What Do Questions Reveal?

The Detection, Recognition, and Implications of Topic-Specific Expertise in Capital Markets*

Ling Cen	Yanru Han	Jarrad Harford
CUHK	Stevens Institute of Technology	University of Washington
ling.cen@cuhk.edu.hk	yhan47@stevens.edu	jarrad@uw.edu

Abstract:

We provide a generalizable methodology to identify analysts' topic-specific expertise based on their past earnings conference call questions. When their topic-specific expertise aligns with the content of new information, such expertise leads to more accurate forecasts, more influential stock recommendations, and better career prospects. While analyst expertise diversity improves the ability of prices to reflect future earnings, it discourages buy-side participation in earnings conference calls and weakens corporate investment sensitivity to stock prices.

Keywords: Topic-Specific Expertise; Information Processing; Market Efficiency

JEL Classifications: G24; D83; J24

---

* We are grateful for comments and suggestions from Philip Bond, Roger Loh, Tamara Nefedova, Kelly Shue, Yan Xiong, Liyan Yang, Jianfeng Yu, and seminar participants at the 2022 AsianFA, 5th Future of Financial Information Conference, 2023 CEIBS Finance & Accounting Academic Symposium, 2023 CUFE Summer Conference, 2025 University of Kansas Finance Conference, Boston College, The Chinese University of Hong Kong, Wake Forest University, the University of Houston, and the University of Mississippi. Ling Cen acknowledges the funding support from GRF 14500119 offered by the Hong Kong Research Grants Council.

I. Introduction

Understanding how new information is processed by key intermediaries and registered into price is central to financial economics (Fama, Fisher, Jensen, and Roll, 1969; Fama, 1970). A defining characteristic of complex financial markets is the stochastic nature of information flow, where news arrives randomly and often clusters around distinct topical themes, such as macroeconomic shocks, industry-specific developments, technological trends, or geopolitical events. These topical themes of new information evolve over time, generating time-series variation in their market impact, while heterogeneous firm characteristics drive cross-sectional differences in exposure to these topical themes. For instance, during the 2018 US-China Trade War, U.S. firms with greater import exposure to China experienced more pronounced information shocks compared to their less-exposed peers.

Plenty of evidence suggests that investors, when faced with the inherent complexity of information arrival, actively seek out and integrate the perspectives of expert information processors into their trading decisions (e.g., Stickel, 1992, 1995; Gleason and Lee, 2003; Loh and Stulz, 2011). The identification of such experts is generally less costly for industry-specific news, as key information processors, such as financial analysts, are organized by industry sector, and their expertise is publicly signaled through recognitions like star analyst status. However, a significant portion of new information in financial markets transcends industry boundaries, as seen in our earlier example of the US-China Trade War. Further, topic-specific expertise is often shaped by an individual's unique background, experience, and network (e.g., Bae, Stulz, and Tan, 2008; Cohen, Frazzini, and Malloy, 2010; Bradley, Gokkaya, and Liu, 2017), which are not readily observable, and because specific expertise is valuable only when relevant, it is not necessarily revealed through the average historical accuracy of standard outputs like earnings forecasts or stock recommendations. This presents a critical challenge: how can investors reliably identify which information processors possess genuine expertise on a particular emerging topic?

This paper introduces a novel and generalizable methodology to evaluate the topic-specific expertise of sell-side financial analysts, an important group of information processors in capital markets. Our approach quantifies analysts' interests and expertise by measuring the frequency of topic-specific questions they pose during earnings conference call Q&A sessions. Using this measure, we examine whether analysts with topic-specific expertise generate more accurate forecasts than their peers and whether such expertise enhances expert analysts' recognition by investors in capital markets and by employers in labor markets, especially when their specialized expertise coincides with emerging topical themes in new information arrival. Finally, we show that identified topic-specific expertise can generate economic implications in the capital market. Specifically, we investigate the impact of analyst expertise diversity, i.e., whether a firm is covered by analysts with heterogeneous expertise, on price efficiency and the real effects in financial markets.

