10-2017

Do security analysts learn from their colleagues?

Kenny PHUA

T. Mandy THAM
Singapore Management University, mandytham@smu.edu.sg

Chi Shen WEI
Singapore Management University, cswei@smu.edu.sg

Do security analysts learn from their colleagues?

Kenny Phua, Mandy Tham, and Chishen Wei*

July 31, 2017

Abstract

We examine how learning from colleagues affects security analyst forecast outcomes. We represent the brokerage house as an information network of analysts connected through industry overlaps in their coverage portfolios. Analysts who are more centrally connected in their brokerage network produce more accurate forecast estimates and generate more influential forecast revisions. Consistent with learning, more central analysts tend to unwind their colleagues' recent forecast errors in their forecast revisions. Learning appears to benefit all colleagues, as working at more interconnected brokerages (i.e., denser networks) improves forecast accuracy for all analysts.

JEL Classification Code: D83, G17, G24

Keywords: Learning, Networks, Analyst Forecast Accuracy

* Corresponding author: Chishen Wei (cswei@smu.edu.sg) +65 6828-9650. Phua is at Nanyang Business School, Nanyang Technological University, Singapore. Tham is at Wealth Management Institute (Singapore). Wei is at the Lee Kong Chian School of Business, Singapore Management University. We are grateful for helpful comments from Gennaro Bernile, Ekkehart Boehmer, Steve Dimmock, Umit Gurun, Jarrad Harford, Harrison Hong, Byoung-Hyoun Hwang, Chuan-yang Hwang, Roger Loh, Abhiroop Mukherjee, Chayawat Ornthanalai, Rui Shen, Avanidhar Subrahmanyam, Rong Wang, Bohui Zhang, Zhuo Zhong, and seminar and conference participants at the Asian FA Conference (Bangkok), CICF (Xiamen), Complexity Institute at NTU, EFA (Oslo), FIRCG (Melbourne), University of Melbourne, Nanyang Business School, Singapore Management University, and SFS Cavalcade (Nashville). All errors are our own. The authors acknowledge research support from Nanyang Technological University.

Do security analysts learn from their colleagues?

ABSTRACT

We examine how learning from colleagues affects security analyst forecast outcomes. We represent the brokerage house as an information network of analysts connected through industry overlaps in their coverage portfolios. Analysts who are more centrally connected in their brokerage network produce more accurate forecast estimates and generate more influential forecast revisions. Consistent with learning, more central analysts tend to unwind their colleagues' recent forecast errors in their forecast revisions. Learning appears to benefit all colleagues, as working at more interconnected brokerages (i.e., denser networks) improves forecast accuracy for all analysts.

A large body of research shows that securities analysts play an important role as information intermediaries in financial markets. For example, security analysts supply valuable information (e.g., Womack, 1996; Chen, Francis, and Jiang, 2005; Bradley, Clarke, Lee, and Ornthanalai, 2014; Loh and Stulz, 2011, 2017) and shape the information environment (e.g., Balakrishnan, Billings, Kelly, and Ljungqvist, 2014; Lou, Cohen, and Malloy, 2014; Merkley, Michaely, and Pacelli, 2017).[1] This naturally raises the question of how forecast estimates are generated (e.g., Bradshaw, 2011). Extant studies find that forecast accuracy is a function of individual attributes including experience, personal characteristics, personal ties, and professional connections. However, even the best security analysts do not work in isolation. Their brokerage houses provide not only operational and back-office resources, but also a group of colleagues. These colleagues provide a potentially valuable network of knowledge and information.

For example, an analyst covering Google may provide useful industry insights to a colleague[2] covering Apple. This type of information sharing is valuable because evidence suggests that industry knowledge helps produce better forecast estimates.[3] Indeed, our conversations with analysts suggest that colleagues often workshop ideas and solicit feedback from their colleagues. Lehman Brother's research department in the early 1990s is an example of the importance of knowledge sharing. To foster communication and learning, Lehman Brothers instituted a policy that every analyst's presentation must reference the work of at least two other colleagues. During that period, Lehman Brothers was regularly ranked among the top brokerage firms.[4]

We hypothesize that analysts produce higher-quality equity research if they are better able to tap into the knowledge base of their in-house colleagues. To test our hypothesis, we map the information network within a brokerage house using industry overlaps among analyst coverage portfolios. Brown et al. (2015) write, "*Industry knowledge is the single*

[1] Securities analyst also influence financial policy and valuation (e.g., Lang, Lins, and Miller, 2004; Chang, Dasgupta, and Hilary, 2006; Derrien and Kecskes, 2013).
[2] Throughout the paper, we use colleague to refer to analysts that work at the same brokerage house.
[3] See for example Boni and Womack (2006), Clement, Koonce, and Lopez (2007), Kadan, Madureira, Wang, and Zach (2012), Hilary and Shen (2013), Brown et al. (2015), Bradley, Gokkaya, and Liu (2017).
[4] "The Risky Business of Hiring Stars," Harvard Business Review, May 2004

most useful input to analysts' earnings forecasts and stock recommendations." Our premise is that information exchange is more likely to occur between a pair of analysts if there is industry sector overlap in their coverage portfolios.

Figure 1 provides an example of an actual brokerage network in our sample. A larger node represents an analyst who shares more overlaps in industry coverage relative to a periphery colleague represented as a small node. For example, the largest node in red represents an analyst with direct connections with every other colleague in the network, whereas the smallest node is only connected to five other nodes. To measure each analyst's position in their brokerage network, we create an *Analyst Centrality* score that encapsulates four individual established network measures that capture various aspects of knowledge exchange. An analyst located in a central nodal position has a high *Analyst Centrality* score and is at the epicenter of information exchange within the brokerage network. If analysts can learn from their colleagues, we hypothesize that central analysts are in the best position to benefit from such knowledge exchange.

Our first set of tests indicate that analysts with higher *Analyst Centrality* scores produce better equity research. Central analysts make more accurate forecasts and their forecasts are more influential. Economically, an analyst who is in the 75th percentile of *Analyst Centrality* is about 10.5% more accurate than one in the corresponding 25th percentile (relative to the median forecast error).[5] Conditional on the deviation from the consensus forecast, an inter-quartile increase in *Analyst Centrality* is also associated with a 0.39% higher abnormal return to revision announcements. Our test specifications include firm-year fixed effects to ensure that our findings are not due to underlying heterogeneity in the coverage firm. We also conduct tests with brokerage-year fixed effects. This specification is important because it rules out common alternative explanations relating to brokerage prestige, sector specialization, resources, and other important yet unobservable brokerage level characteristics. The evidence is consistent with the view that

[5] To put this into perspective, a similar inter-quartile increase in General Experience corresponds to a 14.8% improvement in forecast accuracy.

analysts with higher *Analyst Centrality* have better access to the information transmitted through their brokerage network.

Analysts face other incentives that temper the benefits of in-brokerage information exchange. Although colleagues never cover the same company, they do compete for bonus pools or promotions within the brokerage. This type of tournament competition may weaken cooperation and diminish the incentives to share knowledge between colleagues. We find that such competitive pressures, which are more prominent at large investment banks, generally weaken the relation between analyst centrality and research quality.

We design additional tests to understand the nature of information sharing among colleagues. Central analysts are better able to forecast earnings on complex companies as measured by multi-industry segments (e.g., Cohen and Lou, 2012). This suggests that central analysts are experts at piecing together information from different industries. In addition, central analyst can more accurately forecast hard-to-value firms as measured by R&D. This is consistent with the view that central analysts acquire more difficult to uncover insights from their colleagues.

Next, we explore a possible way that central analysts learn such insights following the following learning paradigm (Clement, Hales, and Xu, 2011). In this paradigm, an analyst observes the *ex-post* forecast errors of her colleagues and in response, revises her own forecasts to incorporate the newly revealed information and to unravel the information previously gleaned from her colleagues. Consistent with this paradigm, we find that analysts with higher *Analyst Centrality* are more likely to issue revisions that adjust for their colleagues' *ex-post* forecast errors. This evidence shows a potential channel of information propagation through the network as central analysts take into account their colleagues' mistakes.

An interesting question remains whether information flows among *all* analysts in the brokerage network. Using a network measure, we estimate the amount of interconnection among colleagues (i.e., 'density' per network theory). We also expect greater information flow in brokerages with higher Consistent with this hypothesis, we find that denser brokerages produce more accurate forecasts.

An unaddressed issue is the matching process between the analyst and the brokerage house. Concerns relating to brokerage characteristics are partially addressed in our main tests with the inclusion of brokerage-year fixed effects in our baseline specification. However, it remains that an analyst with expertise in a particular industry is more likely to match with a brokerage that specializes in that industry. To explore the causal relation between *Analyst Centrality* and forecast accuracy, we exploit shocks to *Analyst Centrality* using a sample of brokerage mergers (e.g., Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012) from years 2000 to 2007. These mergers introduce exogenous changes to the acquirer's brokerage structure, shocking the *Analyst Centrality* scores of existing analysts. We find that analysts who experienced increases in their *Analyst Centrality* scores subsequently issued more accurate forecasts relative to analysts whose *Analyst Centrality* scores declined. Results from a generalized difference-in-difference model (Autor, 2003) show that 1) increases/declines in *Analyst Centrality* are effectively random with respect to pre-treatment forecast accuracy, and 2) gaps in forecast accuracy between both groups exist only in the post-treatment periods.

One may wonder to what extent the *Analyst Centrality* measure reflects an analyst's ability or effort. Of course, developing higher centrality is completely consistent with greater talent or skill. What distinguishes our learning hypothesis is that *Analyst Centrality* has incremental explanatory power beyond measures of ability or effort found in the prior literature. Notably, our main specifications include proxies for ability (i.e., experience) and effort (i.e., revision frequency). In addition, we proxy for analyst ability using analyst membership in the Institutional Investor All-America Research Team (i.e. All-American status). The effect of *Analyst Centrality* on forecasting outcomes is unchanged with the inclusion of this control in our main tests.

Our paper contributes to a growing literature that seeks to penetrate the 'black box' of information generation of sell-side financial analysts (e.g., Bradshaw, 2011; Brown et al., 2015). Extant studies find that forecast accuracy is a function of experience, location, personal characteristics, personal ties, and professional connections.[6] Our paper emphasizes the value of information transfer between colleagues. Related to our findings is the evidence in Hwang, Liberti, and Sturgess (2016) of information transmission among in-house analysts in the mergers and acquisition setting. Equity analysts may also acquire information from their colleagues in debt/macro research (Hugon, Lin, and Markov, 2016; Hugon, Kumar, and Lin, 2016), on the asset management side (Irvine, Simko, and Nathan, 2004), and on the brokerage trading desk (Li, Mukherjee, and Sen, 2017). In contrast, we measure the general propensity of an analyst to receive information from her equity research colleagues using a network theoretic approach.

Our study also links to a literature that explores how analyst incorporate various signals into their forecasts. Chen and Jiang (2006) show how analysts weight public and private information into their forecasts. Clement, Hales, and Xu (2011) find that analysts revise forecast following the revisions of their direct competitors who cover the same company. Our evidence suggests that colleagues are a valuable source of information.

Our findings also contribute to the growing literature on information transmission through networks in financial markets. Venture capitalists and stock market investors who hold central positions in the network generate better investment returns (Hochberg, Ljungqvist, and Lu, 2007; Ozsoylev, Walden, Yavuz, and Bildik, 2013). Anjos and Fracassi (2015) find that internal information markets help conglomerates deliver more innovative and greater value. Studies also identify information transmission through workplace connections, personal networks, and social connections using educational backgrounds or social affiliations.[7]

[6] See for example Clement (1999), Jacob, Lys, and Neale (1999), Malloy (2005), Bae, Stulz, and Tan (2008), Bradley, Gokkaya, and Liu (2017), Chen and Jiang (2006), Chen and Matsumoto (2006), Clement, Koonce, and Lopez (2007), Cohen, Frazzini, and Malloy (2010), Kumar (2010), Clement, Hales, and Xue (2011), Hilary and Shen (2013), Law (2013), Green, Jame, Markov, and Subasi (2014a, 2014b), Malloy (2005), Soltes (2014), Jiang, Kumar, and Law (2016). For a recent review, see Bradshaw (2011).

[7] For example, social ties with management affect both analyst forecast behavior and coverage decisions (e.g., Westphal and Clement, 2008; Cohen, Frazzini, and Malloy, 2010; Brochet, Miller, and Srivinasan, 2014). Information also transmits through social networks among investors (e.g., Hong, Kubik, and Stein, 2004, 2005; Pool, Stoffman, and Yonker, 2015), between boards and managers (e.g., Kuhnen, 2009), between corporate and bank managers (Engelberg, Gao, and Parsons, 2012), across corporate boards (e.g., Shue,

1. Sample and network definitions

This section describes our methodology and variable construction. We discuss our data and present summary statistics on our sample.

1.1. Defining Network Connections

Studies show that industry-specific information is valuable in making accurate earnings forecast.[8] Therefore, we presume that information transfer and dissemination is likely to occur between analysts that share industry sector overlaps in their coverage portfolios. This may occur for direct economic reasons. First, given the significant benefits of industry knowledge, there is economic value to discuss and solicit feedback from colleagues in the same industry. Second, given economies of scale in underlying resources such as sharing support staff, research assistants, and data sources, these types of interactions are likely to naturally develop among analysts in the same sector. Our study focuses primarily on these channels of information dissemination, but we discuss the possibility of other related channels in the next section.

We construct the connections within-brokerage network based on the forecast data for fiscal year one from the Detailed History file of I/B/E/S. We use the Global Industrial Classification Standard (GICS) codes to classify industries because these are commonly used among practitioners (e.g., Bhojraj, Lee, and Oler, 2003). To ensure that our networks are meaningfully large, we require that a brokerage in a given year has at least 5 analysts.

We allow the network structures within the brokerage to vary through time to allow for changes in an analyst's coverage portfolio. For example, consider two analysts, A and B, at the same brokerage. A connection exists in year t if A makes forecasts in GICS sectors 20 and 45, and B makes forecasts in GICS sector 45. By construction, the

2013; Larcker, So, and Wang, 2013, Fracassi, 2016), and between corporate boards and mutual fund managers (Cohen, Frazzini, and Malloy, 2008).

[8] See: Boni and Womack (2006), Clement, Koonce, and Lopez (2007), Kadan et al. (2012), Hilary and Shen (2013), Brown et al. (2015), Bradley, Gokkaya, and Liu (2017). Brown, Call, Clement, and Sharp (2015) write, "industry knowledge is the single most useful input to analysts' earnings forecasts and stock recommendations."

connection weight is independent of the number of unique firms covered in the overlapping GICS sectors by either analyst. While it is reasonable to expect that the intensity of information exchange increases with the number of firms covered in the overlapping sectors of connected analysts, our simple counting approach is more parsimonious and conservative.

[Insert Figure 1]

As an example, we illustrate the network structure of a brokerage (I/B/E/S identifier: 481) in the year 2005. Each circular node represents an analyst in the brokerage network. The numbers below each node identify the GICS sectors covered by the analyst. Most analysts in this brokerage network cover up to two sectors, and only two analysts span four sectors in their coverage portfolios. A line between a pair of nodes denotes a direct connection between a pair of analysts. Nodes that are larger in size and illustrated with more intense colors (progressively, from light green to bright red) have more direct connections.[9] Figure 1 shows that analysts can have different levels of connectedness even if they cover the same number of economic sectors. For example, an analyst who covers GICS sectors 25 and 45 has more direct connections than her colleague who covers GICS sectors 45 and 50.[10] This suggests that the relation between an analyst's coverage portfolio complexity and potential for information exchange is non-linear and dependent on the composition of her colleagues' coverage portfolios.

1.2. Defining Analyst Centrality

There are many dimensions of the information exchange process. Aside from different types of information (e.g., firm-specific, industry news, and macroeconomic news), there is variation in the amount (i.e., volume) and speed of information transmission. Our approach uses the four most common measures of network centrality to describe

[9] While we focus on direct connections in Figure 1, network theory informs us that a node's connectedness is also influenced by her indirect connections in the network. In the next section, we explore some measures that account for dynamics arising from analysts' indirect connections.
[10] This difference arises because there are more analysts in the brokerage network who cover GICS sector 25 than the ones who cover GICS sector 50.

information exchange (e.g., Larker, So, and Wang, 2013). Each measure of network centrality captures a distinct feature of connectivity and information exchange. As such, any single measure is unlikely to capture all facets of information exchange (Newman, 2003). Characterizing the information exchange process can be challenging (Borgatti, 2005), but we describe below the intuition behind the types of inter-colleague information exchange that each measure is likely to capture and reserve the details for Appendix I.

- *Degree Centrality* (Newman, 1979) and *Eigenvector Centrality* (Bonacich, 1972) are conventionally used to capture information flow in a network (Borgatti, 2005). *Degree centrality* counts the number of directly connected neighbors. An analyst with more direct connections has more opportunities for interaction and is likely to receive more information. Since our analysts are connected by industry overlaps, degree centrality likely captures information that is industry-specific or related to competition dynamics.
- *Eigenvector Centrality* is a close cousin of *Degree Centrality* but places more weight on neighbors who are themselves more central whereas *Degree Centrality* places equal weight. This captures more relevant or valuable industry-specific information and additional information flow that is more macroeconomic in nature.
- *Closeness Centrality* (Freeman, 1979) captures the speed (time-to-arrival) of receiving information that is produced anywhere in the network (Borgatti, 2005). In financial markets, the value of information is time-sensitive, and speed of information exchange helps an analyst incorporate colleagues' information into her forecasts in a timely fashion.
- *Betweenness Centrality* (Freeman, 1979) captures the relative position of analyst within the network structure. It captures *cross*-industry information exchange[11] that is valuable to the analyst and relates to the positional benefit of 'strategic complementarity' in information production from colleagues working in different industries. An analyst with high *Betweenness Centrality* is more likely to reap cross-industry synergistic gains from her colleagues' information production. Technically, such cross-industry synergies are unlike the canonical type of information exchange

[11] Kini et al. (2009) find that cross-industry information is valuable to forecast accuracy.

(i.e., degree and eigenvector centrality). Where there is an opportunity for an analyst to exploit synergies between two industries, the information exchange should be targeted and not via a circuitous path.[12]

A large literature has stressed the importance of information transmission through informal or social connections. In contrast, we focus on professional linkages based on sector overlap in their coverage portfolios. It is likely that the analysts we study are connected socially too, beyond these professional relations (e.g., Hwang and Kim, 2011). Data limitations prevent us from capturing these non-professional ties, but we argue that information exchange outside of our constructed networks is likely to work against finding empirical support for our hypothesis.

1.3. Principal Component Analysis

Given the complexity of information in a network, any single measure of network centrality is unable to completely capture analyst connectedness (Newman, 2003; Valente, Coronges, Lakon, and Costenbader, 2008). Each measure of centrality may be better at capturing different types of information exchange, which we further discuss below.

Our four centrality measures have cross-correlations ranging from 30.2% to 89.1%. Simultaneously using all four centrality measures is likely to induce multicollinearity issues. Therefore, we perform a principal component analysis (PCA) on degree, betweenness, closeness, and eigenvector centrality. From the first principal component, we extract its standardized factor score and define it as *Analyst Centrality*. We report more details of the PCA in Section 1.6.

1.4. Forecast Accuracy

Following Clement (1999), we construct *Normalized Forecast Error* from the latest firm-year forecast values by calculating the absolute difference between the analyst's firm-year

[12] Borgatti (2005) makes clear that, due to the algorithmic implications of betweenness centrality, it is incompatible with the canonical type of untargeted information exchange. We also do not adopt the typical interpretation that betweenness centrality represents information brokering capacity. Instead, we stress that analysts with synergistic coverage portfolios will mechanically be situated between many colleagues.

forecast value and the corresponding earnings-per-share (EPS) of the firm-year, scaled by the firm-year mean absolute forecast error for comparability across observations. For analyst i covering firm f in year t, we define her *Normalized Forecast Error* as in (1).

$$Normalized\ Forecast\ Error_{i,f,t} = \frac{\left|EPS\ Forecast_{i,f,t} - Actual\ EPS_{f,t}\right|}{\frac{1}{N_{f,t}}\sum_{j}^{N_{f,t}} \left|\ EPS\ Forecast_{j,f,t} - Actual\ EPS_{f,t}\ \right|} \tag{1}$$

1.5. Control Variables

We include controls for the determinants of forecasting performance found in prior research.[13] We provide a brief description of the key variables and provide full construction details in Appendix II. All continuous variables are winsorized at the 1st and 99th percentile values to limit the influence of outliers.

