s-c+(N-1)(c-\sigma'_{iw})}} \quad . \tag{A5}$$

Note that under our assumptions, $0 < \alpha < 1$.

From this, the following comparative static results follow:

For *N>2*:

$$(i)\,\frac{\partial\alpha}{\partial s} > 0;\;\; (ii)\;\frac{\partial\alpha}{\partial\sigma_i^2} < 0\;. \tag{A6}$$

The results are intuitive: an analyst assigns a lower weight to his own signal (lower α_i), i.e., a higher weight to the mean forecast of the co-analysts, if these analysts have higher precision of their signals (lower s). Lower precision of the analyst's own signal (higher σ_i^2) causes the analyst to place higher weight on the co-analysts' mean forecast

Hypothesis 3: The first implications of part 1, i.e., H3(i), follows from the comparative static results in Equation (A6) since the closer the other analysts are to the firm being followed, the more precise is their signal (lower s). This leads to both lower α_i and lower $\hat{\sigma}_i$ (recall that we do not need to consider the effect of α_i on $\hat{\sigma}_i$ since the latter is already optimized with respect to α_i). From Equation (A4), it follows that since α_i and $\hat{\sigma}_i$ both decrease as s falls, the sensitivity of the *IEER* to *OHHI* increases.

Considering H3(ii), result (ii) in Equation (A6) implies that α_i will be smaller when σ_i^2 is higher, which we argue is the case when the analyst is more distant from the same industry peer firms. However, the sensitivity of *IERR* to *OHHI* as per Equation (A4) is not unambiguously higher since $\hat{\sigma}_i$ also increases. We fix the number of analysts at *N=10* and normalize the variance/covariance magnitudes by assuming $s = 1$. The remaining parameter values are as follows: $c = 0.5, \sigma_{iw}' \leq 0.5$, and $\sigma_i^2 \in (\sigma'_{iw}, 5)$. The expression $\frac{1}{2\hat{\sigma}_i}(1 - \alpha_i)^2 \frac{N-2}{N-1}$ representing the sensitivity of the *IERR* with respect to *c* is monotonically increasing in σ_i^2 in this range, α_i is monotonically decreasing, and $\hat{\sigma}_i$, while increasing, is almost flat (see Figure A1 below). Varying *c* does not affect this pattern.

Comment: Note that the comparative static results in Equation (A6) directly imply the effect of parameter changes on α_i. We do not try to test for these effects in our empirical design for the following reason. From Equation (A2), we can write an individual analyst's forecast as $\hat{x}_{it} = (1 - \alpha_i)\left(f + \frac{\sum_{w \neq i} \bar{\epsilon_{wt}}}{N-1}\right) + \alpha_i(f + \epsilon_{it})$. The researcher observes the forecast, $\hat{x}_{it}$, and the mean of the other analysts' forecasts, $f + \frac{\sum_{w \neq i} \bar{\epsilon_{wt}}}{N-1}$, which is the Consensus based on the latest earlier forecasts by other analysts, C_{-i}. Unfortunately, the coefficient β of C_{-i} from a regression $\hat{x}_{it} = \beta C_{-i} + Analyst\ Controls + \eta_i$ does not recover (1-α) because the error $\eta_i = \alpha_i(f + \epsilon_{it})$ is correlated with C_{-i}, creating an endogeneity bias.

Chen and Jiang (2006) regress the difference in an analyst's forecast and the actual earnings on the difference of the Consensus and the actual earnings (called *Deviation*), which avoids this endogeneity bias. However, in that setting, the coefficient of *Deviation* captures the extent to which the weight an analyst assigns to his own signal deviates from the correct Bayesian weight reflecting the information content of the signals. Since the correct Bayesian weight changes when the information content of signals change, this framework cannot be used to address how the weight assigned by the analyst changes.