To control for forecast characteristics associated with accuracy, we include proxies for revision activity (*Revision Frequency*), timeliness based on the closeness to the actual earnings announcement (*Horizon*), lowballing behavior (*Lowball*)[14], and deviation from the consensus forecast (*Boldness*). As lowballing behavior is associated with forecast accuracy, we also create a lowball measure following Hilary and Hsu (2013). Since centrality is likely associated with ability, we proxy for analyst ability using total work experience (*General experience*) and experience covering the firm (*Firm Experience*). To account for the analyst's portfolio complexity, we measure number of unique firms (*Firm Breadth*) and the number of 2-digit GICS sectors (*Industry Breadth*) covered by the analyst during the year. We measure *Brokerage Size* as the logarithm of the number of analysts employed by the brokerage.

[13] See for example, Mikhail, Walther, and Willis (1997); Clement (1999); Jacob, Lys, and Neale (1999); and Brown (2001), Ivković and Jegadeesh (2004).

[14] Hilary and Hsu (2013) show that analysts strategically increase forecast error consistency via lowballing behavior for firms with high institutional ownership. The lowballing strategy inevitably decreases forecast accuracy. Appendix II contains full details of the construction of lowball.

We also account for the influence of firm heterogeneity on forecasting outcomes by controlling for the number of analysts (*Analyst Coverage*), market capitalization, book-to-market ratio, leverage, ROA volatility, and negative earnings (*Loss*)[15].

1.6. Descriptive Statistics

Our sample comprises 5539 firms, 317 brokerages, 9170 analysts, and 274,671 firm-year forecasts from the years 1996 to 2014 using the I/B/E/S vintage from WRDS in May 2015. Panel A presents summary statistics of unique analyst-year pairs. The average (median) analyst has 17 (12) connections and covers 1.68 (1) GICS sectors. The median brokerage employs 44 unique analysts. By definition, *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality*, and *Eigenvector Centrality* are non-negative. However, many analysts in the sample have negative values of *Analyst Centrality* because we used standardized variables in the PCA-extraction.[16]

[Insert Table 1]

Panel B of Table 1 presents Pearson correlations between the network centrality measures and analyst/brokerage characteristics. Naturally, *Analyst Centrality* is positively correlated with its component measures (54.6% to 94.9%). *Analyst Centrality* is also positive correlated with *Industry Breadth* (57.1%), but weakly related to *Firm Breadth* (16.5%), as analysts generally tend to cover firms in the same industry. *Analyst Centrality* has low correlations with measures of experience including *General Experience*, *Firm Experience*, and *Brokerage Experience*. Interestingly, *Analyst Centrality* is negatively correlated with *All-American* status, which suggests that central analyst are unlikely to be voted as All-Americans.

Panel C reports a principal component analysis of the four network centrality measures. The first principal component captures approximately 66.5% of the variance in the four network centrality measures and is the only eigenvalue greater than one. The incremental variance explained by the next principal component is only about 20.9%. Other loadings

[15] See Hwang, Jan, and Basu (1996) and Brown (1997; 1998).
[16] Standardization accounts for the different scales of the four network centrality variables.

on some centrality measures turn negative in principal components, thus making interpretation difficult. In light of these considerations, we construct *Analyst Centrality* based on the first principal component.

[Insert Figure 2]

1.7. Central analysts and industry coverage

As discussed above, central analysts cover more industries but not necessarily more firms. Figure 2 shows the types of industries that central analyst cover. Analysts with high centrality primarily cover information technology, consumer discretionary, and industrials. In contrast, low centrality analysts tend to focus on health care, financials, and energy.

2. Do central analysts produce better forecasts?

Our learning hypothesis proposes that analysts generate better equity research if they can learn valuable information from their colleagues. In this section, we test the learning hypothesis by focusing on the analyst's centrality in the network. Our main prediction is that central analysts make better forecasts because they have better access to intra-brokerage information flow. We test this along two dimensions: 1) forecast accuracy, and 2) market reactions to revision announcements.

2.1. Forecast accuracy

We test whether central analyst produce more accurate forecasts by estimating specification (2).

$$\textit{Normalized Forecast Error}_{i,f,t} = a + \beta_1 \cdot \textit{Analyst Centrality}_{i,t} + \vartheta \cdot \textit{Controls}_{i,f,t} + \varepsilon_{i,f,t} \quad (2)$$

Our forecast accuracy measure is *Normalized Forecast Error* which allows for comparability across firms. The control variables include forecast horizon, revision frequency, forecast boldness, analyst experience, coverage experience, lowballing behavior, and complexity of coverage portfolio (i.e., firm breadth of coverage and industry breadth of coverage). Additionally, we include the following firm characteristics: earnings volatility, transitory earnings, leverage, growth opportunities, firm size, and analyst coverage (e.g., Heflin, Subrahmanyam, and Zhang, 2003; Hilary and Hsu, 2013).

Our baseline specification includes brokerage-year fixed effects to capture unobserved brokerage characteristics such as prestige and resources (e.g., Stickel, 1995; Clement, 1999; Hugon, Kumar, and Lin, 2016). We also estimate a specification with firm-year fixed effects to absorb unobserved heterogeneity of the coverage firm that affects all analysts' forecasting performance in the firm-year. We carefully estimate the standard errors as follows. Since fixed effects do not fully capture correlations of regression residuals, we cluster the standard errors along two dimensions following Petersen (2009) and Gow, Ormazabal, and Taylor (2012). First, we cluster by analyst-firm because an analyst's forecast error on a particular firm may be correlated over time. The second clustering dimension is either 1) firm-year to capture dependence in forecast errors of competing analysts in a firm-year or 2) brokerage-year to capture correlations of analysts' forecasting performance within each brokerage-year.

[Insert Table 2]

The results in Table 2 show that central analysts produce more accurate forecasts than their periphery colleagues. The specification in Column (1) shows that, controlling for analyst characteristics, central analysts produce more accurate forecasts than their periphery colleagues in the same brokerage. In Column (2), we show that the positive relation between forecast accuracy and *Analyst Centrality* is robust to the inclusion of controls for firm characteristics. Economically, an analyst in the 75th percentile of *Analyst Centrality* is 10.5% more accurate (relative to the median forecast error) than an analyst in the 25th percentile. For comparison, an interquartile increase (p25th to p75th corresponding to an interquartile range from 23 months to 90 months) in *General Experience* corresponds to a 14.8% improvement in forecast accuracy.[17] Column (3) shows

[17] From Table 1, the difference in *Analyst Centrality* between the 75th and 25th percentiles of its distribution is 1.337. This difference corresponds to a decrease in *Normalized Forecast Error* of $1.337 \times 0.014 = 0.019$. We de-normalize this value by the mean raw forecast error to obtain $0.019 \times \$0.222 = \$0.0042$, or about 10.5% of the median raw forecast error ($0.04). An interquartile increase in *General Experience* (from 23 months to 90 months) is a 290% increment. This has an impact of $-0.020 \times \log(1+2.9) = 0.027$ decrease

that the results are similar with the inclusion of firm-year fixed effects which necessarily omits controls for firm characteristics. The results suggest that central analysts make more accurate forecasts than their competitors who cover the same firm in the year.

Consistent with prior literature, we find that forecast accuracy is associated with shorter forecasting horizons, more firm-specific experience, bolder forecasts, less low balling behavior (e.g., Clement and Tse, 2005; Hilary and Hsu, 2013). Greater effort, as proxied by higher revision frequency is also associated with greater accuracy. Higher analyst coverage also creates greater accuracy, consistent with greater competition spurring more effort. (e.g., Merkley, Michaely, and Pacelli, 2017; Loh and Stulz, 2017).

We also find that analysts that cover more industries (*Industry Breadth*) tend to have greater forecast errors as their coverage portfolio becomes more complex (e.g., Clement, 1999; Clement and Tse, 2005). This is noteworthy because *Analyst Centrality* measure is related to *Industry Breadth*, but the two measures have opposing effects on forecast accuracy. An interquartile increase in *Industry Breadth* reduces accuracy by 3.3% relative to the median forecast error.[18] This is consistent with a tradeoff between the benefits of intra-brokerage information exchange and the costs of undertaking more complex tasks.

Overall, the evidence is consistent with the view that analysts with higher values of *Analyst Centrality* have greater information advantages. The findings suggest that central analysts are likely learning valuable information from their intra-brokerage colleagues. The relation is unlikely due to unobservable brokerage-related factors or heterogeneity in coverage-firm specific characteristics.

2.2. Forecast revision informativeness: Market reactions

in *Normalized Forecast Error*. On average, this translates to $0.027 \times \$0.222 = \$0.0059$, or about 14.8% of the median raw forecast error ($0.04).

[18] From Table 1, the difference in *Industry Breadth* between the 75^{th} and 25^{th} percentiles of its distribution is 1. This difference corresponds to an increase in *Normalized Forecast Error* of $1 \times 0.006 = 0.006$. We de-normalize this value by the mean raw forecast error to obtain $0.006 \times \$0.222 = \$0.0013$, or about 3.3% of the median raw forecast error ($0.04).

If central analysts benefit from information exchange, we expect that their forecast revisions contain more novel and value-relevant information that will attract greater market reactions. We test this prediction by estimating specification (3).

$$\begin{aligned} AbsCAR\ [\text{-}1,+1]_{r,i,f,t} = a + \beta_1 \cdot Analyst\ Centrality_{i,t} + \\ \beta_2 \cdot Analyst\ Centrality_{i,t} \times Consensus\ Deviation + \vartheta \cdot Controls_{i,f,t} + \varepsilon_{i,f,t} \end{aligned} \tag{3}$$

The unit of analysis is each forecast revision on a firm by a particular analyst. Standard errors are double-clustered by year-week to capture common time-varying macroeconomic shocks and also by firm because market reactions to forecast revisions may be correlated over time for a given firm. We control for *Consensus Deviation* because we expect stronger market reactions when the analyst's revision significantly deviates more from the prevailing consensus forecast (e.g., Clement and Tse, 2003; Hilary and Shen, 2013). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing forecast consensus, normalized by the absolute value of the forecast consensus. We also control for stock performance during the run-up to the forecast revision date to capture potential auto-correlation in returns.

[Insert Table 3]

Table 3 shows that central analysts command larger market reactions around their forecast revisions. Column (1) shows a statistically positive association between *Analyst Centrality* and 3-day market reactions around forecast revisions. Furthermore, the interaction term shows that, conditional on the deviation from the consensus forecast, the revisions of a central analyst command higher market reactions than those of a peripheral analyst. This suggests that the market perceives the revisions of central analysts to contain more relevant and novel information, even after controlling for *Consensus Deviations*.

The results are similar in Column (2) with the inclusion of controls for analyst characteristics and stock performance during run-ups to the forecast revisions. Evaluated at the median *Consensus Deviation*, an inter-quartile increase in *Analyst Centrality* is associated with higher returns of about 0.39% over 3 days.[19] Consistent with Bradley, Gokkaya and Liu (2017), the forecast revisions of analysts who have higher general experience and who work in larger brokerages also attract larger market reactions. Overall, the evidence is consistent with the view that analysts with higher centrality possess more novel and value-relevant information.

A potential concern is that confounding events including material firm announcements may coincide with analysts' forecast revisions. To alleviate this concern, we collect all dates on which firms file SEC Form 8-Ks[20] from the EDGAR database. Thereafter, we drop a forecast revision from our sample if a Form 8-K is filed in the (-1, +1) day window of the revision. Additionally, we exclude a forecast revision from our sample if the firm has earnings announcements in the (-1, +1) day window of the revision. These filters reduce our sample size by about 48%, suggesting that analysts' revisions are often motivated by the disclosures of new information by firms. We repeat our analysis in this reduced sample and present the results in Column (3). Notwithstanding a marginal drop in economic magnitude, we continue to find that the forecast revisions of central analysts attract larger market reactions.

Finally, we repeat our analysis on a further-reduced sample of standalone forecast revisions (Gleason and Lee, 2003; Chen and Matsumoto, 2006).[21] Under the caveat that this filter inevitably introduces a look-ahead bias, we continue to find a positive relation between *Analyst Centrality* and market reactions around forecast revisions in Column (4).

Despite our exclusions of forecast revisions that coincide with issuances of Form 8-Ks and earnings announcements, an alternative explanation remains that central analysts only issue forecast revisions around significant news events. We address this possibility

[19] The difference between the 25th and 75th percentiles of the *Analyst Centrality* distribution is 1.337. The median *Consensus Deviation* in the sample is 0.046. Evaluated at the median level of *Consensus Deviation*, an inter-quartile increase in *Analyst Centrality* corresponds to an increase in market reactions of $0.009\times0.046\times1.337+0.288\times1.337=0.358\%$

[20] In addition to annual and quarterly reports, public companies are required to report certain material corporate events on a more current basis via the Form 8-K. A comprehensive list of the event types that trigger a firm's obligation to file a Form 8-K is available on the SEC website.

[21] Standalone forecast revisions have no other revisions released in the (-1, +1) day window of the revision.

by replacing *Consensus Deviation* with *Self Deviation* in our model. The latter variable captures the magnitude of deviation between an analyst's revision and her previous forecast value. Our results in Table 1 of the Internet Appendix show that, conditional on the magnitudes of such deviations, central analysts still command larger market reactions. Notably, we also find that market reactions to forecast revisions increase unconditionally with *Analyst Centrality.* Overall, the results are consistent with the view that central analysts have access to better information and learn from their brokerage colleagues. As a result, they make more informative forecast revisions that generate larger market reactions.

2.3. Competition and learning

Sell-side analysts work in highly competitive environments. Our main tests control for cross-brokerage competition effects using the number of analysts who covers the same firm (e.g., Merkley, Michaely, and Pacelli, 2017; Loh and Stulz, 2017). However, analysts also face direct in-house competition for year-end bonus or promotions (e.g. Groysberg, Healy, and Maber, 2011). Under intense competition within the brokerage, analysts may abstain from collaboration or withhold feedback, and thus weaken information exchange.

To explore these dynamics, we proxy for the intensity of in-house competition using the existence of large investment banking businesses. We rank investment banks (IB) by their IPO and SEO deal values in every year using data from SDC platinum. Investment banks are added to the *Large IBs* pool sequentially until the pool accounts for 75% of the total deal value in the market of that year.[22] The remaining brokerages (i.e. smaller IBs and non-IBs) are assigned to the *Non-Large IBs* pool.

[Insert Table 4]

[22] We perform this ranking procedure separately for IPO and SEO deals so that a brokerage with substantial IPO deal flow but weak SEO deal flow (or vice versa) is still classified as a *Large IB*. The market for underwriting services is dominated by a few IBs – the 15 *Large IBs* out of 101 IBs in our sample account for at least 75% of the IPO/SEO market share.

The results in Table 4 are consistent with the notion that the learning mechanism weakens in highly competitive environments. In Panel A, we find a positive relation between *Analyst Centrality* and forecast accuracy in the *Non-Large IBs* sample in Column (1); however, this relation is absent among the largest IBs in Column (2). This suggests that central analysts do not benefit strongly from inter-colleague information exchange at the large investment banks. We also use *Brokerage Size* as an alternative measure for in-house competition because larger brokerages tend to offer higher compensation and employment prestige. In every year, we sort brokerages into terciles according to their size. Columns (3) through (5) present results on sub-samples increasing in *Brokerage Size*. The link between *Analyst Centrality* and forecast accuracy appears to increase in the first two *Brokerage Size* terciles, but diminishes among the largest brokerages.

We repeat the above analysis on market reactions surrounding analysts' forecast revisions in Panel B. In a sample that excludes revisions around issuances of Form 8-Ks and earnings announcements, we find that the positive relation between market reactions and *Analyst Centrality* is present among *Non-Large IBs* but not among the largest IBs. Similar to the patterns in Panel A, we also find that the effect of *Analyst Centrality* decays with *Brokerage Size* in the context of market reactions.

Overall, these findings support the salience of competition as a factor in designing an optimal collaborative structure for information exchange among workers.

3. The learning channel

Our evidence thus far is consistent with the view that intra-brokerage information transfer affects forecast performance. In this section, we examine the potential learning channel more carefully.

3.1. Learning information on hard-to-value stocks

We hypothesize that central analyst are more likely to learn information that is more complex in nature from their colleagues. Koh and Reeb (2015) show that firms that fail to report R&D expenditures have higher uncertainty about their true level of innovation and future growth prospects. Information incorporation is also more complicated for firms with highly dispersed operations across numerous industry segments (Cohen and Lou, 2012).

We construct *High R&D,* which is an indicator equal one if either the R&D intensity of the firm is above the yearly median or the firm has missing R&D expenditure data in Compustat[23], and *Conglomerate*, which is the Herfindahl index of a firm's sales across its industry segments. To facilitate interpretation, we multiply the Herfindahl index by minus one so that higher values of *Conglomerate* correspond to higher firm complexity. We interact *Missing R&D* and *Conglomerate* with *Analyst Centrality* respectively to examine incremental effects of colleague learning on stocks that are difficult to value.

Table 5 shows that central analysts are better able to forecast earnings for such hard-to-value stocks. Column (1) of Table 5 shows a statistically negative coefficient on the interaction term *Analyst Centrality* x *High R&D*. Column (2) shows similar results using *Firm Complexity* measure. Column (3) shows that our conclusions are unchanged when including both *High R&D* and *Firm Complexity* (and their interactions with *Analyst Centrality*) in the same specification. Interestingly, the statistically negative loadings on both interaction terms suggest that *High R&D* and *Firm Complexity* proxy for different dimensions of valuation for difficult-to-value stocks.

This evidence provides some color on the nature of information analysts learn from their colleagues. The findings suggests that central analyst are able to gain particularly valuable and rare insights to help forecast uncertain or complex firms.

3.2. Learning from our colleague's mistakes

We examine whether central analysts are able to learn from their colleague's mistakes under the following paradigm. A central analyst observes the ex-post forecast errors of her colleagues when actual earnings are announced in the colleague's coverage portfolio.

[23] R&D intensity is defined as the ratio of R&D expenditures to total assets. Our results also hold if we require sample firms to have non-missing R&D expenditure data.

In response, the central analyst revises her own forecasts to incorporate the newly revealed information in her colleagues' forecast error. We hypothesize that after a colleague's forecast is revealed to be optimistic, the central analyst is likely to learn from this mistake and revise her forecast downwards. Our test focuses on an analyst's revision behavior upon the revelations of *ex-post* forecasting performance of her brokerage colleagues.

To measure the re-adjustment of the analyst's prior forecast, we construct *Analyst Revision*, a signed measure defined as the difference between an analyst's forecast revision value and her prior forecast value, deflated by the absolute value of the latter. Therefore, a positive (negative) value of *Analyst Revision* reflects an increase (a decrease) in the analyst's forecast value from her previous forecast.

For a forecast revision of a given analyst, we collect all instances of her colleagues' realized forecast errors that occurred within the past 30 days. We only retain the realized forecast errors of 1) colleagues who are directly connected to the analyst (i.e. who cover the same industries), and 2) colleagues who cover either the major suppliers or major customers of the analyst's firms.[24] Each realized forecast error is classified as optimistic if the forecasted value is above the actual reported EPS. We define *Colleague Optimism* as the proportion of optimistic forecast errors in the 30-day window. Given an analyst i in brokerage G covering firm f and making a forecast revision on date d, we let $r(j,d)$ equate to unity if colleague j has a realized forecast error 30 days prior to d with the below characteristics, and equate to zero otherwise.

$$Colleague\,Optimism_{i,f,d} = \sum_{j \neq i, j \in G} \frac{optimistic\,revisions_{j,d}}{total\,revisions_{j,d}} \tag{4}$$

[24] We identify customer and supplier pairs from the business segment files of Compustat. In accordance with SFAS 14, public firms are required to disclose sales to their principal customers, defined as customers that contribute to at least 10 percent of the total revenue of the firm or if sales to a customer are material to the business of the firm. Principal customer names are manually matched to Compustat GVKEYs following the approach in Fee, Hadlock, and Thomas (2006). For customer names that are abbreviated, we hand-match and use industry affiliations to determine whether the customer is in Compustat. For the remaining unmatched customers, we check their corporate websites in the Directory of Corporate Affiliation (DCA) database to determine if the customer is a subsidiary of a listed firm. If so, we assign the customer to its parent's GVKEY. To ensure accuracy, we discard any customer name that cannot be unambiguously matched to a GVKEY.

To test the learning channel, we estimate specification (5).

$$\begin{aligned} Analyst\ Revision_{i,f,d} = {} & a + \beta_1 \cdot Analyst\ Centrality_{i,t} \times Colleague\ Optimism_{i,d} \\ & + \beta_2 \cdot Colleague\ Optimism_{i,d} + \beta_3 \cdot Analyst\ Centrality_{i,t} \\ & + \vartheta \cdot Controls_{i,f,t} + \varepsilon_{i,f,t} \end{aligned} \tag{5}$$

To assess the conditional effect of *Analyst Centrality* on *Analyst Revision*, we include the interaction of *Analyst Centrality* with *Colleague Optimism* as the key variable. While analysts may unconditionally revise their forecasts downwards after their colleagues' forecasts are revealed to be optimistic, we expect this effect to be stronger for central analysts because they participate in inter-colleague information exchange more intensely. We include controls for analyst and brokerage characteristics, and the stock performance of the firm leading up to the forecast revision. Standard errors are double-clustered by year-week and firm to capture dependence in analysts' revisions.

[Insert Table 6]

Table 6 shows that central analysts are more likely to re-adjust their forecasts in response to their colleagues' *ex-post* forecasting mistakes. Column (1) shows that *Colleague Optimism* predicts more negative forecast revisions. This suggests that, unconditionally, analysts' forecast revisions tend to be more negative in response to a higher incidence of optimistic errors made by their colleagues. In Column (2), the interaction between *Analyst Centrality* and *Colleague Optimism* is significantly negative. This implies that central analysts issue more negative revisions in response to their colleagues' revealed optimism. This suggests that central analysts may have previously incorporated more of their colleagues' information in their forecasts and now unravel erroneous information.

An alternative interpretation is that when an industry shock occurs, central analysts are too busy to update their forecast due to the complexity of their coverage portfolio. Their revisions would then appear to lag their colleagues' revisions. However, given that the revisions of central analysts generate larger market reactions (see Table 3), they are more likely to contain novel information. If the revisions of central analysts are delayed after the news shock, it is unlikely that those revisions would generate significantly larger market reactions.

The results may also reflect a central analyst's superior ability to process public information because *ex-post* forecast errors are publicly available information. To rule out the information-processing hypothesis, we introduce *Global Optimism*, the global analog of *Colleague Optimism*. To construct *Global Optimism*, we collect all realized forecast errors 1) in the 30 days leading up to the analyst's revision, and 2) that are not made by the analyst's colleagues. *Global Optimism* is the proportion of optimistic forecast errors made by non-colleagues in the 30-day window.[25]

We add *Global Optimism* and its interaction with *Analyst Centrality* to our specification in Column (3). Unconditionally, analysts' revisions tend to be more negative as *Global Optimism* increases. This is unsurprising because realized forecast errors, even those made by non-colleagues, may be informative. However, the interaction between *Analyst Centrality* and *Global Optimism* no longer predicts the analyst's revision activity. This shows that central analysts do not re-adjust their forecasts incrementally to forecast errors outside their brokerages. This finding is inconsistent with the hypothesis that central analysts have superior ability to process public information. In contrast, the interaction between *Analyst Centrality* and *Colleague Optimism* remains negative and statistically significant, supporting the learning hypothesis.

3.3. Do all analysts within a brokerage benefit from information exchange?

An interesting question is whether gains from learning also accrue to periphery analysts. Both central and periphery analysts may collectively benefit via two-way feedback in a brokerage. To examine the collective benefit from learning, we create a *Network Density* measure (Newman, Watts, and Strogatz, 2002).[26] A network is dense (sparse) if its nodes are strongly (weakly) interconnected among one another, thus allowing for higher

---

[25] We provide additional evidence in Internet Appendix Table 2 that our results hold when we increase the window length from 30 days to 60 days in the constructions of *Colleague Optimism* and *Global Optimism*.
[26] Appendix I presents the technicalities of *Network Density* and working examples.

information exchange and flow (Smith-Doerr and Powell, 2010; Gibbert and Durand, 2009). Higher *Network Density* is likely to improve analyst performance as the rate of information exchange increases among analysts.

For each analyst-year, we construct *Outperformance* (%) as the proportion of her forecasts with realized forecast errors that are lower than their firm-year averages.[27] Analysts with higher values of *Outperformance (%)* are more accurate than their competitors from other brokerages. We estimate a Tobit model following equation (6) and include controls for analyst experience and portfolio complexity. By construction, *Network Density* is normalized by network size but we also explicitly control for brokerage size. We cluster standard errors at the brokerage-year level.

$$Outperformance\ (\%)_{i,t} = a + \beta_1 \cdot Analyst\ Centrality_{\,i,t\text{-}1} + \beta_2 \cdot Network\ Density_{\,i,t\text{-}1} + \varepsilon_{\,i,t} \quad (6)$$

[Insert Table 7]

The evidence in Table 7 suggests that both central and periphery analysts benefit from inter-colleague information exchange. In Column (1), we verify that our baseline findings hold at the analyst-year level. We find a positive and significant association between *Analyst Centrality* and *Outperformance (%)*. Consistent with our earlier findings, central analysts have higher forecast accuracy than their periphery colleagues. We include *Network Density* to the specification in Column (2) and find that analysts who reside in denser brokerage networks are more accurate than competing analysts in sparser brokerage networks. This suggests that analysts benefit from a network structure that promotes inter-colleague information exchange for both central and periphery analysts. The positive loading on *Analyst Centrality* in Column (2) suggests that, while analysts benefit collectively from information exchange, the gains to central analysts are greater than those to their peripheral colleagues.

[27] This construction helps to account for firm-specific heterogeneity of forecast difficulty across analysts in a brokerage. In untabulated results, our findings are qualitatively and quantitatively similar when we measure an analyst's performance as her median *Normalized Forecast Error* in the year.

By virtue of our network construction methodology, brokerages that cover only one or fewer 2-digit GICS sectors will have mechanically high values of *Network Density*. To alleviate this concern, we repeat our analysis in Column (3) on a reduced sample of brokerages that cover at least three 2-digit GICS sectors.[28] We find that our results in Column (3) remain quantitatively and qualitatively similar.

Overall, we show that brokerage structures that facilitate inter-colleague information exchange benefit not only central analysts but also their periphery colleagues. However, central analysts benefit disproportionately more from learning in brokerage networks.

4. Causal effects of analyst centrality on forecast accuracy

An unaddressed issue is the matching process between the analyst and the brokerage house. Concerns relating to brokerage characteristics discussed above are partially addressed in our main tests with the inclusion of brokerage-year fixed effects in our baseline specification. However, it remains that an analyst with expertise in a particular industry may be more likely matched with a brokerage that specializes in that industry. To establish a causal relation between *Analyst Centrality* and forecast accuracy, we exploit brokerage mergers (Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012) from years 2000 to 2007 as exogenous shocks to *Analyst Centrality* scores.[29] For each merger event, we track all analysts who work at the same acquirer brokerages pre-event and post-event. We further require that each analyst covers the same firm before and after the merger. Therefore, our unit of observation in this quasi-natural experiment is an analyst-firm. Derrien and Kecskes (2013) find that most brokerage mergers are usually motivated by business reasons, suggesting that analysts who leave post-merger, if any, are not systematically different in forecast accuracy. Further, we exclude closures because it is plausible that only high-ability analysts manage to find new employment post-closure. Restricting our analysis to analysts in brokerage mergers alleviates concerns of selection

[28] Our results are quantitatively and qualitatively similar when we adopt different values for this arbitrary cutoff.

[29] In Appendix III, we document the list of 17 brokerage mergers that we are able to match to our final sample.

bias due to unobserved analyst ability. Essentially, the shocks to *Analyst Centrality* of analysts emanate from changes in their within-brokerage network structure due to the mergers.

The treatments in this test are *Analyst Centrality Up* and *Analyst Centrality Down.* *Analyst Centrality Up* is an indicator that equal to one if the analyst's average post-merger *Analyst Centrality* is higher than her average pre-merger value, and to zero otherwise. *Analyst Centrality Down* is defined symmetrically. Since brokerage mergers are scattered temporally, we use a difference-in-difference model, generalized to accommodate multiple treatment groups, and multiple shocks across time, following Autor (2003). We estimate the specification in (7).

$$Normalized\,Forecast\,Error_{g,t} = \gamma_g + \tau_t + \sum_{j=-m,j\neq 0}^{+n} \beta_j \cdot D_{g,t}\left(t = k + j\right) + \vartheta \cdot \delta_{g,t} + \varepsilon_{g,t} \quad (7)$$

where γ_g represents the group (analyst-firm) fixed effects and τ_t represents the year fixed effects. k is the time at which the brokerage merger occurs, the term $D_{g,t}$ is an indicator which switches to one in year t if the group receives the treatment. Note that this generalized model allows k to vary in different g. This is important because brokerage mergers in our sample occur at various points in time. We let $j \neq 0$ because we skip the year of the brokerage merger. Visual inspection of the parallel trend assumption is tenuous in a model with shocks spread across time. Therefore, we include temporal leads and lags of the treatment in the model to test the assumption econometrically. Building m leads and n lags of the treatment effect β_j into the model allows us to estimate the pre-treatment dynamics (m leads) and post-treatment dynamics (n lags). The parallel trend assumption is fulfilled if β_j is not statistically significant for $j < 0$ – this suggests the absence of anticipatory effects of the treatment.

We use a 6-year window centered on the brokerage merger event. We choose *Normalized Forecast Error* as the dependent variable because it is a normalized measure that allows for comparison of forecast accuracy across different analyst-firms. The key independent variables are the three temporal leads (*Pre-Treatment*) and three temporal lags (*Post-Treatment*) of the treatment. The m^{th} temporal *Pre-Treatment* is an indicator that equates to unity only in the m^{th} year before the brokerage merger and only if the analyst-firm is treated, and zero otherwise. Similarly, the n^{th} temporal *Post-Treatment* is an indicator that equates to unity only in the n^{th} year and only if the analyst-firm is treated, and zero otherwise. Apart from year dummies and analyst-firm dummies, we also add analyst time trends to help control for confounding heterogeneity.

[Insert Table 8]

We present results from the generalized difference-in-difference model in Table 8. In Column (1), we find that the coefficients of the *Pre-Treatment* indicators are statistically insignificant. This implies that before employment shocks, analysts who experienced *Analyst Centrality Up* display no systematic differences in forecast accuracy compared to analysts whose centrality scores did not improve. Econometrically, anticipatory effects of the treatment are unlikely to be the spurious driver of our results – this helps us to validate the parallel trend assumption. Crucially, this suggests that the assignment of analysts, into more central or less central network positions subsequent to the brokerage mergers, is independent on their pre-shock performance. On the other hand, we find that the positive effects of higher *Analyst Centrality* on forecast accuracy occur only after treatment has been administered. The weak statistical significance of the *t+1 Post-Treatment* indicator suggest the presence of post-merger adjustment costs which delay the gains to learning.

Switching the treatment to *Analyst Centrality Down* in Column (2) yields symmetrically similar conclusions. In summary, our findings suggest that analysts who experience an increment in *Analyst Centrality* display higher forecast accuracy than those who did not. Moreover, we show that the differences in forecast accuracy manifest largely in the post-treatment period, and not in the pre-treatment period.

5. Discussion and robustness tests

In this section, we discuss and analyze alternative explanations to the learning hypothesis. We also perform a series of tests to ensure that our results are robust.

5.1. Does individual ability or talent explain our findings?

Analyst Centrality is likely to capture a component of talent or ability as a talented analyst may develop high centrality after surviving intense labor market competition. High-ability analysts are also likely to acquire greater coverage responsibilities over time. Therefore, our main tests include control for measures of analyst ability (i.e., experience and lowballing frequency).

To further disentangle the learning hypothesis from the ability explanation, we proxy for another dimension of ability with membership in the All-American Research Team (e.g., Leone and Wu, 2007). Studies show that membership in the Institutional Investor All-American Research Team (All-American) reflects analyst ability.[30] We re-estimate the forecast accuracy and market reaction tests presented in Tables 2 and 3. Due to our limited data on All-American Research Team membership, we end the sample period in 2008 for these additional tests.

[Insert Table 9]

Table 9 shows that our main results are unchanged after controlling for All-American status. In Column (1) of Panel A, we first show that our main centrality results hold in this subsample period. Column (2) includes the *All-American* indicator. The coefficient estimate on *Analyst Centrality* remains significantly negative and the economic effect is not statistically distinguishable across Columns (1) and (2). Consistent with prior literature, we also find that All-American analysts are more accurate. The market reaction tests in Panel B yield similar inferences. We first show in Column (1) that during this

[30] All-American analysts produces more accurate forecasts (Stickel, 1992), elicit stronger market reactions around their forecasts (Gleason and Lee, 2003), exhibit performance persistence (Leone and Wu, 2007), and attract more investment banking deal flows (Clarke et al., 2007)

sub-sample, central analysts command larger market reactions around their forecast revisions. Column (2) shows our main conclusions are unchanged with the inclusion of the *All-American* indicator.

Together, the results suggest that the superior performance of central analyst is unlikely to be completely driven by greater individual ability or talent alone. Of course, it is plausible that talented and highly-skilled analysts acquire higher centrality over time. For instance, the management may design a brokerage network around high-ability analysts to leverage on their expertise. What distinguishes our learning hypothesis is that *Analyst Centrality* has incremental explanatory power beyond measures of ability found in the prior literature.

5.2. Regulation Fair Disclosure

Regulation Fair Disclosure (Reg. FD) eliminated the practice of selective disclosure of information. As access to non-public communications with firm managers was traditionally an important information acquisition channel of analysts, the implementation of Reg. FD posed a paradigmatic shift in analysts' operations. Mohanram and Sunder (2006) argues that, in the post-Reg. FD regime, analysts may heighten information discovery efforts to offset the impact of reduced access to firm management. Following this argument, insights from inter-colleague information exchange may fill the void that emerged after Reg. FD.

[Insert Table 10]

Table 10 shows how the effect of *Analyst Centrality* changes with the implementation of Reg. FD. In Column (1), we find that, controlling for brokerage-year fixed effects, central analysts do not exhibit higher forecast accuracy than their periphery colleagues in the pre-Reg. FD period. In contrast, Column (2) shows that *Analyst Centrality* has a positive and statistically significant impact on forecast accuracy after the implementation of Reg. FD. Economically, the effect of *Analyst Centrality* in Column (2) is about 17 times the size that in Column (1).

We repeat the split-sample analysis on market reactions to analysts' forecast revisions. Following our empirical setup in Table 3, we also exclude forecast revisions if they coincide with issuances of Form 8-Ks or earnings announcements for this test. Columns (3) and (4) show that the forecast revisions of central analysts command higher market reactions both before and after the implementation of Reg. FD. Furthermore, the effect of *Analyst Centrality* on market reactions, conditional on deviations from the consensus forecast is also higher in both regimes. However, we do not find statistically significant pre- and post-difference in the effect of *Analyst Centrality*.

5.3. Are only direct connections important?

Analysts typically issue forecast revisions upon receiving value-relevant information. If information sharing occurs within a brokerage, this information shock is likely to diffuse to other colleagues. Such information may even flow when colleagues are more than 'one-step' away. As information propagates through the network, revision-activity may diffuse like a wave through the network. The information propagation may weaken with distance because 1) the information shocks become less value-relevant further away from the origin and 2) analysts have limited information-processing capacities and may not 'pass on' the information with fidelity.

Following Ahern and Harford (2014), we construct a measure called *Closeness-weighted Revision Activity* that captures these waves of revision activity. For each analyst-brokerage pair, we find the earliest and latest forecast announcement dates. In between these two dates, we divide the analyst's tenure at the brokerage by months. We exclude the first and last months of her tenure at the brokerage from our sample to avoid truncation issues. In each month, we construct *Revision Count* as the total number of forecast revisions made by the analyst. Next, we compute *Closeness-weighted Revision Activity* of the analyst as the weighted sum of all colleagues' *Revision Count* in the previous month. We use the reciprocal of the shortest-path distance[31] between an analyst i and a given colleague j in brokerage G as the weight.

$$Closeness\text{-}weighted\ Revision\ Activity_{i,t} = \sum\nolimits_{j \in G, j \neq i} \frac{1}{Distance_{i,j,t}} \times \sum\nolimits_{r} Revision_{r,j,t-1} \tag{8}$$

Since the dependent variable, *Revision Count*, is a discrete count variable, we estimate negative binomial regressions following equation (9). We include controls for analyst characteristics and brokerage size.

$$Revision\ Count_{i,t} = \alpha + \beta_1 \cdot Closeness\text{-}weighted\ Revision\ Activity_{i,t} + \vartheta\ Controls_{i,t} + \varepsilon_{i,t} \tag{9}$$

Table 3 of the Internet Appendix shows that revision activity diffuses within the brokerage network. Column (1) shows a significantly positive relation between *Closeness-weighted Revision Activity* and *Revision Count*. This indicates that the analyst's own revision activity increases in response to colleagues' intense revision activity. Since the sum of colleagues' revision activity mechanically increases with the number of analysts employed at the brokerage, we adopt a size-normalized measure of *Closeness-weighted Revision Activity.* Specifically, we scale the revision activity at each distance tier by the number of colleagues residing in that tier before applying the above distance-weighted aggregation. Additionally, as the number of revisions made by an analyst is expected to increase with the number of firms under her coverage, we include *Firm Breadth* as an exposure variable in the model. Our conclusion remains unchanged when we adopt this alternative measure in Column (2). Economically, a unit increase in size-normalized *Closeness-weighted Revision Activity* is associated with a 1.5% increase in the analyst's *Revision Count*.

In Columns (3) and (4), we show that analysts are influenced more strongly by the revision activity of proximate (shorter paths) colleagues in the network. We define *Distance1* (*Distance2*) *Revision Activity* of an analyst as the total *Revision Count* of all colleagues who are at one path-length (two path-lengths) away in the network, scaled by

[31] Returning to the simplistic triad of analysts A, B, and C in the above example, analyst A has shortest-path distances of one and two to analysts B and C respectively.

the number of colleagues residing at one path-length (two path-lengths). Comparing their estimated coefficients in Columns (3) and (4), *Distance1 Revision Activity* has an economic magnitude approximately two times larger than that of *Distance2 Revision Activity.*[32] The Welch-Satterthwaite t-test indicates that this difference in economic magnitudes is statistically significant at the 1% level. Column (3) shows that an analyst has a 3.2% increase in his revision activity for an additional forecast revision made by a colleague at one path-length away. In Column (4), this effect shrinks to 1.1% for forecast revisions made by colleagues at two path-lengths away. This suggests that the diffusion of revision activity degrades with distance.

5.4. Do central or peripheral analysts drive our findings?

We hypothesize that central analysts are better positioned to benefit from information exchange in brokerages. However, it is plausible that only the most peripheral analysts persistently underperform but the other (more central) analysts' forecast accuracy do not increase with *Analyst Centrality.* On average, this pattern will still generate a positive relation between *Analyst Centrality* and forecast accuracy. Crucially, this pattern is consistent with an alternative hypothesis that *Analyst Centrality* spuriously correlates with the amount of brokerage resources allocated to analysts; the most peripheral analysts underperform relatively because brokerages allocate less research support to their coverage portfolios. In contrast, the relation between *Analyst Centrality* and forecast accuracy applies for both peripheral and central analysts under the colleague-learning hypothesis.