Comparative Statics

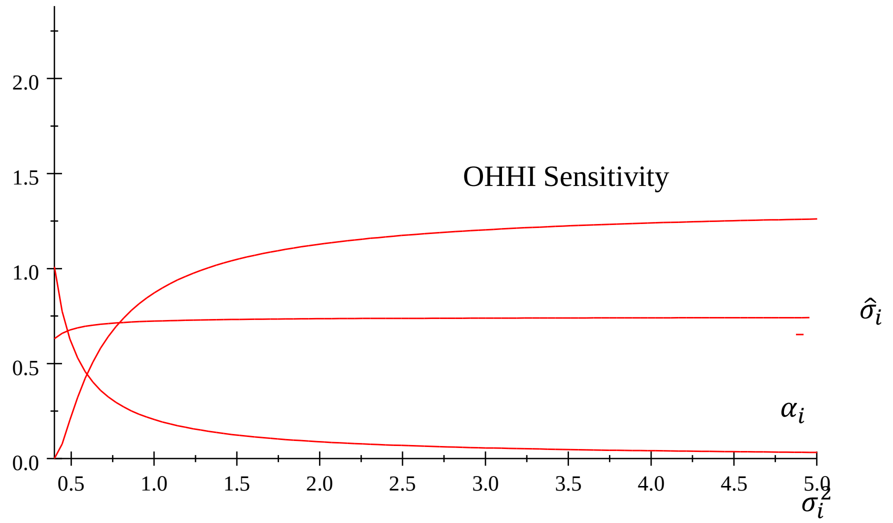


Figure A1. Parameters: $N = 10, s = 1, c = 0.5, \sigma_{iw} = 0.4$, and $\sigma_i^2 \in (\sigma_{iw}, 5)$.

Error covariance and a "primitive" c

The forecast noise covariances between prior co-analysts in our model (denoted by *c*) are affected by these analysts learning from the forecasts of analysts that precede them. It could be argued that the *OHHI* should be related to a more primitive covariance – one that captures the covariance between signal noises before learning has occurred. We show that the forecast noise covariances after learning has occurred increase in such a primitive *c* under plausible conditions.

To do so, we calculate the correlation between the forecasts of prior co-analysts when these forecasts already reflect the learning of these co-analysts from their respective peers (prior co-analysts). Such a covariance should replace the parameter *c* in Equation (A3). We show that, under plausible conditions, this covariance is increasing in a primitive covariance *c*, which is the covariance between analyst-specific signal noises.

For a pair of analysts *i* and *j*, we have

$$Cov\left(\left(\alpha_i\varepsilon_i + (1-\alpha_i)\frac{\sum_{w\neq i}\epsilon_{\bar{w}}}{N-1}\right)\left(\alpha_j\varepsilon_j + (1-\alpha_j)\frac{\sum_{w\neq j}\epsilon_{\bar{w}}}{N-1}\right)\right)$$

$$= \left(\alpha_i\alpha_j Cov(\varepsilon_i,\varepsilon_j) + \alpha_i(1-\alpha_j)\frac{\sum_{w\neq j} cov(\epsilon_{\bar{w}},\varepsilon_i)}{N-1}\right)$$

$$+\left(\alpha_j(1-\alpha_i)\frac{\sum_{w\neq i} cov(\epsilon_{\bar{w}},\varepsilon_j)}{N-1} + (1-\alpha_i)(1-\alpha_j)\frac{cov(\sum_{w\neq i}\epsilon_{\bar{w}},\sum_{w\neq j}\epsilon_{\bar{w}})}{(N-1)^2}\right)$$

$$= \alpha_i\alpha_j c + \alpha_i(1-\alpha_j)\sigma_w' + \alpha_j(1-\alpha_i)\sigma_w' + (1-\alpha_i)(1-\alpha_j)c$$

Here, we have assumed that the common correlation between the signal noises $\varepsilon_i$and ε_j in the last round is the same as that in the prior round – both being denoted by the primitive *c*.

For fixed α_i and α_j this expression is increasing in c. For the symmetric case, since α is increasing in c, it is easy to check that the covariance is increasing in c for α ≥ ½ (a sufficient condition, but hardly necessary). Moreover, we get an increasing relationship between the covariances and the primitive c once we endogenize α as per Equation (A6) for a large set of plausible parameter values. Figure A2 provides an example, in which N is set equal to 10, and firms *i* and *j* are symmetric. For *c* in the range indicated below, the common α ranges from 0.16 at *c*=0.3 to 0.52 at *c*=0.7.