To test whether our findings apply generally to all analysts, we introduce a tiered-decomposition of *Analyst Centrality.* Specifically, we assign *High* (*Low*) *Centrality* indicators to analysts in the top (bottom) tercile of *Analyst Centrality* in every year. Thereafter, we repeat our baseline analysis with these indicators in lieu of *Analyst*

[32] We refrain from including both *Distance1 Revision Activity* and *Distance2 Revision Activity* in the same specification because they are highly collinear with each other. We only present results up to a path-length of two because analysts in many brokerages are connected to any colleague by two path-lengths.

Centrality as key variables. This empirical design allows us to benchmark the performance of the most peripheral and most central analysts with their moderately central colleagues.

The results in Table 4 of the Internet Appendix indicate that the relation between *Analyst Centrality* and forecast accuracy applies generally to both peripheral and central analysts. In Column (1), we estimate the effect of *Low Centrality* on forecast accuracy in the absence of *High Centrality.* Thus, the benchmark category in Column (1) comprises analysts who are either moderately central or highly central. We find that *Low Centrality* loads positively and significantly on forecast error. In line with our expectation, peripheral analysts exhibit lower accuracy than their more central colleagues do. Following this, we replace *Low Centrality* with *High Centrality* in Column (2). Consistent with our prior findings that analysts who are more central have higher forecast accuracy, we find a negative and statistically significant loading on *High Centrality.* In Column (3), we estimate the effects of *Low Centrality* and *High Centrality* on forecast error simultaneously. With moderately central analysts as the benchmark category, we expect *High Centrality* to load insignificantly on forecast error if our prior findings are driven exclusively by the most peripheral analysts. Our results indicate otherwise; *Low Centrality* and *High Centrality* load positively and negatively on forecast error when they are jointly included in the regression specification. In Columns (4) through (6), we repeat the preceding analysis but *Low* (*High*) *Centrality* corresponds to the bottom (top) quartile of *Analyst Centrality* instead. Our conclusions remain unchanged from Columns (1) through (3).

Overall, our evidence does not support the alternative hypothesis that the effect of *Analyst Centrality* is exclusively driven by a relative underperformance of the most peripheral analysts in brokerages. Our evidence is also inconsistent with the notion that brokerages allocate fewer resources to analysts with low *Analyst Centrality.*

5.5. Robustness: Alternative industry classification schemes in network construction

The broadness of the 2-digit GICS industry classification scheme may create imprecision and generate noise in our measure of analysts' propensities to learn from their colleagues.

However, noise tends to reduce the fidelity of the network and is likely to bias against uncovering evidence of inter-colleague learning. Alternatively, we could adopt a more granular industry classification scheme in our network construction but this approach risks the possibility of failing to register inter-colleague connections that are in reality present. Therefore, the optimal choice of industry classification schemes reflects a trade-off between being too broad or too narrow.

To show that our baseline findings are not due to our choice of industry classification scheme, we use a more granular industry classification scheme – the 2-digit SIC scheme – in our network construction methodology.

Table 5 of the Internet Appendix reports results on the relations between *Analyst Centrality (SIC-2D)* and 1) forecast accuracy, 2) market reactions around forecast revisions. Control variables are included per the model specifications in Tables 2 and 3 but their estimated coefficients are not tabulated for brevity. The definition of *Analyst Centrality (SIC-2D)* remains unchanged from Sections 1.2 and 1.3, except that it is computed on within-brokerage networks in which two analysts are connected if they cover at least one 2-digit SIC industry in common. Our conclusions remain unchanged with the adoption of *Analyst Centrality (SIC-2D)* in our tests. We continue to find that, controlling for brokerage-year fixed effects, more central analysts exhibit higher forecast accuracy. Market reactions around forecast revisions are also increasing in *Analyst Centrality (SIC-2D)*, both conditional and unconditional on analysts' deviations from the prevailing consensus. The economic magnitudes on the effect of analyst centrality are also comparable to those in Tables 2 and 3. Overall, we show that our main findings are not sensitive to the granularity of industry classification schemes adopted in our network construction methodology.[33]

[33] However, when compared to Tables 2 and 3, our sample sizes in Table 11 are marginally smaller with the use of 2-digit SIC in our network construction methodology. The increased granularity of the industry classification scheme causes some brokerages in our sample to have no within-brokerage connections among their analysts. Since *Analyst Centrality* is undefined for all analysts in these brokerages, they are omitted from our sample.

6. Conclusion

In knowledge-based professions, the quality and ability of one's co-workers can have a significant impact on an employee's productivity. Our study focuses specifically on the effects of learning from one's colleagues by examining financial analysts. This setting is particularly well-suited because information is a key input to an analyst's performance.

Using GICS sector overlaps in their coverage portfolios, we build an information network within a brokerage house to measure potential information flow among colleagues. Our evidence suggests that analysts who are more centrally located in their brokerage networks produce higher quality equity research. We provide more direct evidence that high centrality analysts are more likely to learn from their colleagues. If high centrality analysts are tapping into their colleagues' knowledge and expertise, we expect that they would be more likely to include such information in their forecast revisions. Consistent with this argument, we find that high centrality analysts are more likely to revise their forecasts after their colleagues' forecasting mistakes are known. Controlling for analyst membership in the Institutional Investor All-America Research Team, supplementary analysis suggests that the positive effect of *Analyst Centrality* on superior forecasting outcomes is incremental to that of innate analyst ability.

The formation of within-brokerage networks may be endogenous. Therefore, to better understand the causal relation between within-brokerage network centrality and forecast accuracy, we exploit exogenous brokerage mergers from 2000 to 2007. These mergers introduce exogenous changes to the brokerage structures of the acquirers and targets, and hence impact the *Analyst Centrality* of their analysts. We find that analysts who experienced increases in within-brokerage network centrality improve their forecast accuracy.

REFERENCES

Ahern, K., Harford, J., 2014. The importance of industry links in merger waves. Journal of Finance 69, 527-576.

Anjos, F., Fracassi, C., 2015. Shopping for information? Diversification and the network of industries. Management Science 61, 161-183.

Autor, D., 2003. Outsourcing at Will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. Journal of Labor Economics 21, 1-42.

Bae, K-H, Stulz R., 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, 581-606.

Balakrishnan, K., Billings, M., Kelly, B., Ljungqvist, A., 2014. Shaping liquidity: On the causal effects of voluntary disclosure. Journal of Finance 69, 2237-2278.

Bhojraj, S., Lee, C., Oler, D., 2003. What's my line? A comparison of industry classification schemes for capital market research. Journal of Accounting Research 41, 745-774.

Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. Journal of Mathematical Sociology 2, 113-120.

Boni, L., Womack, K., 2006. Analysts, industries, and price momentum. Journal of Financial and Quantitative Analysis 41, 85–109.

Borgatti, S., 2005. Centrality and network flow. Social networks 27, 55-71.

Bradley, D., Clarke, J., Lee, S., Ornthanalai, C., 2014. Are analysts' recommendations informative? Intraday evidence on the impact of time stamp delays. Journal of Finance 69, 645-673.

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

Bradshaw, M., 2011. Analysts' forecasts: What do we know after decades of work? Unpublished working paper, Boston College.

Brochet, F., Miller, G., Srinivasan, S., 2013. Do analysts follow managers who switch companies? An analysis of relationships in the capital markets. The Accounting Review 89, 451-482.

Brown, L., 1997. Analyst forecasting errors: Additional evidence. Financial Analyst Journal 53, 81–88.

Brown, L., 1998. Managerial behavior and the bias in analysts' earnings forecasts. Unpublished working paper, Georgia State University.

Brown, L., 2001. A temporal analysis of earnings surprises: Profits versus losses. Journal of Accounting Research 39, 221-241.

Brown, L., Call, A., Clement, M., Sharp, N., 2015. Inside the "black box" of sell-side financial analysts. Journal of Accounting Research 53, 1-47.

Chang, X., Dasgupta, S., Hilary, G., 2006. Analyst following and financing decisions. Journal of Finance 61, 3009-3048.

Chen, Q., Francis, J., Jiang, W., 2005. Investor learning about analyst predictive ability. Journal of Accounting and Economics 39, 3-24.

Chen, Q, Jiang W., 2006. Analysts' weighting of private and public information. Review of Financial Studies 19, 319-355.

Chen, S., Matsumoto, D., 2006. Favorable versus unfavorable recommendations: The impact on analyst access to management ‐ provided information. Journal of Accounting Research 44, 657-689.

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

Clement, M., 1999. Analyst Forecast Accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics 27, 285-303.

Clement, M., Tse, S., 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? The Accounting Review 78, 227–249.

Clement, M., Tse, S., 2005. Financial analyst characteristics and herding behavior in forecasting. Journal of Finance 60, 307-341.

Clement, M., Koonce, L., Lopez, T., 2007. The roles of task-specific forecasting experience and innate ability in understanding analyst forecasting performance. Journal of Accounting and Economics 44, 378–398.

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

Cohen, L., Frazzini, A., Malloy, C., 2008. The small world of investing: Board connections and mutual fund returns. Journal of Political Economy 116, 951–979.

Cohen, L., Frazzini, A., Malloy, C., 2010. Sell-side school ties. Journal of Finance 65, 1409-1437.

Das, S., Levine, C., Sivaramakrishnan, K., 1998. Earnings predictability and bias in analysts' earnings forecasts. The Accounting Review 2, 277-294.

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

Engelberg, J., Gao, P., Parsons, C., 2012. Friends with money. Journal of Financial Economics 103, 169-188.

Fracassi, C., 2016. Corporate finance policies and social networks. Management Science forthcoming.

Francis, J., Philbrick, D., 1993. Analysts' decisions as products of a multi-task environment. Journal of Accounting Research 31, 216-230.

Freeman, L., 1979. Centrality in networks: Conceptual clarification. Social Networks 1, 215–239.

Gibbert, M., Durand, T., 2009. Strategic networks: Learning to compete. John Wiley & Sons.

Gleason, C., Lee, C., 2003. Analyst forecast revisions and market price discovery. The Accounting Review 78, 193-225.

Gow, I., Ormazabal, G., Taylor, D., 2010. Correcting for cross-sectional and time-series dependence in accounting research. The Accounting Review 85, 483-512.

Green, T., Jame, R., Markov, S., Subasi, M., 2014a. Broker-hosted investor conferences. Journal of Accounting and Economics 58, 142–166.

Green, T., Jame, R., Markov, S., Subasi, M., 2014b. Access to management and informativeness of analyst research. Journal of Financial Economics 114, 239–255.

Groysberg, B., Healy, P., Maber, D., 2011. What drives sell-side analyst compensation at high-status investment banks? Journal of Accounting Research 49, 969-1000.

Heflin, F., Subramanyam, K., Zhang, Y., 2003. Regulation FD and the financial information environment: Early evidence. The Accounting Review 78, 1-37.

Hilary, G., Hsu, C., 2013. Analyst forecast consistency, Journal of Finance 68, 271-297.

Hilary, G., Shen, R., 2013. The role of analysts in intra-industry information transfer. The Accounting Review 88, 1265-1287.

Hochberg, Y., Ljungqvist, A., Lu, Y., 2007. Whom you know matters: Venture capital networks and investment performance. Journal of Finance 62, 251-301.

Honoré, B., 1992. Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects. Econometrica 60, 533-565.

Hong, H., Kubik, J., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecast. RAND Journal of Economics 31, 121-144.

Hong, H, Kubik, J., 2003. Analyzing the analysts: Career concerns and biased forecasts. Journal of Finance 58, 313-351.

Hong, H., Kubik, J., Stein, J., 2004. Social interaction and stock-market participation. Journal of Finance 59, 137-163.

Hong, H., Kubik, J., Stein, J., 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. Journal of Finance 60, 2801-2824.

Hong, H., Kacperczyk, M., 2010. Competition and bias. Quarterly Journal of Economics 125, 1683-1725.

Hugon, A., Lin, A-P, Markov, S., 2016. Do equity analysts benefit from access to high quality debt research? Unpublished working paper, Arizona State University.

Hugon, A., Kumar, A., Lin, A-P., 2016. Analysts, macroeconomic news, and the benefit of active in-house economists. The Accounting Review 91, 513-534.

Hwang, L., Jan, C., Basu, S., 1996. Loss firms and analysts' earnings forecast errors. Journal of Financial Statement Analysis 1, 18–30.

Hwang, B-H, Kim, S., 2009. It pays to have friends. Journal of Financial Economics 93, 138-158.

Hwang, B-H, Liberti, J., Sturgess, J., 2016. Information sharing and spillovers: Evidence from financial analysts. Unpublished working paper, Northwestern University.

Irvine, P., Simko, P.J., Nathan, S., 2004. Asset management and affiliated analysts' forecast. Financial Analyst Journal May/June 67-78

Ivković, Z., Jegadeesh, N., 2004. The timing and value of forecast and recommendation revisions. Journal of Financial Economics 73, 433-463.

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

Jiang, D., Kumar, A., Law, K., 2016. Political contributions and analyst behavior. Review of Accounting Studies 21, 37-88.

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

Ke, B., Yu, Y., 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. Journal of Accounting Research 44, 965-999.

Kelly, B., Ljungqvist, A., 2012. Testing asymmetric-information asset pricing models. Review of Financial Studies 25, 1366-1413.

Kini, O, Mian, S., Rebello, M., Venkateswaran, A., 2009. On the structure of analyst research portfolios and forecast accuracy. Journal of Accounting Research 47, 867-909.

Kuhnen, C., 2009. Business networks, corporate governance, and contracting in the mutual fund industry. Journal of Finance 64, 2185-2220.

Kumar, A., 2010. Self-selection and the forecasting abilities of female equity analysts. Journal of Accounting Research, 48, 393–435

Lang, M., Lins, K., Miller, D., 2004. Concentrated control, analyst following, and valuation: Do analysts matter most when investors are protected least? Journal of Accounting Research 42, 589-623

Larcker, D., So, E., Wang, C., 2013. Boardroom centrality and firm performance. Journal of Accounting and Economics 55, 225-250.

Law, K., 2013. Career imprinting: The influence of coworkers in early career. Unpublished working paper, Nanyang Technological University.

Leone, A., Wu, J., 2007. What does it take to become a superstar? Evidence from Institutional Investor rankings of financial analysts. Unpublished working paper, University of Rochester

Li, F., Mukherjee, A., Sen, R., 2017. Inside brokers. Unpublished working paper, Hong Kong University of Science and Technology.

Lim, T., 2001. Rationality and analysts' forecast bias. Journal of Finance 56, 369-385.

Loh, R., Mian, G., 2006. Do accurate earnings forecasts facilitate superior investment recommendations? Journal of Financial Economics 80, 455–483.

Loh, R., Stulz, R., 2011. When are analyst recommendation changes influential? Review of Financial Studies 24, 593-627.

Loh, R., Stulz, R., 2017. Is sell-side research more valuable during bad times? Journal of Finance, forthcoming.

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

Marsden, P., 1990. Network data and measurement. Annual Review of Sociology, 435-463.

Matsumoto, D., 2002. Management's incentives to avoid negative earnings surprises. The Accounting Review 77, 483-514.

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. Journal of Finance, forthcoming.

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

Mikhail, M., Walther, B., Willis, R., 1999. Does forecast accuracy matter to security analysts? The Accounting Review 74, 185–200.

Mohanram, P., Sunder, S., 2006. How has Regulation FD affected the operations of financial analysts? Contemporary Accounting Research 23, 491-525.

Newman, M., 2003. The structure and function of complex networks. SIAM Review 45, 167-256.

Newman, M., Watts, D., Strogatz, S., 2002. Random graph models of social networks. Proceedings of the National Academy of Sciences 99, 2566-2572.

Ozsoylev, H., Walden, J., Yavuz, M., Bildik, R., 2014. Investor networks in the stock market. Review of Financial Studies 27, 1323-1366.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Review of Financial Studies 22, 435-480.

Pool, V., Stoffman, N., Yonker, S., 2015. The people in your neighborhood: Social interactions and mutual fund portfolios. Journal of Finance 70, 2679-2732.

Richardson, S., Teoh, S., Wysocki, P., 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. Contemporary Accounting Research 21, 885-924.

Shue, K., 2013. Executive networks and firm policies: Evidence from the random assignment of MBA peers. Review of Financial Studies 26, 1401-1442.

Smith-Doerr, L., Powell, W., 2010. The Handbook of Economic Sociology. Princeton University Press.

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

Stickel, S., 1992. Reputation and performance among security analysts. Journal of Finance 47, 1811-1836.

Valente, T., Coronges, K., Lakon, C., Costenbader, E., 2008. How correlated are network centrality measures? Connections (Tor) 28, 16-26.

Westphal, J., Clement, M., 2008. Sociopolitical dynamics in relations between top managers and security analysts: Favor rendering, reciprocity, and analyst stock recommendations. Academy of Management Journal 51, 873-897.

Womack, K., 1996. Do brokerage analysts' recommendations have investment value? Journal of Finance 51, 137-167.

Yin, H., Zhang, H., 2014. Tournaments of financial analysts. Review of Accounting Studies 19, 573-605.

Appendix I

II.A. Degree Centrality

Degree centrality is related to the number of colleagues that an analyst is immediately connected to in a brokerage network. For example, in Figure 1, analyst $\boldsymbol{A}$ – represented by the red node – is immediately connected to four colleagues. Since degree centrality is increasing in the number of analysts $\boldsymbol{N}$ in a brokerage $\boldsymbol{G}$, we normalize degree centrality by $\boldsymbol{N_G - 1}$, or the maximum possible number of direct connections an analyst can have in a network.

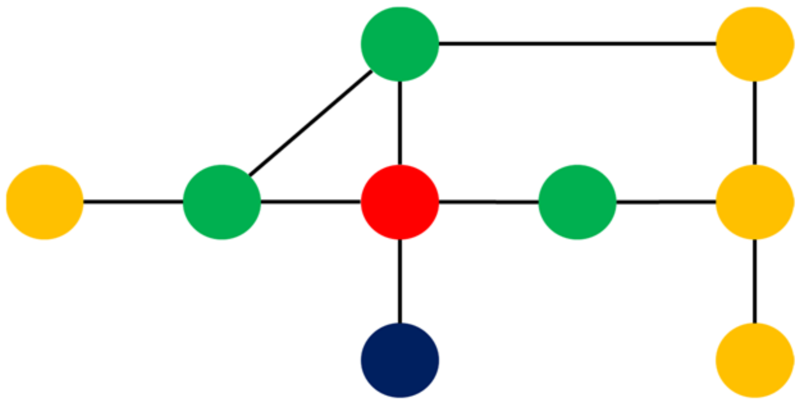

Figure 2

Degree centrality favors analysts who have relatively more opportunities. Consider the network in Figure 2. Analysts $\boldsymbol{A}$ and $\boldsymbol{Z}$ are represented as red and blue nodes respectively. If $\boldsymbol{Z}$ cannot provide resources to $\boldsymbol{A}$, $\boldsymbol{A}$ has the opportunity to ask her other 3 neighbors – represented as green nodes. However, $\boldsymbol{Z}$ does not have an alternative connection if A is unable to provide resources. Under degree centrality, $\boldsymbol{A}$ possesses more social power than $\boldsymbol{Z}$ because the former is less dependent on any single colleague. Being connected to more colleagues also increases the likelihood that $\boldsymbol{A}$ will receive any information being circulated in the network. The formal mathematical definition of degree centrality for a given analyst $\boldsymbol{i}$ is as follows.

$$Centrality_{Degree}(i,\ G) = \frac{\sum_{j \neq i, j \in G}^{N_G - 1} f(i,\ j)}{N_G - 1} \qquad N_G = Number\ of\ analysts\ in\ network\ G$$

$$f(i,\ j) = \begin{cases} +1, & i\ is\ connected\ to\ j \\ 0, & i\ is\ not\ connected\ to\ j \end{cases}$$

Following the above discussion, we show that analyst $\boldsymbol{A}$ has a higher degree centrality than analyst $\boldsymbol{Z}$.

$$Centrality_{Degree}\ of\ analyst\ \boldsymbol{A} = \frac{(3 \times 1_{RedGreen}) + (1\ \times 1_{RedBlue}) + (3 \times 0_{RedYellow})}{9-1} = 0.500$$

$$Centrality_{Degree}\ of\ analyst\ \boldsymbol{Z} = \frac{(1 \times 1_{BlueRed}) + (0\ \times 1_{BlueGreen}) + (0 \times 0_{BlueYellow})}{9-1} = 0.125$$

II.B. Closeness Centrality

Closeness centrality is related to the distances between an analyst and all her colleagues (both immediately or not immediately connected) in a brokerage network. For example, in Figure 3a, the closeness centrality of analyst $\boldsymbol{A}$ – represented by the red node – is the reciprocal of the sum of its shortest-path distances to all other brokerage colleagues. Since the sum of shortest-path distances is increasing in the number of analysts, closeness centrality is normalized by the minimum possible sum of shortest-path distances, $\boldsymbol{N_G - 1}$. For an analyst whose normalized closeness centrality equates to unity, all her colleagues are immediately connected to her.

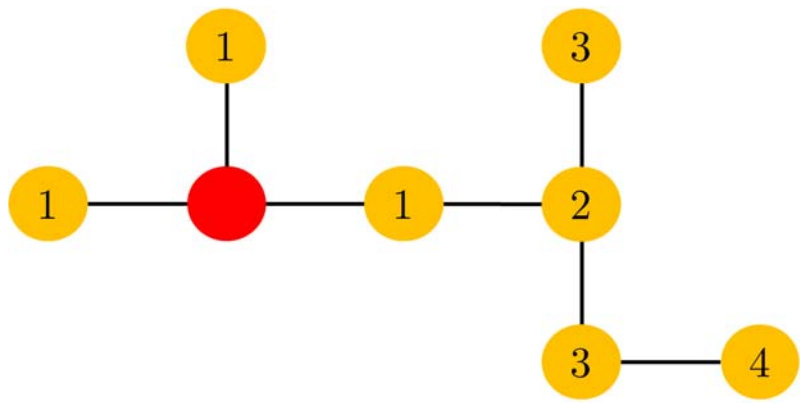

Figure 3a

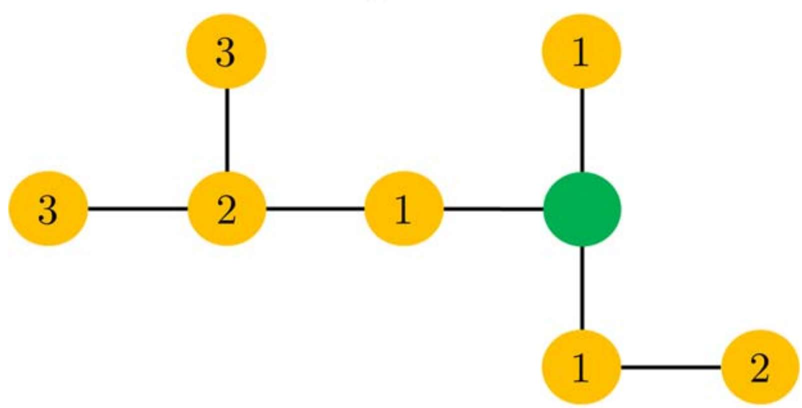

Figure 3b

Closeness centrality favors analysts who can access their colleagues, or are reachable by their colleagues, at relatively shorter path lengths. An advantage of closeness centrality over degree centrality is that the former can also account for indirect connections in the network. This advantage is salient if there are isolated or disconnected components (cluster of nodes) in the network. When such components are present, closeness centrality, unlike degree centrality, can differentiate global (network-wide) centrality from local centrality. To illustrate the difference between closeness and degree centralities, consider a brokerage network $\boldsymbol{G}$ in which analysts $\boldsymbol{A}$ and $\boldsymbol{Z}$ are represented as a red node in Figure 3a and a green node in Figure 3b respectively. The shortest-path distance of each colleague from analysts $\boldsymbol{A}$ and $\boldsymbol{Z}$ is indicated in the orange nodes. We show below that analyst $\boldsymbol{Z}$ has a higher closeness centrality than analyst $\boldsymbol{A}$ even though both analysts have the same values of degree centrality. The formal mathematical definition of closeness centrality for a given analyst $\boldsymbol{i}$ is as follows.

$$Centrality_{Closeness}(i,\ G) = \frac{N_G - 1}{\sum_{j \neq i, j \in G}^{N_G - 1} d(i,j)} \qquad N_G = Number\ of\ analysts\ in\ network\ G \qquad d(i,j) = Shortest\ path\ distance\ between\ i\ and\ j$$

We now show that analysts $\boldsymbol{A}$ and $\boldsymbol{Z}$ have different values of closeness centrality despite having the same values of degree centrality.

$$Centrality_{Closeness}\ of\ analyst\ \boldsymbol{A} = \frac{8-1}{1+1+1+2+3+3+4} = 0.467$$

$$Centrality_{Closeness}\ of\ analyst\ \boldsymbol{Z} = \frac{8-1}{1+1+1+2+2+3+3} = 0.538$$

II.C. Betweenness Centrality

Betweenness centrality is related to the number of geodesics (shortest paths) in the brokerage network that pass through an analyst.

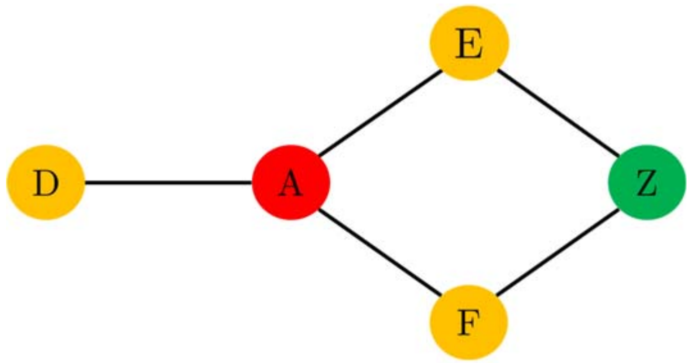

Figure 4

We first elaborate on the definition of geodesics. For a pair of analysts, a geodesic between them is a path of the shortest possible length. An analyst pair may have more than 1 geodesic. Consider analysts $\boldsymbol{E}$ and $\boldsymbol{F}$. Analyst $\boldsymbol{E}$ may reach analyst $\boldsymbol{F}$ via 2 paths – $\boldsymbol{EZF}$ and $\boldsymbol{EAF}$. Since both paths have lengths of 2, and the shortest possible path length between the analyst pair is 2, both $\boldsymbol{EZF}$ and $\boldsymbol{EAF}$ qualify as geodesics.

The betweenness centrality of analyst $\boldsymbol{A}$ is the sum of proportions of all geodesics (not involving $\boldsymbol{A}$) which pass through $\boldsymbol{A}$. Revisiting the example of analyst pair $\boldsymbol{E}$ and $\mathbf{F}$, there are 2 geodesics $\boldsymbol{EZF}$ and $\boldsymbol{EAF}$ between them but only path $\boldsymbol{EAF}$ passes through analyst $\boldsymbol{A}$, yielding a proportion of 0.5. If we repeat this computation for all possible analyst pairs with reference to analyst $\boldsymbol{A}$, we will obtain her betweenness centrality. Since these sums of proportions are increasing in the number of analysts, betweenness centrality is normalized by $\frac{2}{(N_G-1)(N_G-2)}$ the number of unique analyst pairs not involving $\boldsymbol{A}$.

Betweenness centrality favors analysts who have brokering capacity. In Figure 4, analyst $\boldsymbol{A}$ is in an advantageous brokering position relative to analysts $\boldsymbol{D}$ and $\boldsymbol{E}$. Should $\boldsymbol{D}$ and $\boldsymbol{E}$ choose to interact with each other, they must do so via $\boldsymbol{A}$. In contrast, if $\boldsymbol{A}$ chooses to interact with either $\boldsymbol{D}$ or $\boldsymbol{E}$, he may do so without the need to pass through any colleagues.

The formal mathematical definition of betweenness centrality for a given analyst $\boldsymbol{i}$ is as follows.

$$Centrality_{Betweenness}(i,\ G) = \frac{2}{(N_G - 1)(N_G - 2)} \sum\nolimits_{x \neq i, y \neq i, x, y \in G} \frac{s(x, y|i)}{s(x, y)}$$

$N_G = Number\ of\ analysts\ in\ network\ G$

$s(x, y) = Number\ of\ geodesics\ between\ any\ pair\ x\ and\ y$

$s(x, y|i) = Number\ of\ geodesics\ between\ any\ pair\ x\ and\ y\ passing\ through\ i$

We now show that analyst $\boldsymbol{A}$ has a higher betweenness centrality than analyst $\boldsymbol{Z}$. The identities of analyst pairs are denoted by subscripts in the denominators of the fractions.

$$Centrality_{Betweenness}\ of\ analyst\ \boldsymbol{A} = \frac{1}{6}\left(\frac{1}{1_{DE}} + \frac{1}{1_{DF}} + \frac{2}{2_{DZ}} + \frac{1}{2_{EF}} + \frac{0}{1_{EZ}} + \frac{0}{1_{FZ}}\right) = 0.583$$

$$Centrality_{Betweenness}\ of\ analyst\ \boldsymbol{Z} = \frac{1}{6}\left(\frac{0}{1_{DA}} + \frac{0}{1_{DE}} + \frac{0}{1_{DF}} + \frac{0}{1_{AE}} + \frac{0}{1_{AF}} + \frac{1}{2_{EF}}\right) = 0.083$$

II.D. Eigenvector Centrality

Eigenvector centrality is related to the notion that the centrality of an analyst is high if her connected colleagues are highly central in the brokerage network. A notable application of eigenvector centrality is the PageRank algorithm used by the Google search engine to determine the importance of websites on the Internet. The underlying logic of the algorithm is that important websites are more likely to receive more web-links from other important websites. Similarly, an analyst has a high eigenvector centrality if her connected colleagues also possess high eigenvector centralities. Under eigenvector centrality, not only does the *quantity* of connections determine one's prominence in the network, but the *quality* of those connections matters as well.

To motivate the mathematical intuition behind eigenvector centrality, consider a simple network structure in Figure 5 and its corresponding adjacency matrix $\mathbf{M_G}$ below. The row-wise and column-wise sequences of the elements follow $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{S}$. Where there is a connection, the element values equate to unity, and equate to zero otherwise. For example, element $\mathbf{m_{1,2}}$ equates to unity because $\boldsymbol{P}$ and $\boldsymbol{Q}$ are connected. On the other hand, element $\mathbf{m_{1,3}}$ equates to zero because there is no connection between $\boldsymbol{P}$ and $\boldsymbol{R}$.

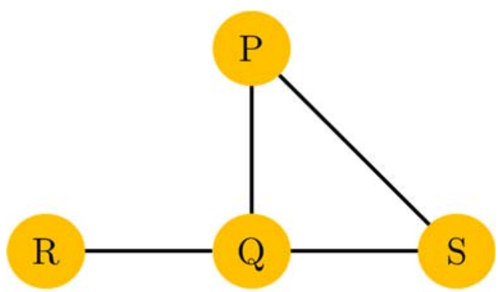

Figure 5

$$\mathbf{M_G} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Next, suppose there is a 4 by 1 vector $\mathbf{K_G}$ of centrality values. For the purpose of exposition, we begin by choosing $\mathbf{K_G}$ to be a vector of un-normalized degree centralities. We arbitrarily choose un-normalized degree centralities as a starting point for its simplicity. For all purposes of this exposition, we could have defined $\mathbf{K_G}$ to be a vector of any other centrality values.

Formally, we write $\mathbf{K_G}$ as follows.

$$\mathbf{K_G} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 2 \end{bmatrix}$$

$$where\ \mathbf{k_{1,1}}, \mathbf{k_{2,1}}, \mathbf{k_{3,1}}, and\ \mathbf{k_{4,1}}\ are\ the\ unnormalized\ degree\ centralities\ of\ \boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}, and\ \boldsymbol{S}\ respectively$$

Now we perform the below matrix multiplication, we obtain another 4 by 1 matrix.

$$\mathbf{M_G} \cdot \mathbf{K_G} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \times 2 + 1 \times 3 + 0 \times 1 + 1 \times 2 \\ 1 \times 2 + 0 \times 3 + 1 \times 1 + 1 \times 2 \\ 0 \times 2 + 1 \times 3 + 0 \times 1 + 0 \times 2 \\ 1 \times 2 + 1 \times 3 + 0 \times 1 + 0 \times 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 3 \\ 5 \end{bmatrix}$$

For each analyst in the network, the matrix multiplication sums up only the centralities of colleagues whom he is directly connected to. Alternatively, this multiplication is not only causing each analyst to receive her connected colleagues' centralities, but also causing her to distribute her centrality to connected colleagues concurrently. From the above example, the element [1, 1], corresponding to $\boldsymbol{P}$, of the resulting matrix carries a value of 5, the cumulative centrality of her connections. This value is derived from $\boldsymbol{P}$'s connections – $\boldsymbol{Q}$ and $\boldsymbol{S}$ – who have degree centralities of 3 and 2 respectively. Now, we can interpret this matrix multiplication as 'spreading' the initial vector $\mathbf{K_G}$ across the network.

Suppose we repeat the multiplication to spread the initial vector $\mathbf{K_G}$ further, we obtain more 4 by 1 matrices.

$$\boldsymbol{M_G} \cdot \boldsymbol{M_G} \cdot \boldsymbol{K_G} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 5 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 13 \\ 5 \\ 10 \end{bmatrix} \quad \mathbf{M_G} \cdot \mathbf{M_G} \cdot \mathbf{M_G} \cdot \mathbf{K_G} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ 13 \\ 5 \\ 10 \end{bmatrix} = \begin{bmatrix} 23 \\ 25 \\ 13 \\ 23 \end{bmatrix}$$

While we observe that the values of elements in the matrices continue to become larger, one may speculate that there may exist an equilibrium where the proportion of total centralities held by each analyst is constant remains constant through additional stages of multiplication. At such an equilibrium, the centrality value of each analyst fully reflect the centralities of all connected colleagues. We can search for this equilibrium by choosing the initial vector $\mathbf{K_G}$. Upon closer inspection, the search for this equilibrium solution is in fact a search for the eigenvector of the adjacency matrix $\mathbf{M_G}$.

If we had replaced each element of the centrality vector $\mathbf{K_G}$ with values of eigenvector centralities, the brokerage network can be described as follows.

$$Centrality_{Eigenvector}(i,\ G) = element\ v_i\ of\ matrix\ \mathbf{V_G}$$

$$\mathbf{V_G}\ is\ an\ eigenvector\ of\ the\ network's\ adjacency\ matrix\ \mathbf{M_G}$$

$$or, \qquad \mathbf{M_G} \cdot \mathbf{V_G} = \boldsymbol{\lambda} \cdot \mathbf{V_G}$$

$$where\ \ \boldsymbol{\lambda}\ \text{is a scalar}$$

Suppose we perform a matrix multiplication between $\mathbf{M_G}$ and its eigenvector $\mathbf{V_G}$.

$$\mathbf{M_G} \cdot \mathbf{V_G} = \boldsymbol{\lambda} \cdot \mathbf{V_G}$$

And multiply the resulting vector with $\mathbf{M_G}$.

$$\mathbf{M_G} \cdot \boldsymbol{\lambda} \cdot \mathbf{V_G} = \boldsymbol{\lambda} \cdot \mathbf{M_G} \cdot \mathbf{V_G} = \boldsymbol{\lambda} \cdot \boldsymbol{\lambda} \cdot \mathbf{V_G} = constant \cdot \mathbf{V_G}$$

We observe that even with incremental steps of matrix multiplication, the resulting vector is always a scalar inflation of the starting vector $\mathbf{V_G}$. Thus, we can say that the vector $\mathbf{V_G}$ fully represents the cumulative centrality (or prominence) of analysts and their colleagues in the brokerage network.

II.E. Network Density

Network density is a measure that describes the overall interconnectedness of analysts in a brokerage network. Distinct from centrality measures which are analyst-centric, network density is a brokerage-level measure that relates to the network structure of the brokerage. Analysts in a brokerage with high network density would be strongly interconnected to one another. To define interconnectedness among analysts, we rely on the concept of transitivity (Newman, Watts, and Strogatz, 2002). The concept of transitivity readily maps to our intuition of interconnectedness when we consider the colloquialism – "we are interconnected if the friend[34] of my friend also happens to be my friend". Notice that the colloquialism elicits an image of a triangle with a node at each of its 3 vertices. This idea is at the heart of defining network transitivity[35] (or network density).

Consider a simple network $\boldsymbol{P}$ in Figure 6.

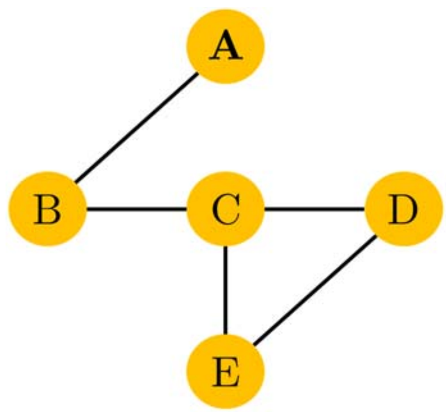

Figure 6

We begin by counting the number of paths of length = 2. In total, there are 12 paths of length = 2: one originating from $\boldsymbol{A}$ ($\boldsymbol{ABC}$), two originating from $\boldsymbol{B}$ ($\boldsymbol{BCE}$; $\boldsymbol{BCD}$), three originating from $\boldsymbol{C}$ ($\boldsymbol{CBA}$; $\boldsymbol{CDE}$; $\boldsymbol{CED}$), three originating from $\boldsymbol{D}$ ($\boldsymbol{DCB}$; $\boldsymbol{DCE}$; $\boldsymbol{DEC}$), and three originating from $\boldsymbol{E}$ ($\boldsymbol{ECB}$; $\boldsymbol{ECD}$; $\boldsymbol{EDC}$). Notice that paths are sequences so that, for example, $\boldsymbol{ECB}$ is distinct from $\boldsymbol{BCE}$.

From the paths of length = 2, a subset of them form closed triangles (recall the colloquialism above). In this network, there are 6 closed triangles: two originating from $\boldsymbol{C}$ ($\boldsymbol{CDEC}$; $\boldsymbol{CEDC}$), two originating from $\boldsymbol{D}$ ($\boldsymbol{DCED}$; $\boldsymbol{DECD}$), and two originating from $\boldsymbol{E}$ ($\boldsymbol{ECDE}$; $\boldsymbol{EDCE}$). As the case with paths of length = 2, closed triangles are also sequences so that, for example, $\boldsymbol{ECDE}$ is distinct from $\boldsymbol{EDCE}$.

[34] To be clear, a "friend" from the perspective of an analyst is a colleague who shares a connection with her.
[35] Network transitivity is also known as global clustering coefficient.

Notice that closed triangles can only be formed from three nodes that sit on the same path of length = 2. Following this, the transitivity of the network in Figure 6 is the ratio of the number of closed triangles to the number of paths of length = 2.

$$Transitivity\ of\ \boldsymbol{P} = \frac{6}{12} = 0.500$$

Consider another simple network $\boldsymbol{Q}$ in Figure 7.