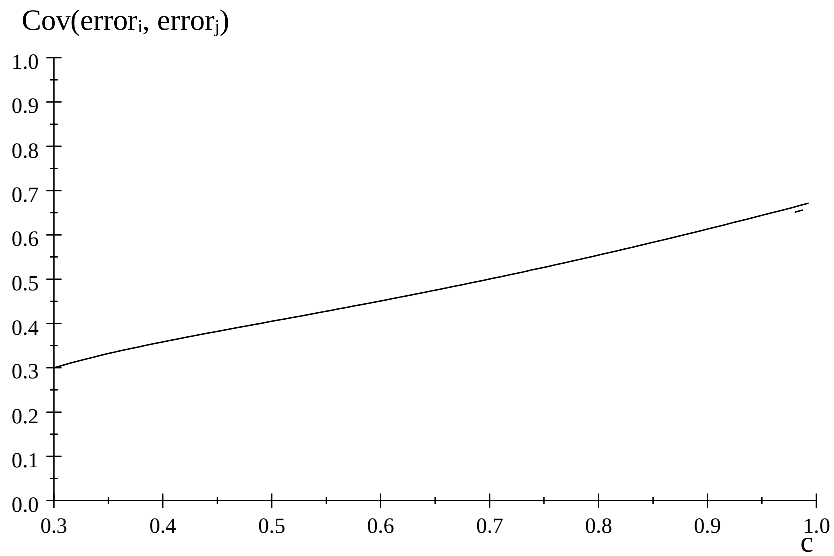


Figure A2. Parameter values: $s = 1, \sigma_i^2 = 0.7,\ \sigma_w' = 0.3.$

Appendix C

Figure A3. U.S. Distribution of Analysts (2006)

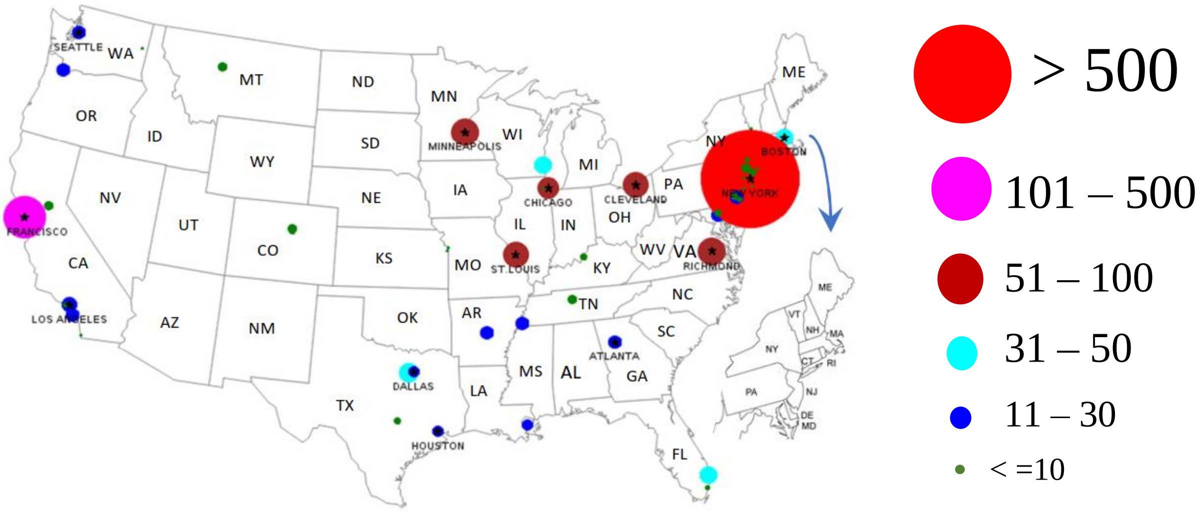
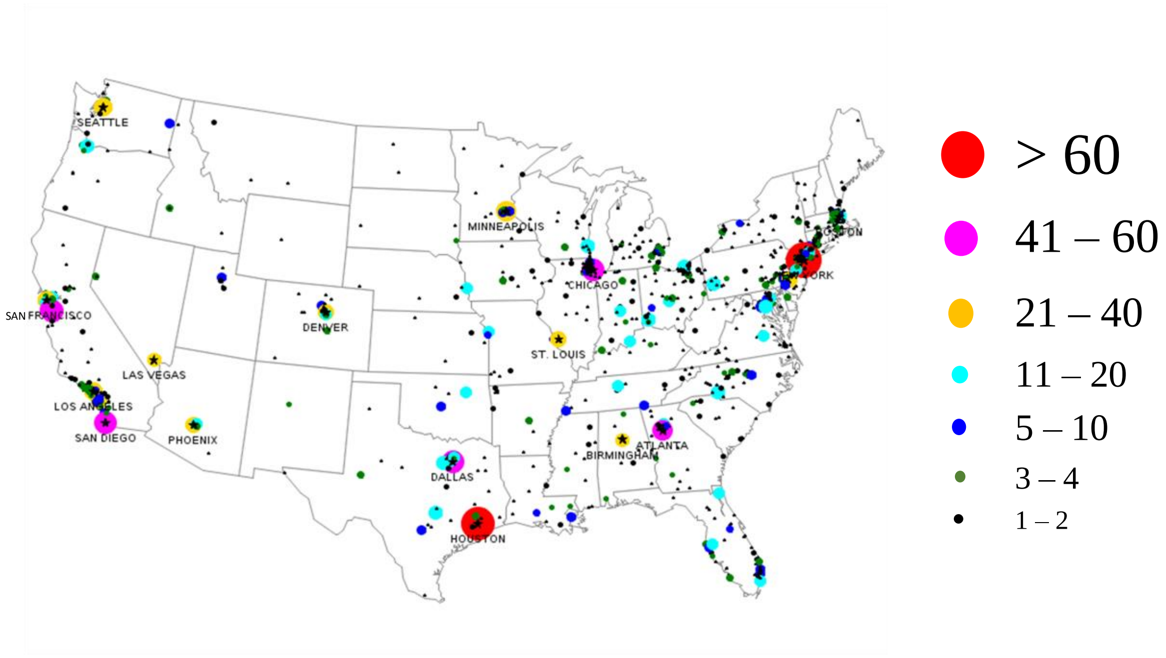