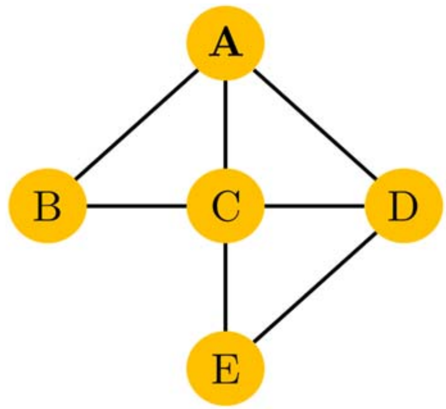

Figure 7

Again, we begin by counting the number of paths of length = 2. In total, there are 28 paths of length = 2: six originating from $\boldsymbol{A}$ ($\boldsymbol{ABC}$; $\boldsymbol{ACB}$; $\boldsymbol{ACD}$; $\boldsymbol{ACE}$; $\boldsymbol{ADC}$; $\boldsymbol{ADE}$), five originating from $\boldsymbol{B}$ ($\boldsymbol{BAC}$; $\boldsymbol{BAD}$; $\boldsymbol{BCA}$; $\boldsymbol{BCD}$; $\boldsymbol{BCE}$), six originating from $\boldsymbol{C}$ ($\boldsymbol{CAB}$; $\boldsymbol{CAD}$; $\boldsymbol{CBA}$; $\boldsymbol{CDA}$; $\boldsymbol{CDE}$; $\boldsymbol{CED}$), six originating from $\boldsymbol{D}$ ($\boldsymbol{DAB}$; $\boldsymbol{DAC}$; $\boldsymbol{DCA}$; $\boldsymbol{DCB}$; $\boldsymbol{DCE}$; $\boldsymbol{DEC}$), and five originating from $\boldsymbol{E}$ ($\boldsymbol{ECA}$; $\boldsymbol{ECB}$; $\boldsymbol{ECD}$; $\boldsymbol{EDA}$; $\boldsymbol{EDC}$).

Of the set of paths with length = 2, 18 of them form closed triangles: four originating from $\boldsymbol{A}$ ($\boldsymbol{ABCA}$; $\boldsymbol{ACBA}$; $\boldsymbol{ACDA}$; $\boldsymbol{ADCA}$), two originating from $\boldsymbol{B}$ ($\boldsymbol{BACB}$; $\boldsymbol{BCAB}$), six originating from $\boldsymbol{C}$ ($\boldsymbol{CABC}$; $\boldsymbol{CBAC}$; $\boldsymbol{CADC}$; $\boldsymbol{CDAC}$; $\boldsymbol{CDEC}$; $\boldsymbol{CEDC}$), four originating from $\boldsymbol{D}$ ($\boldsymbol{DACD}$; $\boldsymbol{DCAD}$; $\boldsymbol{DCED}$; $\boldsymbol{DECD}$), and two originating from $\boldsymbol{E}$ ($\boldsymbol{ECDE}$; $\boldsymbol{EDCE}$).

$$Transitivity\ of\ \boldsymbol{Q} = \frac{18}{28} = 0.643$$

From a quick visual inspection of Figures 6 and 7, we observe that the nodes in $\boldsymbol{Q}$ are more interconnected than those in $\boldsymbol{P}$. This observation agrees with our calculations that $\boldsymbol{Q}$ has a higher transitivity score than $\boldsymbol{P}$.

Appendix II. Variable Definitions

Analyst Centrality	Standardized PCA-extracted factor score of 4 network centrality measures – Degree Centrality, Closeness Centrality, Betweenness Centrality and Eigenvector Centrality.
Degree Centrality	Degree Centrality is related to the number of colleagues that the analyst is immediately connected to in the brokerage network (see Appendix I for details on centrality measures).
Closeness Centrality	Closeness Centrality is related to the path distances between the analyst and all her colleagues in the brokerage network (see Appendix I for details on centrality measures).
Betweenness Centrality	Betweenness Centrality is related to the number of geodesics (shortest paths) in the brokerage network that pass through the analyst (see Appendix I for details on centrality measures).
Eigenvector Centrality	Eigenvector Centrality is related to the notion that the centrality of the analyst is high if her connected colleagues are also well-connected in the brokerage network (see Appendix I for details on centrality measures).
Revision Frequency	Total number of revisions made by the analyst in the firm-year.
Horizon	Number of days elapsed between the analyst's firm-year forecast and the actual earnings announcement. We exclude all forecasts that are more than 365 days or less than 30 days from the actual earnings announcement.
Boldness	The deviation of an analyst's forecast from the consensus forecast in the previous year.
General Experience	Number of months elapsed between an analyst's first appearance in I/B/E/S and her firm-year forecast.
Firm Experience	Number of months elapsed between an analyst's earliest forecast of the firm in I/B/E/S and her firm-year forecast.
Brokerage Experience	Number of months elapsed between an analyst's first appearance in the brokerage in I/B/E/S and her firm-year forecast.
Firm Breadth	Number of unique firms covered by the analyst in the year.
Industry Breadth	Number of unique 2-digit GICS (Global Industry Classification Standard) sectors covered by the analyst in the year.
Lowball	Number of times over the past 3 years that lowballing forecasts were issued for the firm by the analyst. For a forecast to be classified as a lowballing one, 3 conditions must be met. First, the forecast value must be below the actual EPS value. Second, the forecast error (absolute difference between forecast value and actual EPS value) must be either greater than 3 cents or higher than 5% of the actual EPS value. Third, to reduce the likelihood of mistaking forecasting difficulty for lowballing behavior, the difference between the forecast value and the consensus value must be greater than 3 cents or higher than 5% of the consensus value.
Brokerage Size	Logarithm of the number of analysts employed by the brokerage in the year.
Loss	Indicator that equates to unity if the actual EPS of the firm is negative, and equates to zero otherwise.
Forecast Dispersion	Standard deviation of firm-year forecast values among analysts.
Analyst Coverage	Number of analysts who have contributed at least 1 forecast in the firm-year.
Leverage	Sum of short-term debt and long-term borrowings, scaled by total assets of the firm.
Market-to-Book Ratio	Ratio of firm book value to its market capitalization in the year.
Total Assets	Total assets of the firm.
ROA Volatility	Volatility of the ratio of net income to total firm assets over the past 36 months.

Appendix II. (Continued)

Analyst Revision	Difference between the forecast revision value and the previous forecast value, scaled by the absolute value of the latter variable. *Analyst Revision* is a signed variable – positive (negative) values of *Analyst Revision* reflect increments (declines) in the analyst's forecasted value from her previous forecast.
Forecast Error	Absolute difference between the analyst's final firm-year forecast and the actual firm-year EPS value.
Normalized Forecast Error	*Forecast Error* scaled by the average *Forecast Error* across analysts in the firm-year.
Number of Forecasts	The number of competitor analysts who issue a forecast in the (-1, +1) window surrounding an analyst's forecast.
Consensus Deviation	Absolute difference between the analyst's revision value and the prevailing firm-year consensus value, scaled by the absolute value of the latter variable.
Self Deviation	Absolute difference between the analyst's revision value and her prior forecast value, scaled by the absolute value of the latter variable.
Number of Forecasts	Number of firm-year forecasts issued by all analysts in the week of the analyst's forecast revision.
Colleague Optimism	For a forecast revision of a given analyst, we collect instances of her colleagues' realized forecast errors that occurred within the past 30 days. We only retain the realized forecast errors of 1) colleagues who are directly connected to the analyst (i.e. who cover the same industries), and 2) colleagues who cover either the major suppliers or major customers of the analyst's firms. For each forecast error of the analyst's colleagues, we classify them as *optimistic* if the forecasted value is above the actual earnings. If the forecast error is zero, it is neither optimistic nor pessimistic but is still counted in the window. *Colleague Optimism* is the proportion of *optimistic* forecast errors issued by colleagues in the 30-day window.
Global Optimism	For a given forecast revision of an analyst, we collect all instances (including non-colleagues) realized forecast errors in the same 2-digit GICS sector within the past 30 days. Alternatively, the firms covered by all analysts (including non-colleagues) in the same 2-digit GICS sector must have announced their actual earnings in the same 30-day window. For each forecast error, we classify it as optimistic if the forecasted value is above the actual earnings. If the forecast error is zero, it is neither optimistic nor pessimistic but is still counted in the window. *Global Optimism* is the proportion of optimism forecast errors issued by all non-colleague analysts in the 30-day window.
All-American	Indicator that equates to unity if the analyst belongs to the Institutional Investor All-America Research Team in the year, and equates to zero otherwise.
Herding Rate	Following Clement and Tse (2005), an analyst's forecast revision in the firm-year is classified as non-herding only when it is either above both her prior forecast and pre-revision consensus forecast or below both her prior forecast and pre-revision consensus forecast. Otherwise, the revision is classified as herding. Herding Rate is the ratio of herding revisions to the total number of revisions made by an analyst in the firm-year.
Abs. CAR [-1, +1]	Market-adjusted cumulative abnormal returns, centered on the forecast revision date.
Outperformance	A firm-year forecast is defined as an *outperforming* forecast if its forecast error is lower than the mean forecast error. *Outperformance (%)* of an analyst is the percentage of her forecasts in a year that are categorized as *outperforming* forecasts.

Appendix II. (Continued)

Firm Complexity	Using segment data from Compustat, we determine the number of 2-digit SIC industries that each firm operates in, following the procedure in Cohen and Lou (2012). For each firm, we then compute the Herfindahl-Hirschman index (HHI) of its industry segment sales. A high HHI indicates that firm sales are concentrated in one or few industries, and implies that the firm has low complexity. A symmetrical reasoning may be applied to a low HHI. To facilitate interpretation, we multiply the HHI by -1 and define this measure as *Firm Complexity*. A firm with a more positive *Firm Complexity* score indicates that firm sales are dispersed across multiple industry segments, and that the firm is complex.
High R&D	An indicator that equates to unity if either 1) R&D expenses are missing, or 2) R&D intensity (ratio of R&D to total assets) is above the year median, and equates to zero otherwise.
Revision Count	For each analyst-brokerage pair, we find her earliest and latest forecast announcement dates. In between these two dates, we divide the analyst's tenure at the brokerage by months. We exclude the first and last months of her tenure from our sample to avoid truncation issues. In each month, Revision Count is the total number of forecast revisions made by the analyst.
Closeness-weighted Revision Activity	For a given analyst, we find the shortest-path length between her and her brokerage colleagues. For example, if analyst A is directly connected to analyst B, and analyst C is in turn connected to analyst B, the shortest-path length between analysts A and C is two. Using the reciprocals of shortest-path lengths as weights, *Closeness-weighted Revision Activity* of an analyst is the weighted sum of her colleague's *Revision Count* in the previous month.
Distance1 Revision Activity	*Distance1 Revision Activity* of an analyst is the total *Revision Count* of her colleagues who are separated by a shortest-path length of one, normalized by the number of such colleagues.
Distance2 Revision Activity	*Distance2 Revision Activity* of an analyst is the total *Revision Count* of her colleagues who are separated by a shortest-path length of two, normalized by the number of such colleagues.

Appendix III. Brokerage Merger Events in Table 8

We compile the list of brokerage mergers from Hong and Kacperczyk (2010), and Kelly and Ljungqvist (2012). The table below documents the 17 mergers in our final sample.

Acquirer Brokerage	**Target Brokerage**	**Date**	
Salomon Smith Barney	Schroder Wertheim	Apr	2000
Soundview	Wit Capital	May	2000
Paine Webber	J.C. Bradford	Jun	2000
Credit Suisse First Boston	Donaldson Lufkin & Jenrette	Oct	2000
UBS Warburg Dillon Reed	Paine Webber	Dec	2000
J.P. Morgan	Chase Manhattan	Dec	2000
J.P. Morgan Chase	Hambrecht & Quist	Jan	2001
Dresdner Bank	Wasserstein Perella	Feb	2001
Fahnestock	Josephthal Lyon & Ross	Sep	2001
First Union	Wachovia Securities	Oct	2001
UBS	Schwab Soundview	Oct	2004
Janney Montgomery Scott	Parker/Hunter	Mar	2005
Merrill Lynch	Advest	Dec	2005
Merrill Lynch	Petrie Parkman	Dec	2006
Stifel Financial	Ryan Beck & co	Jan	2007
Fox-Pitt Kelton	Cochran, Caronia Securities	Sep	2007
Fahnestock	CIBC World Markets	Jan	2008

Figure 1. Example of a Within-brokerage Network

This figure maps the network structure of I/B/E/S brokerage identifier (481) in the year 2005. Each node represents an analyst in the brokerage. The numbers below each node identify the GICS sectors covered by the analyst.[36] Larger sized nodes with more intense colors (green to red) reflect a higher numbers of direct connections.

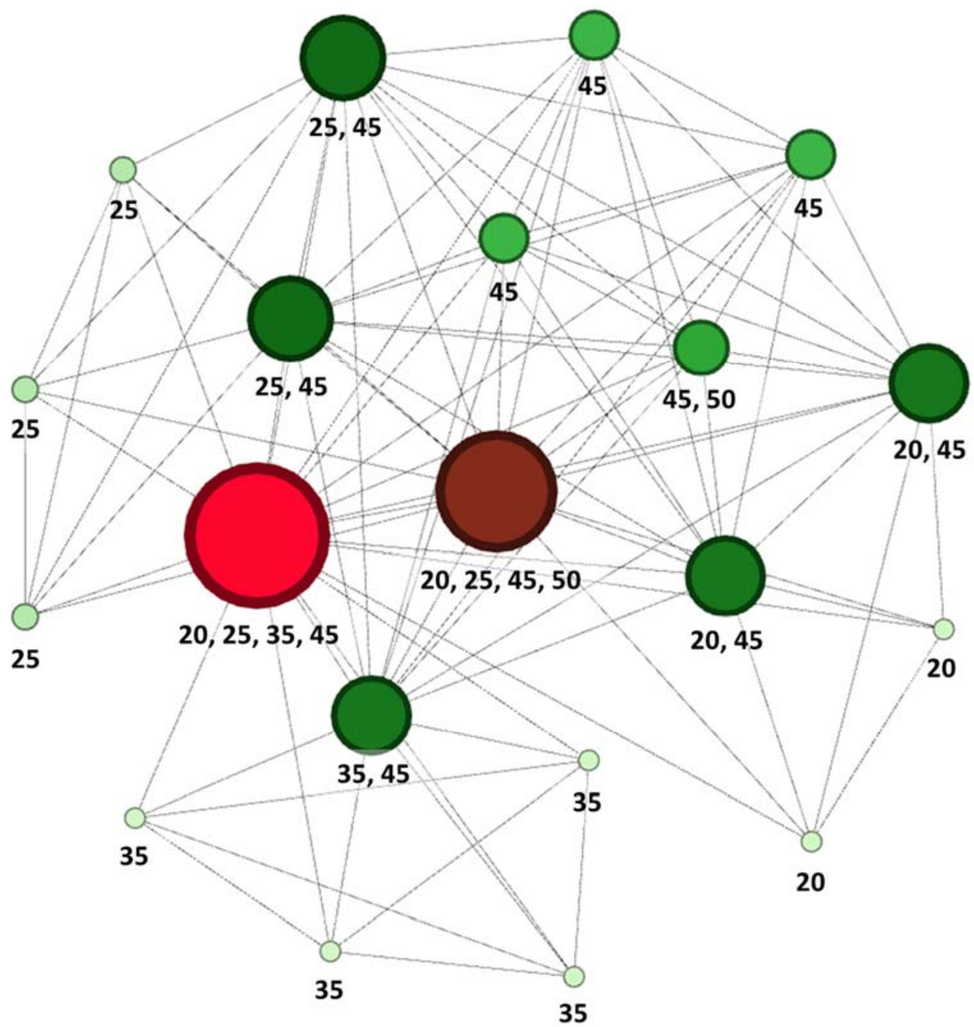

Figure 1

[36] The 2-digit GICS sector codes map to the economic sectors as follows: 20 (industrials), 25 (consumer discretionary), 35 (healthcare), 45 (information technology), and 50 (telecommunication services). For a full map of all eleven 2-digit GICS sector codes, refer to https://www.msci.com/gics.

Figure 2. Distribution of GICS Sector Coverage

This figure shows the distribution of GICS sector coverage across analyst centrality quartiles. In each year, we sort unique analysts by their analyst centrality scores and allocate them to quartiles. Within each quartile, we track the distribution of analyst-firms under coverage across GICS sectors.

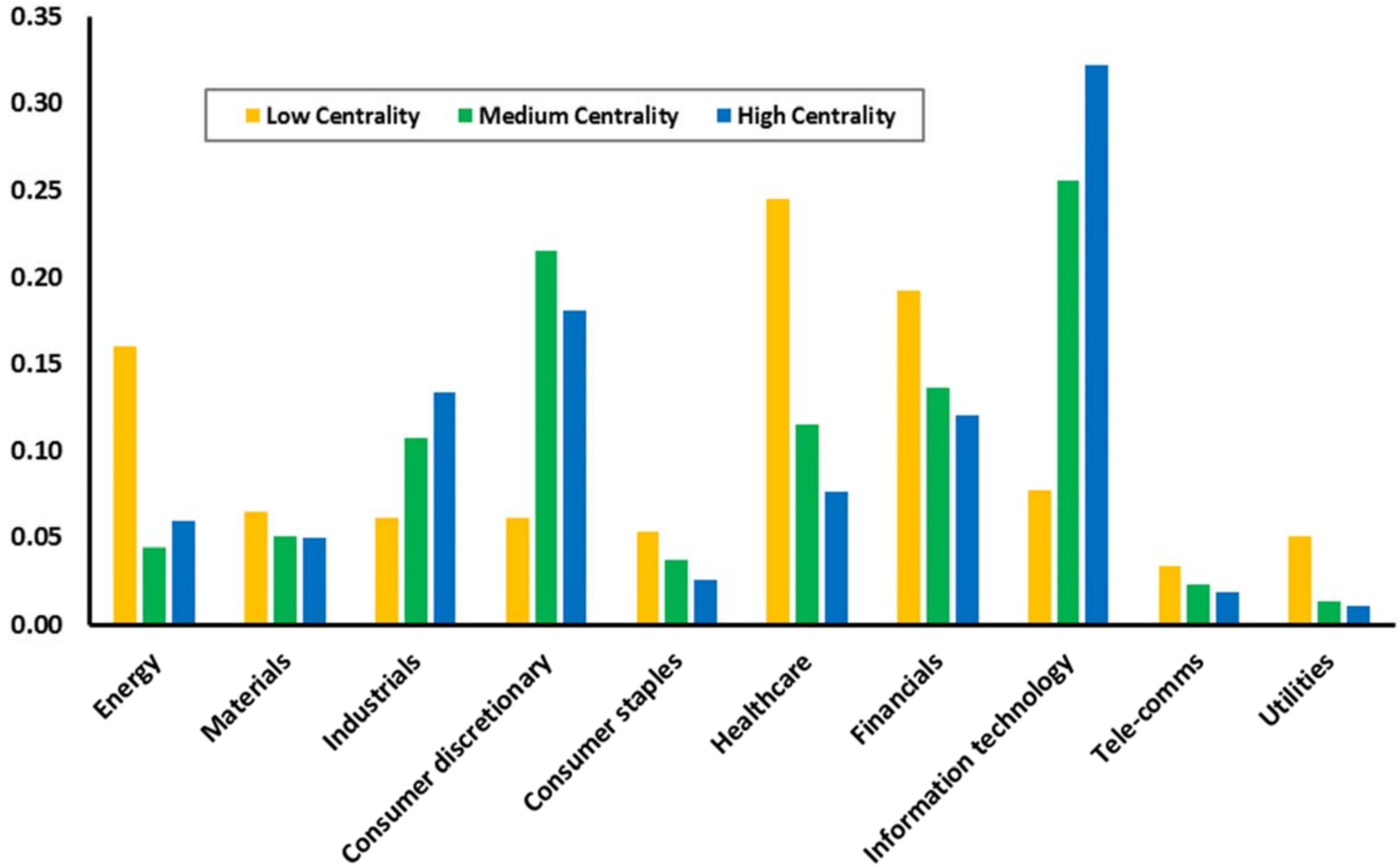

Table 1.
Descriptive Statistics

Panel A reports the summary statistics of unique analyst-year pairs. Since an analyst covers multiple firms in a year, we report summary statistics at the analyst-year level to show analyst characteristics that drive the formation of within-brokerage networks. Panel B reports Pearson correlations among selected variables of unique analyst-year pairs. Panel C reports results of principal component analysis on the four centrality measures. Panel D reports the summary statistics of the baseline sample used in Table 2 Column (1). Due to data limitations, descriptive statistics of *All-American* is based on a sample truncated in the year 2008. In Panel A and B, *Firm Experience* and *Brokerage Experience* are averaged at the analyst-year level.

Panel A. Summary Statistics (unique analyst-year pairs)

	N	Mean	StDev	p10	p25	p50	p75	p90
Number of Connections	60694	16.023	14.565	3	6	12	22	35
Industry Breadth	60694	1.564	0.914	1	1	1	2	3
Firm Breadth	60694	9.660	7.188	1	4	9	14	19
General Experience (mth)	60694	52.286	51.445	2	12	37	78	129
Firm Experience (mth)	60694	22.671	26.588	0	2	14	34	59
Brokerage Experience (mth)	60694	2.470	2.981	0	0	12	48	72
Brokerage Size	60694	58.753	51.107	11	19	42	88	123
All-American	42558	0.102	0.303	0	0	0	0	1
Analyst Centrality	60694	0.007	0.984	-1.034	-0.720	-0.229	0.597	1.512
Degree Centrality	60694	0.371	0.257	0.103	0.169	0.295	0.511	0.774
Closeness Centrality	60694	0.553	0.193	0.345	0.452	0.536	0.646	0.807
Betweenness Centrality	60694	0.021	0.063	0	0	0	0.010	0.058
Eigenvector Centrality	60694	0.151	0.137	0.006	0.033	0.115	0.232	0.354

Table 1. (Continued)
Panel B. Pearson Correlations (unique analyst-year pairs)

		[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]
Analyst Centrality	[A]	1										
Degree Centrality	[B]	0.949	1									
Closeness Centrality	[C]	0.889	0.887	1								
Betweenness Centrality	[D]	0.518	0.324	0.277	1							
Eigenvector Centrality	[E]	0.848	0.746	0.592	0.378	1						
Industry Breadth	[F]	0.571	0.494	0.437	0.521	0.476	1					
Firm Breadth	[G]	0.165	0.158	0.18	0.128	0.078	0.307	1				
General Experience	[H]	0.008	-0.004	-0.009	0.048	0.011	0.061	0.344	1			
Firm Experience	[I]	-0.053	-0.066	-0.062	0.014	-0.037	0.013	0.287	0.801	1		
Brokerage Experience	[J]	0.021	0.015	0.029	0.045	-0.01	0.111	0.339	0.641	0.628	1	
Brokerage Size	[K]	-0.376	-0.409	-0.176	-0.144	-0.46	-0.119	-0.042	0.001	0.051	0.085	1
All-American	[L]	-0.121	-0.133	-0.07	-0.031	-0.141	0.01	0.203	0.192	0.245	0.208	0.261

Panel C. Principal Component Analysis

	Component (1)	Component (2)	Component (3)	Component (4)
Degree Centrality	0.948	-0.208	0.085	-0.225
Closeness Centrality	0.886	-0.258	0.345	0.171
Betweenness Centrality	0.512	0.851	0.118	-0.002
Eigenvector Centrality	0.846	-0.011	-0.528	0.074
Eigenvalue	2.661	0.834	0.419	0.085
Variance explained (%)	66.5	20.9	10.5	2.1
Cumulative variance (%)	66.5	87.4	97.9	100

Table 1. (Continued)
Panel D. Summary Statistics (baseline sample)

	N	Mean	S.D	p10	p25	p50	p75	p90
Analyst Centrality	274673	0.083	0.958	-0.983	-0.619	-0.107	0.671	1.459
Degree Centrality	274673	0.385	0.247	0.111	0.192	0.326	0.538	0.750
Closeness Centrality	274673	0.569	0.177	0.377	0.469	0.555	0.663	0.796
Betweenness Centrality	274673	0.028	0.068	0.000	0.000	0.000	0.023	0.082
Eigenvector Centrality	274673	0.155	0.130	0.008	0.045	0.126	0.237	0.346
Horizon	274673	116.442	90.613	21	55	98	125	273
Brokerage Size	274673	60.444	48.143	12	22	47	91	119
Revision Frequency	274673	3.664	2.868	1	2	3	5	7
Boldness	274673	-0.002	0.256	-0.110	-0.025	0.001	0.023	0.093
General Experience	274673	80.503	54.008	20	37	69	114	160
Firm Experience	274673	42.068	38.132	10	13	28	58	96
Firm Breadth	274673	15.451	7.737	7	11	15	19	24
Industry Breadth	274673	1.789	1.039	1	1	1	2	3
Lowball	274673	0.159	0.366	0	0	0	0	1
Loss	274673	0.125	0.330	0	0	0	0	1
Analyst Coverage	274673	16.945	9.702	6	9	15	23	31
Forecast Error	274673	0.222	33.947	0.005	0.013	0.040	0.110	0.290

Table 2.
The Effect of Analyst Centrality on Forecast Accuracy

This table presents results from OLS regressions on the effect of analyst centrality on forecast accuracy. The dependent variable is *Normalized Forecast Error*, defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The key independent variable is *Analyst Centrality. Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). Columns (2) and (4) exclude firm financial variables because they are invariant at the firm-year level. Detailed definitions of other variables are in Appendix II. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 2. (Continued)

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error	(3) OLS: Normalized Forecast Error
Analyst Centrality	-0.019***	-0.014***	-0.009***
	(0.003)	(0.003)	(0.003)
Revision Frequency	-0.006***	-0.005***	-0.009***
	(0.001)	(0.001)	(0.002)
Boldness	-0.019***	-0.019***	-0.020**
	(0.006)	(0.006)	(0.009)
Horizon	0.003***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)
General Experience	-0.019***	-0.020***	-0.019***
	(0.002)	(0.002)	(0.003)
Firm Experience	0.003	0.004	-0.011***
	(0.003)	(0.003)	(0.003)
Firm Breadth	-0.000	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)
Industry Breadth	0.014***	0.006**	0.012***
	(0.003)	(0.003)	(0.002)
Lowball	0.067***	0.066***	0.084***
	(0.005)	(0.005)	(0.006)
Loss		-0.011**	
		(0.005)	
Analyst Coverage		-0.003***	
		(0.000)	
Leverage		0.010	
		(0.007)	
Book-to-Market Ratio		0.002	
		(0.002)	
Total Assets		0.008***	
		(0.001)	
ROA Volatility		-0.000	
		(0.000)	
Brokerage Size			0.003
			(0.002)
Observations	274,673	274,673	274,673
R-squared	0.176	0.177	0.277
Brokerage-Year fixed effects	Yes	Yes	No
Firm-Year fixed effects	No	No	Yes
Brokerage-Year cluster	Yes	Yes	No
Firm-Year cluster	No	No	Yes
Analyst-Firm cluster	Yes	Yes	Yes

Table 3.
The Effect of Analyst Centrality on Market Reactions to Forecasts

This table presents results from OLS regressions. The dependent variable in this table is the absolute 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date. The key independent variable is *Analyst Centrality* and its interaction with *Consensus Deviation*. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing consensus, normalized by the absolute value of the latter variable. Stock return variables are unitized in percentage points. Detailed definitions of other variables are in Appendix II. In Column (3), we exclude a forecast revision if the firm files SEC 8-Ks in the (-1, 0) day window of the revision. In Column (4), we exclude a forecast revision if the firm either files SEC 8-Ks or has an earnings announcement in the (-1, 0) day window of the revision. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 3. (Continued)

	(1) OLS: Abs. CAR [-1, +1]	(2) OLS: Abs. CAR [-1, +1]	(3) OLS: Abs. CAR [-1, +1]	(4) OLS: Abs. CAR [-1, +1]
Exclude 8-K events	No	No	Yes	Yes
Exclude earnings announcements	No	No	Yes	Yes
Standalone forecasts only	No	No	No	Yes
Analyst Centrality [A]	0.336***	0.288***	0.209***	0.145***
	(0.022)	(0.023)	(0.021)	(0.023)
[A] x Consensus Deviation	0.014***	0.009***	0.022**	0.055***
	(0.005)	(0.003)	(0.009)	(0.011)
Consensus Deviation	0.011**	0.007**	0.011**	0.026***
	(0.005)	(0.003)	(0.004)	(0.005)
General Experience		0.113***	0.050***	0.028
		(0.017)	(0.017)	(0.021)
Firm Experience		-0.063***	-0.049***	-0.016
		(0.009)	(0.009)	(0.012)
Firm Breadth		-0.027***	-0.021***	-0.019***
		(0.002)	(0.002)	(0.002)
Industry Breadth		-0.139***	-0.142***	-0.091***
		(0.021)	(0.018)	(0.023)
Brokerage Size		0.225***	0.231***	0.177***
		(0.020)	(0.019)	(0.023)
Number of Forecasts		0.206***	0.308***	
		(0.008)	(0.014)	
Abs. CAR [-5, -2]		-0.022***	-0.028***	-0.030***
		(0.005)	(0.004)	(0.006)
Leverage		0.723***	0.564***	0.455***
		(0.135)	(0.121)	(0.139)
Book-to-Market Ratio		0.389***	0.350***	0.304***
		(0.091)	(0.085)	(0.108)
Total Assets		-0.784***	-0.632***	-0.599***
		(0.020)	(0.017)	(0.019)
ROA Volatility		0.001	0.004*	0.007
		(0.002)	(0.002)	(0.005)
Observations	1,211,243	1,211,243	637,010	130,626
R-squared	0.004	0.125	0.126	0.093
Week cluster	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes

Table 4.
Does Competition Weaken Learning?

This table presents results from OLS regressions. The dependent variable in Panel A is *Normalized Forecast Error*, defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The dependent variable in Panel B is the absolute 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date. The key independent variables are *Analyst* and its interaction with *Consensus Deviation*. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing consensus, normalized by the absolute value of the latter variable. We split the sample by the investment banking business size of the brokerages in Columns (1) and (2). In every year, we rank investment banks (IB) by their combined IPO and SEO deal values. Investment banks are added to the *Large IBs* pool sequentially until the pool accounts for 75% of the total deal value in the market of that year. All other brokerages (i.e. smaller IBs and non-IBs) are assigned to the *Non-Large IBs* pool. We split the sample by *Brokerage Size* in Columns (3) through (5). In every year, we sort brokerages into terciles according to their *Brokerage Size*. Column (3) contains the smallest brokerages while Column (6) contains the largest brokerages. We include control variables used in Column (2) of Table 2 or Column (3) of Table 3 but suppress their estimated coefficients to facilitate presentation. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 4. (Continued)
Panel A. Competition, Analyst Centrality, and Forecast Accuracy

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error	(3) OLS: Normalized Forecast Error	(4) OLS: Normalized Forecast Error	(5) OLS: Normalized Forecast Error
Sample	Non-Large IBs	Large IBs	Small brokerages	Medium brokerages	Large brokerages
Analyst Centrality	-0.016***	0.018	-0.013***	-0.023***	0.000
	(0.003)	(0.014)	(0.004)	(0.006)	(0.009)
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	235,744	38,927	87,352	91,031	96,288
R-squared	0.179	0.166	0.199	0.173	0.159
Brokerage-year fixed	Yes	Yes	Yes	Yes	Yes
Brokerage-year cluster	Yes	Yes	Yes	Yes	Yes
Analyst-firm cluster	Yes	Yes	Yes	Yes	Yes

Table 4. (Continued)
Panel B. Competition, Analyst Centrality, and Market Reactions

	(1)	(2)	(3)	(4)	(5)
	OLS: Abs. CAR [-1, +1]	OLS: Abs. CAR [-1, +1]	OLS: Abs. CAR [-1, +1]	OLS: Abs. CAR [-1, +1]	OLS: Abs. CAR [-1, +1]
Sample	Non-Large IBs	Large IBs	Small brokerages	Medium brokerages	Large brokerages
Exclude 8-K events and earnings announcements	Yes	Yes	Yes	Yes	Yes
Analyst Centrality [A]	0.197***	0.527***	0.136***	0.390***	0.513***
	(0.019)	(0.075)	(0.018)	(0.042)	(0.083)
[A] x Consensus Deviation	0.020**	0.011	0.018*	0.022	0.001
	(0.009)	(0.022)	(0.009)	(0.023)	(0.002)
Consensus Deviation	0.010**	0.036**	0.069***	0.070***	0.001
	(0.004)	(0.017)	(0.015)	(0.018)	(0.001)
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	529,280	107,712	224,329	196,050	216,613
R-squared	0.128	0.133	0.125	0.137	0.128
Week cluster	Yes	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes	Yes

Table 5.
Analyst Centrality and Hard-to-Value Firms

This table presents results from OLS regressions. The dependent variable is *Normalized Forecast Error,* defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The key independent variable is *Analyst Centrality* and its interactions with *High R&D* and *Firm Complexity*. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *High R&D* is an indicator that equates to unity if either 1) R&D expenses are missing, or 2) R&D intensity (ratio of R&D to total assets) is above the year median, and equates to zero otherwise. To construct *Firm Complexity*, we compute the Herfindahl index of firm sales across the industry segments in which the firm operates (Cohen and Lou, 2012). *Firm Complexity* is this Herfindahl index multiplied by negative one. Therefore, a more positive value of *Firm Complexity* correspond to higher complexity in a firm's operations. Control variables used in the Column (2) of Table 2 are also included in the model estimations. However, their estimated coefficients are suppressed to facilitate presentation. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 5. (Continued)

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error	(3) OLS: Normalized Forecast Error
Analyst Centrality [A]	-0.009**	-0.013***	-0.008*
	(0.004)	(0.003)	(0.004)
[A] x High R&D	-0.007*		-0.007**
	(0.004)		(0.004)
[A] x Firm Complexity		-0.020**	-0.022**
		(0.009)	(0.010)
High R&D	-0.009**		-0.009**
	(0.004)		(0.004)
Firm Complexity		-0.005	-0.007
		(0.009)	(0.009)
Revision Frequency	-0.006***	-0.005***	-0.006***
	(0.001)	(0.001)	(0.001)
Boldness	-0.017***	-0.017***	-0.017***
	(0.006)	(0.006)	(0.006)
Horizon	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)
General Experience	-0.020***	-0.020***	-0.020***
	(0.003)	(0.003)	(0.003)
Firm Experience	0.003	0.004	0.003
	(0.003)	(0.003)	(0.003)
Firm Breadth	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Industry Breadth	0.005**	0.006**	0.005**
	(0.003)	(0.003)	(0.003)
Lowball	0.067***	0.067***	0.067***
	(0.005)	(0.005)	(0.005)
Loss	-0.009*	-0.010**	-0.009*
	(0.005)	(0.005)	(0.005)
Analyst Coverage	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)
Leverage	0.009	0.010	0.009
	(0.007)	(0.007)	(0.007)
Book-to-Market Ratio	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Total Assets	0.008***	0.009***	0.008***
	(0.001)	(0.001)	(0.001)
ROA Volatility	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Observations	274,673	274,673	274,673
R-squared	0.178	0.178	0.178
Brokerage-Year fixed effects	Yes	Yes	Yes
Brokerage-Year cluster	Yes	Yes	Yes
Analyst-Firm cluster	Yes	Yes	Yes

Table 6.
Learning from Colleagues' Ex-Post Forecast Errors

This table presents results from OLS regressions. The dependent variable is *Analyst Revision*. *Analyst Revision* is a signed variable which equates to the difference between the revised forecast value and the previous forecast value, scaled by the absolute value of the latter. Hence, a positive (negative) value of *Analyst Revision* reflects an increment (a decline) in the analyst's forecasted value from her previous forecast. The key independent variables are *Colleague Optimism*, *Global Optimism*, and their respective interactions with *Analyst Centrality*. For a forecast revision of a given analyst, we collect instances of her colleagues' realized forecast errors that occurred within the past 30 days. We only retain the realized forecast errors of 1) colleagues who are directly connected to the analyst (i.e. who cover the same industries), and 2) colleagues who cover either the major suppliers or major customers of the analyst's firms. In other words, the firms covered by the analyst's colleagues must have announced their actual earnings in the same 30-day window. For each forecast error of the analyst's colleagues, we classify them as *optimistic* if the forecasted value is above the actual earnings. *Colleague Optimism* is the proportion of *optimistic* forecast errors in the 30-day window. To construct *Global Optimism*, we collect all realized forecast errors outside of the analyst's brokerage in the same 30-day window. Following this, *Global Optimism* is the proportion of *optimistic* forecast errors made by non-colleagues in the 30-day window. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). Detailed definitions of other variables are in Appendix II. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 6. (Continued)

	(1) OLS: Analyst Revision	(2) OLS: Analyst Revision	(3) OLS: Analyst Revision
Colleague Optimism [A]	-0.848***	-0.819***	-0.462***
	(0.207)	(0.207)	(0.173)
Global Optimism [B]			-7.289***
			(1.727)
[A] x Analyst Centrality		-0.402**	-0.356**
		(0.159)	(0.158)
[B] x Analyst Centrality			-0.055
			(0.810)
Analyst Centrality	-0.314**	-0.168	-0.179
	(0.152)	(0.154)	(0.300)
Number of Colleague Forecasts	-0.004	-0.005*	-0.005*
	(0.003)	(0.003)	(0.003)
Number of Global Forecasts			0.000
			(0.000)
General Experience	0.128	0.123	0.143*
	(0.084)	(0.083)	(0.083)
Firm Experience	-0.064	-0.065	-0.059
	(0.054)	(0.054)	(0.054)
Firm Breadth	-0.022**	-0.022**	-0.021**
	(0.009)	(0.009)	(0.009)
Industry Breadth	-0.062	-0.072	-0.063
	(0.120)	(0.119)	(0.119)
Brokerage Size	-0.045	-0.002	-0.042
	(0.126)	(0.133)	(0.130)
Analyst Coverage	0.098***	0.098***	0.099***
	(0.009)	(0.009)	(0.009)
Abs. CAR [-5, -2]	-0.359***	-0.358***	-0.355***
	(0.031)	(0.031)	(0.030)
Observations	698,697	698,697	698,697
R-squared	0.012	0.012	0.013
Week cluster	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes

Table 7.
Do Analysts Perform Better in Denser Networks?