Figure A4. U.S. Distribution of Headquarters of Listed Firms (2006)



Appendix D. Importance of Analysts' Locations for Their Forecasts

Table A1. Sensitivity of Analysts' Forecasts to Number of Risk News and Total Number of News Related to their States

Columns (1a) and (1b) consist of firm-analyst observations for which the state in which the analyst is located ("analyst's state") is a relevant state for the business of the firm covered by the analyst (Garcia and Norli, 2012). *FNUM* is the analyst's weekly number of forecasts. *NEWSNUM* is the number of news stories about the analyst's state in the previous week. *BIAS RELATIVE TO MEAN(MEDIAN) CONSENSUS* is the difference between the analyst's last forecast and the mean(median) of last forecasts of all analysts in the same firm year, scaled by the share price of the firm as of the end of the previous fiscal year. *RISK_NEWSNUM* (in 10,000s) is the number of news stories about the analyst' that mention terms associated with risk/uncertainty in the previous week, scaled by the total number of risk/uncertainty news stories for all 50 states; the variable takes a value of 0 when the analyst's state is not a relevant state of the firm. *RISK_NEWSNUM_GROWTH* is the growth of *RISK_NEWSNUM* to the previous week from the average of the immediate prior 4 weeks, and takes a value of 0 when the analyst's state is not a relevant state of the firm. Refer to Appendix A for the definitions of the other variables. Columns (2a), (2b), (3a) and (3b) consist of forecasts published more than 90 days prior to the fiscal year end. Estimated coefficients and the robust standard errors (in parentheses) are reported. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
	LN(FNUM)	FNUM	BIAS RELATIVE TO MEAN CONSENSUS		BIAS RELATIVE TO MEDIAN CONSENSUS	
LN(NEWSNUM)	0.160***					
	(0.038)					
NEWSNUM		0.384**				
		(0.187)				
RISK_NEWSNUM			-0.081*		-0.082*	
			(0.041)		(0.041)	
RISK_NEWSNUM_GROWTH				-0.012**		-0.013**
				(0.006)		(0.005)
LN(FORECAST_EXP)			0.015**	0.015**	0.007	0.007
			(0.006)	(0.006)	(0.005)	(0.005)
LN(FIRM_EXP)			0.000	0.000	0.000	0.000
			(0.002)	(0.002)	(0.002)	(0.002)
NFIRM			-0.000	-0.000	0.000	0.000
			(0.000)	(0.000)	(0.001)	(0.001)
NSIC3			-0.001	-0.001	-0.003	-0.003
			(0.002)	(0.002)	(0.002)	(0.002)
LN(HORIZON)			0.331**	0.331**	0.332**	0.331**
			(0.131)	(0.132)	(0.132)	(0.132)
LOSS			-0.845***	-0.845***	-0.743***	-0.743***
			(0.087)	(0.087)	(0.080)	(0.080)
SIGMA			1.255***	1.253***	1.269***	1.267***
			(0.158)	(0.157)	(0.235)	(0.234)
RETTODATE			0.755***	0.755***	0.692***	0.692***
			(0.057)	(0.057)	(0.059)	(0.059)
CONSTANT	0.191***	-0.632***	-1.966**	-1.973**	-1.920**	-1.927**
	(0.004)	(0.024)	(0.706)	(0.707)	(0.713)	(0.714)
Fixed Effects	Analyst Year-week	Analyst Year-week	Analyst	Analyst	Analyst	Analyst
S.E. Clusters	Analyst Year-week	Analyst Year-week	Analyst Firm, Yr	Analyst Firm, Yr	Analyst Firm, Yr	Analyst Firm, Yr
Observation	1,421,524	1,413,820	574,405	574,405	574,405	574,405
Adjusted/Pseudo R^2	0.163	0.164	0.0783	0.0783	0.0705	0.0704

Appendix E

Figure A5. BUS DIS of Prior Co-Analysts and Sensitivity of ∆IERR to ∆OHHI