This table presents results from Tobit regressions. The dependent variable is *Outperformance (%)*. A firm-year forecast is defined as an *outperforming* forecast if its forecast error is lower than the mean forecast error. *Outperformance (%)* of an analyst is the percentage of her forecasts in a year that are categorized as *outperforming* forecasts. The key independent variables are *Analyst Centrality* and *Network Density. Analyst Centrality* is the standardized PCA-extracted factor score of four network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Network Density* is a brokerage-level variable that measures the density of the within-brokerage network in a year. Refer to Appendix I for details on its construction and working examples. In column (4), we exclude brokerages that cover fewer than three 2-digit GICS sectors. Detailed definitions of other variables are in Appendix II. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

	(1) TOBIT: Outperformance (%)	(2) TOBIT: Outperformance (%)	(3) TOBIT: Outperformance (%)
Exclude brokerages that cover less than 3 sectors	No	No	Yes
Analyst Centrality	0.779***	0.619***	0.898***
	(0.214)	(0.225)	(0.276)
Network Density		3.661*	4.644**
		(1.615)	(2.307)
General Experience	0.795***	0.790***	0.819***
	(0.183)	(0.183)	(0.261)
Firm Experience	0.375**	0.376**	0.374*
	(0.181)	(0.181)	(0.223)
Industry Breadth	-2.557***	-2.430***	-2.658***
	(0.219)	(0.226)	(0.213)
Firm Breadth	0.241***	0.243***	0.243***
	(0.025)	(0.025)	(0.028)
Brokerage Size	0.029***	0.030***	0.030***
	(0.003)	(0.003)	(0.005)
Observations	60,694	60,694	58,826
Brokerage-year cluster	Yes	Yes	Yes

Table 8.
The Causal Effect of Analyst Centrality on Forecast Accuracy: Shocks to Analyst Centrality due to Brokerage Mergers

This table presents results from a difference-in-difference model with multiple groups, and multiple shocks across time. We define shocks as brokerage mergers (Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012) from years 2000 to 2007. Appendix III contains the list of 17 brokerage mergers that we are able to match to our final sample. For each merger event, we track all analysts who work at the acquirer before and after the mergers. We further require that each analyst covers the same firm before and after the merger. Therefore, our unit of observation in this quasi-natural experiment is an analyst-firm. The treatment variables are *Increase* (*Decrease*) *in Analyst Centrality*, an indicator variable which equates to unity if the average post-merger *Analyst Centrality* is higher (lower) than the average pre-merger *Analyst Centrality*, and equates to zero otherwise. Since visual inspection to validate the parallel trend assumption is tenuous in a model with shocks spread across time, we include temporal leads and lags of the treatment in the model to test the assumption econometrically (e.g. Autor (2003)). We use a (-3, +3) year window centered on the merger event. The dependent variable in this table is *Normalized Forecast Error*, defined as *Forecast Error* scaled by the average forecast error in the firm-year. The key independent variables are the temporal leads (*Pre-Treatment*) and lags (*Post-Treatment*) of the treatment. We also include analyst time trends to help control for confounding trends. Analyst time trends are *General Experience*, *Brokerage Size*, *Firm Breadth*, and *Industry Breadth.* Detailed definitions of the variables are in Appendix II. To facilitate presentation, coefficient estimates of analyst time trends are not presented. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 8. (Continued)

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error
Treatment indicator	Increase in Analyst Centrality	Decrease in Analyst Centrality
Leads and Lags of Treatment:		
Pre-Treatment (t-3)	0.110	-0.193
	(0.113)	(0.120)
Pre-Treatment (t-2)	0.136	-0.117
	(0.089)	(0.077)
Pre-Treatment (t-1)	-0.060	-0.078
	(0.072)	(0.070)
Post-Treatment (t+1)	-0.111	0.024
	(0.078)	(0.079)
Post-Treatment (t+2)	-0.161*	0.191**
	(0.083)	(0.088)
Post-Treatment (t+3)	-0.209**	0.419***
	(0.094)	(0.095)
Analyst-Firm dummies	Yes	Yes
Year dummies	Yes	Yes
Analyst time trends	Yes	Yes
Analyst-Firm cluster	Yes	Yes
Observations	4,515	4,515
R-squared	0.280	0.285

Table 9.
Analyst Centrality and Analyst Ability

This table presents results from OLS regressions. Due to data limitations on the identities of All-American analysts, our sample period ends in 2008. The dependent variables are *Normalized Forecast Error* and *Abs. CAR*. *Normalized Forecast Error* is defined as *Forecast Error* scaled by the average forecast error in the firm-year. *Abs. CAR* is the absolute 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date. The key independent variables are *All*-American, *Analyst Centrality*, and their interactions with *Consensus Deviation*. *All-American* is an indicator variable that equates to unity if the analyst belongs to the Institutional Investor All-America Research Team in the year, and equates to zero otherwise. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing consensus, normalized by the absolute value of the latter variable. Stock return variables are unitized in percentage points. In Panel B, we exclude a forecast revision if the firm either files SEC 8-Ks or has an earnings announcement in the (-1, 0) day window of the revision. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Table 9. (Continued)
Panel A. Analyst Centrality, Ability, and Forecast Accuracy

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error
Analyst Centrality	-0.011**	-0.010**
	(0.005)	(0.005)
All-American		-0.033***
		(0.007)
Revision Frequency	-0.008***	-0.008***
	(0.001)	(0.001)
Boldness	-0.031***	-0.031***
	(0.010)	(0.010)
Horizon	0.004***	0.004***
	(0.000)	(0.000)
General Experience	-0.018***	-0.017***
	(0.004)	(0.004)
Firm Experience	0.006	0.008
	(0.005)	(0.005)
Firm Breadth	-0.001	-0.000
	(0.000)	(0.000)
Industry Breadth	0.004	0.004
	(0.004)	(0.004)
Lowball	0.064***	0.064***
	(0.008)	(0.008)
Loss	-0.036***	-0.036***
	(0.007)	(0.007)
Analyst Coverage	-0.001**	-0.001**
	(0.000)	(0.000)
Leverage	0.004	0.005
	(0.012)	(0.012)
Book-to-Market Ratio	0.001	0.001
	(0.006)	(0.006)
Total Assets	0.007***	0.007***
	(0.002)	(0.002)
ROA Volatility	0.022	0.021
	(0.016)	(0.016)
Observations	166,023	166,023
R-squared	0.224	0.224
Brokerage-Year fixed effects	Yes	Yes
Brokerage-Year cluster	Yes	Yes
Analyst-Firm cluster	Yes	Yes

Table 9. (Continued)

Panel B. Analyst Centrality, Ability, and Market Reactions

	(1) OLS: Abs. CAR [-1, +1]	(2) OLS: Abs. CAR [-1, +1]
Exclude 8-K events	Yes	Yes
Exclude earnings announcements	Yes	Yes
Analyst Centrality [A]	0.172***	0.173***
	(0.021)	(0.021)
All-American [B]		-0.200***
		(0.039)
[A] x Consensus Deviation	0.210***	0.242***
	(0.079)	(0.080)
[B] x Consensus Deviation		0.450**
		(0.226)
Consensus Deviation	2.192***	2.111***
	(0.113)	(0.115)
General Experience	0.183***	0.185***
	(0.019)	(0.018)
Firm Experience	-0.042***	-0.040***
	(0.008)	(0.008)
Firm Breadth	-0.015***	-0.014***
	(0.002)	(0.002)
Industry Breadth	-0.149***	-0.151***
	(0.017)	(0.017)
Brokerage Size	0.182***	0.207***
	(0.017)	(0.017)
Number of Forecasts	0.202***	0.202***
	(0.008)	(0.008)
Abs. CAR [-5, -2]	-0.016***	-0.016***
	(0.003)	(0.003)
Leverage	0.224***	0.236***
	(0.065)	(0.065)
Book-to-Market Ratio	0.164***	0.163***
	(0.037)	(0.037)
Total Assets	-0.447***	-0.446***
	(0.008)	(0.009)
ROA Volatility	0.004***	0.004***
	(0.001)	(0.001)
Firm financial controls	Yes	Yes
Observations	364,945	364,945
R-squared	0.094	0.094
Week cluster	Yes	Yes
Firm cluster	Yes	Yes

Table 10.
Analyst Centrality and Regulation Fair Disclosure

This table presents results from OLS regressions. The dependent variable in Columns (1) and (2) is *Normalized Forecast Error,* defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The dependent variable in Columns (3) and (4) is the 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date and unitized in percentage points. The key independent variable is *Analyst Centrality* – the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). The pre-Regulation Fair Disclosure sample period is from 1995 to 2000 (based on the fiscal year ends of firms). The post-Regulation Fair Disclosure sample period begins from 2001. Control variables used in the Column (2) of Table 2 or Column (3) of Table 3 are also included in the model estimations. However, their estimated coefficients are suppressed to facilitate presentation. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error	(3) OLS: Abs. CAR [-1, +1]	(4) OLS: Abs. CAR [-1, +1]
Sample period	Pre-RegFD	Post-RegFD	Pre-RegFD	Post-RegFD
Exclude 8-K events	-	-	Yes	Yes
Exclude earnings announcements	-	-	Yes	Yes
Analyst Centrality [A]	-0.001	-0.017***	0.212***	0.188***
	(0.008)	(0.004)	(0.034)	(0.024)
[A] x Consensus Deviation			0.020*	0.017**
			(0.011)	(0.008)
Consensus Deviation			0.081***	0.009**
			(0.017)	(0.004)
Control variables	Yes	Yes	Yes	Yes
Welch-Satterthwaite t-test between [A] in Col. (1) and (2):				t-stat: 1.79 df: 78,564
Welch-Satterthwaite t-test between [A] x Consensus Deviation in Col. (3) and (4):				t-stat: 0.22 df: 355,333
Observations	50,999	223,643	167,135	469,857
R-squared	0.206	0.172	0.111	0.139
Brokerage-Year fixed effects	Yes	Yes	No	No
Brokerage-Year cluster	Yes	Yes	No	No
Analyst-Firm cluster	Yes	Yes	No	No
Week cluster	No	No	Yes	Yes
Firm cluster	No	No	Yes	Yes

Internet Appendix Table 1.
Additional Results: Analyst Centrality and Market Reactions to Forecasts

This table supplements the main results in Table 3. The dependent variable in this table is the absolute 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date. The key independent variable is *Analyst Centrality* and its interactions with *Consensus Deviation* and *Self Deviation. Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing consensus, normalized by the absolute value of the latter variable. *Self Deviation* is the absolute difference between an analyst's revision value and his prior forecast value, normalized by the absolute value of the latter variable. Stock return variables are unitized in percentage points. Detailed definitions of other variables are in Appendix II. We exclude a forecast revision if the firm files SEC 8-Ks in the (-1, 0) day window of the revision or has an earnings announcement in the (-1, 0) day window of the revision. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Internet Appendix Table 1. (Continued)
Analyst Centrality and Market Reactions to Forecasts

	(1) OLS: Abs. CAR [-1, +1]	(2) OLS: Abs. CAR [-1, +1]	(3) OLS: Abs. CAR [-1, +1]	(4) OLS: Abs. CAR [-1, +1]
Exclude 8-K events	No	No	Yes	Yes
Exclude earnings announcements	No	No	Yes	Yes
Standalone forecasts only	No	No	No	Yes
Analyst Centrality [A]	0.337***	0.289***	0.210***	0.157***
	(0.022)	(0.023)	(0.021)	(0.023)
[A] x Self Deviation	0.012	0.008*	0.015**	0.005
	(0.007)	(0.005)	(0.007)	(0.008)
Self Deviation	0.012	0.008	0.008**	0.003
	(0.008)	(0.005)	(0.003)	(0.004)
General Experience		0.113***	0.050***	0.029
		(0.017)	(0.017)	(0.021)
Firm Experience		-0.063***	-0.049***	-0.016
		(0.009)	(0.009)	(0.012)
Firm Breadth		-0.027***	-0.021***	-0.019***
		(0.002)	(0.002)	(0.002)
Industry Breadth		-0.139***	-0.142***	-0.091***
		(0.021)	(0.018)	(0.023)
Brokerage Size		0.225***	0.231***	0.177***
		(0.020)	(0.019)	(0.023)
Number of Forecasts		0.206***	0.308***	
		(0.008)	(0.014)	
Abs. CAR [-5, -2]		-0.022***	-0.028***	-0.030***
		(0.005)	(0.004)	(0.006)
Leverage		0.723***	0.565***	0.461***
		(0.135)	(0.121)	(0.139)
Book-to-Market Ratio		0.390***	0.351***	0.307***
		(0.091)	(0.085)	(0.109)
Total Assets		-0.784***	-0.632***	-0.600***
		(0.020)	(0.017)	(0.019)
ROA Volatility		0.001	0.004*	0.007
		(0.002)	(0.002)	(0.005)
Observations	1,211,243	1,211,243	637,010	130,626
R-squared	0.004	0.125	0.126	0.093
Week cluster	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes

Internet Appendix Table 2.
Additional Results: Learning from Colleagues' Ex-Post Forecast Errors

This table supplements the main results in Table 4. The dependent variable is *Analyst Revision*. *Analyst Revision* is a signed variable which equates to the difference between the revised forecast value and the previous forecast value, scaled by the absolute value of the latter. Hence, a positive (negative) value of *Analyst Revision* reflects an increment (a decline) in the analyst's forecasted value from her previous forecast. The key independent variables are *Colleague Optimism*, *Global Optimism*, and their respective interactions with *Analyst Centrality*. For a forecast revision of a given analyst, we collect instances of her colleagues' realized forecast errors that occurred within the past 60 days. We only retain the realized forecast errors of 1) colleagues who are directly connected to the analyst (i.e. who cover the same industries), and 2) colleagues who cover either the major suppliers or major customers of the analyst's firms. In other words, the firms covered by the analyst's colleagues must have announced their actual earnings in the same 60-day window. For each forecast error of the analyst's colleagues, we classify them as *optimistic* if the forecasted value is above the actual earnings. *Colleague Optimism* is the proportion of *optimistic* forecast errors in the 60-day window. To construct *Global Optimism*, we collect all realized forecast errors outside of the analyst's brokerage in the same 60-day window. Following this, *Global Optimism* is the proportion of *optimistic* forecast errors made by non-colleagues in the 60-day window. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). Detailed definitions of other variables are in Appendix II. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Internet Appendix Table 2. (Continued)
Learning from Colleagues' Ex-Post Forecast Errors

	(1) OLS: Analyst Revision	(2) OLS: Analyst Revision	(3) OLS: Analyst Revision
Colleague Optimism [A]	-0.287***	-0.276**	-0.155
	(0.109)	(0.109)	(0.096)
Global Optimism [B]			-3.362**
			(1.351)
[A] x Analyst Centrality		-0.205*	-0.191*
		(0.109)	(0.108)
[B] x Analyst Centrality			-0.072
			(0.615)
Analyst Centrality	-0.319***	-0.249**	-0.251
	(0.111)	(0.118)	(0.246)
Number of Colleague Forecasts	-0.002	-0.002	-0.001
	(0.003)	(0.003)	(0.002)
Number of Global Forecasts			-0.000
			(0.000)
General Experience	0.283***	0.283***	0.293***
	(0.063)	(0.063)	(0.064)
Firm Experience	-0.159***	-0.159***	-0.156***
	(0.035)	(0.035)	(0.035)
Firm Breadth	-0.002	-0.002	-0.001
	(0.006)	(0.006)	(0.006)
Industry Breadth	-0.056	-0.060	-0.051
	(0.089)	(0.089)	(0.088)
Brokerage Size	-0.103	-0.102	-0.133
	(0.096)	(0.096)	(0.092)
Analyst Coverage	0.087***	0.087***	0.088***
	(0.009)	(0.009)	(0.009)
Abs. CAR [-5, -2]	-0.337***	-0.337***	-0.338***
	(0.022)	(0.022)	(0.022)
Observations	698,697	698,697	698,697
R-squared	0.010	0.010	0.011
Week cluster	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes

Internet Appendix Table 3.
Diffusion of Revision Activity within Brokerage Networks

This table presents results from negative binomial regressions. The dependent variable is *Revision Count*. For each analyst-brokerage pair, we find her earliest and latest forecast announcement dates. In between these two dates, we divide the analyst's tenure at the brokerage by months. We exclude the first and last months of her tenure from our sample to avoid truncation issues. In each month, *Revision Count* is the total number of forecast revisions made by the analyst. The key independent variable is *Closeness-weighted Revision Activity*. For a given analyst, we find the shortest-path length between her and her brokerage colleagues. For example, if analyst A is directly connected to analyst B, and analyst C is in turn connected to analyst B, the shortest-path length between analysts A and C is two. Using the reciprocals of shortest-path lengths as weights, *Closeness-weighted Revision Activity* of an analyst is the weighted sum of her colleague's *Revision Count* in the previous month. In Column (2), we adopt a normalized measure of *Closeness-weighted Revision Activity* where the *Revision Count* in each of the analyst's distance tier is scaled by the number of colleagues. In Column (3), *Distance1 Revision Activity* of an analyst is the total *Revision Count* of her colleagues who are separated by a shortest-path length of one, normalized by the number of such colleagues. In Column (4), we define *Distance2 Revision Activity* analogously. Estimated coefficients are inflated by 100 to facilitate presentation. Detailed definitions of other variables are in Appendix II. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Internet Appendix Table 3. (Continued)

	(1) N-BINOM: Revision Count	(2) N-BINOM: Revision Count	(3) N-BINOM: Revision Count	(4) N-BINOM: Revision Count
Closeness-weighted Revision Activity [A]	0.014***			
	(0.003)			
Size-Normalized [A]		1.865***		
		(0.076)		
Distance1 Revision Activity [B[			3.235***	
			(0.116)	
Distance2 Revision Activity [C]				1.164***
				(0.113)
General Experience	1.489***	1.042***	1.092***	1.339***
	(0.170)	(0.169)	(0.169)	(0.171)
Industry Breadth	-12.582***	-11.897***	-12.274***	-12.348***
	(0.169)	(0.167)	(0.166)	(0.167)
Brokerage Size	0.076***	0.090***	0.094***	0.091***
	(0.004)	(0.003)	(0.003)	(0.003)
Global Revision Activity	24.140***	22.783***	22.825***	23.803***
	(0.170)	(0.172)	(0.169)	(0.170)
Calendar month indicators	Yes	Yes	Yes	Yes
Exposure variable: Firm Breadth (coefficient constrained to 1)				
Welch-Satterthwaite t-test between estimated coefficients [B] and [C]:			t-stat: 12.79	
			df: 956,311	
Observations	478,485	478,485	478,485	478,485

Internet Appendix Table 4.
Do Central or Peripheral Analysts Drive the Findings?

This table presents results from OLS regressions. The dependent variable in Column (1) is *Normalized Forecast Error,* defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The key independent variables are indicators constructed based on *Analyst Centrality*. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). In each year, we sort analysts by their *Analyst Centrality* scores into either quartiles or terciles. *Low Centrality* is an indicator that equates to unity if *Analyst Centrality* is less than the 25th (or 33rd) percentile value of the distribution, and equates to zero otherwise. *High Centrality* is an indicator that equates to unity if *Analyst Centrality* is more than the 75th (or 33rd) percentile value of the distribution, and equates to zero otherwise. In Columns (1) through (3), we sort *Analyst Centrality* into quartiles. In Columns (4) through (6), we sort *Analyst Centrality* into terciles. Detailed definitions of other variables are in Appendix II. In all specifications, we include control variables from Column (2) of Table 2. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

Internet Appendix Table 4. (Continued)
Central and Peripheral Analysts

	(1) OLS: Normalized Forecast Error	(2) OLS: Normalized Forecast Error	(3) OLS: Normalized Forecast Error	(4) OLS: Normalized Forecast Error	(5) OLS: Normalized Forecast Error	(6) OLS: Normalized Forecast Error
Sort analyst centrality	Quartiles	Quartiles	Quartiles	Terciles	Terciles	Terciles
Benchmark category	Above p25	Below p75	p25 to p75	Above p33	Below p66	p33 to p66
Low Centrality	0.011**		0.011**	0.012***		0.010**
	(0.005)		(0.005)	(0.004)		(0.004)
High Centrality		-0.016***	-0.016***		-0.019***	-0.017***
		(0.006)	(0.006)		(0.005)	(0.005)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	274,673	274,673	274,673	274,673	274,673	274,673
R-squared	0.177	0.177	0.177	0.177	0.177	0.177
Brokerage-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brokerage-Year cluster	Yes	Yes	Yes	Yes	Yes	Yes
Analyst-Firm cluster	Yes	Yes	Yes	Yes	Yes	Yes

Internet Appendix Table 5.
Forecast Outcomes in 2-digit SIC Brokerage Networks

This table presents results from OLS regressions. The dependent variable in Column (1) is *Normalized Forecast Error,* defined as *Raw Forecast Error* scaled by the average forecast error in the firm-year. *Raw Forecast Error* is defined as the absolute difference between an analyst's last firm-year forecast and the actual value. The dependent variable in Column (2) is the 3-day market-adjusted cumulative abnormal returns, centered on the forecast revision date and unitized in percentage points. The key independent variable is *Analyst Centrality* and its interaction with *Consensus Deviation*. *Analyst Centrality* is the standardized PCA-extracted factor score of 4 network centrality measures – *Degree Centrality*, *Closeness Centrality*, *Betweenness Centrality* and *Eigenvector Centrality* (see Appendix I for details on centrality measures). *Consensus Deviation* is the absolute difference between an analyst's revision value and the prevailing consensus, normalized by the absolute value of the latter variable. Detailed definitions of other variables are in Appendix II. In Column (1), we include control variables from Column (2) of Table 2. In Column (2), we include control variables from Column (3) of Table 3. Robust standard errors are reported in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

	(1) OLS: Normalized Forecast Error	(2) OLS: Abs. CAR [-1, +1]
Exclude 8-K events	-	Yes
Exclude earnings announcements	-	Yes
Analyst Centrality (SIC-2d) [A]	-0.014***	0.225***
	(0.003)	(0.019)
[A] x Consensus Deviation		0.023**
		(0.010)
Consensus Deviation		0.012**
		(0.005)
Control variables	Yes	Yes
Observations	267,353	612,422
R-squared	0.177	0.130
Brokerage-Year fixed effects	Yes	No
Brokerage-Year cluster	Yes	No
Analyst-Firm cluster	Yes	No
Week cluster	No	Yes
Firm cluster	No	Yes