We begin with an in-depth study of analysts' supply-chain-specific expertise and subsequently generalize our analysis to cover all topics. We first focus on supply-chain-specific expertise primarily because supply-chain information arrivals are plausibly exogenous to both analysts and the supplier firms they cover. For a dependent supplier firm in our sample, firm-specific supply-chain information is significantly affected by its principal customers due to asymmetric bargaining power. Market-wide supply-chain shocks, such as the 2018 US-China Trade War that we rely on as an identification strategy in this paper, are exogenous to any individual firm or its analysts but have generated significant economic impacts on the earnings and asset prices of US firms with international trade partners (Ding, Levine, Lin, and Xie, 2021; Huang, Lin, Liu, and Tang, 2023). We provide a detailed discussion of its advantages in Section 2.3.1.

In our study of supply-chain-specific expertise, we first use a machine-learning algorithm (i.e., Word2Vec, developed by Mikolov, Sutskever, Chen, Corrado, and Dean, 2013) to identify supply-chain-specific questions. Analysts covering the firm who have the highest frequency (i.e., in the top 10^{th} percentile) of raising supply-chain-related questions in earnings conference calls of *other firms* in the last four years are identified as supply-chain experts. We remove the questions in earnings conference calls of the focal firm because those questions are likely triggered by the focal firm's corporate activities and decisions. For example, Steve Smigie from Raymond James Financial made a forecast for Skyworks Solutions before its fiscal year-end of 2016. For this observation, we first identify all 1,331 questions he raised from 2012 up to the day before this forecast announcement date in all conference calls held by other firms (i.e., *excluding* all questions he raised in conference calls held by Skyworks Solutions). We then screen all questions using the supply-chain keyword list and identify 169 of the 1,331 questions as those related to supply-chain topics. The 90th percentile level of the number of supply-chain-related questions across all analyst-firms for 2016 is 103. Therefore, Steve Smigie is identified as a supply-chain expert analyst for Skyworks Solutions in 2016.

When a firm discloses relationships with principal customers and faces heightened risks from customer concentration, supply-chain expert analysts achieve a greater reduction in forecast errors compared to non-expert peers covering the same firm. Our estimates indicate that this difference-in-differences in forecast errors accounts for 4.05% of the average forecast error in the full sample. We posit that the informational advantage stemming from supply-chain expertise should be more pronounced for supplier firms with higher information asymmetry (e.g., Aboody and Lev, 2002) and customer-switching risk (e.g., Intintoli, Serfling, and Shaikh, 2017). Cross-sectional analyses support these predictions. Furthermore, capital markets appear to at least partially recognize supply-chain expertise, as prices respond more to expert analysts' recommendation revisions. After firms disclose relationships with principal customers, expert analysts' upgrades (downgrades) trigger average abnormal returns 45.2 bps higher (46.3 bps lower) than those of non-experts on the day of the recommendation update.

Taking advantage of the 2018 US-China Trade War as a market-wide exogenous shock to global supply-chain risks, we document that supply-chain expert analysts achieved a greater reduction in forecast errors for US firms importing from or exporting to China after the outbreak of the trade war. The labor market value of their supply-chain expertise also increased. Specifically, we find that after the trade war began: (1) among analysts who remained with their current employer, supply-chain experts were more likely to receive internal promotions; (2) among those switching jobs, experts were more likely to transition to (more lucrative) buy-side institutions; and (3) among analysts moving to other sell-side brokerage firms, experts were more likely to join larger firms, relative to non-experts.

Next, we generalize our methodology to accommodate all prespecified and non-prespecified topical themes of new information. We first establish the broader applicability of our approach by examining additional prespecified themes, including environmental, product, labor, and macroeconomic-related topics. Our findings indicate that when firms are exposed to new information pertaining to these themes, topic-specific expert analysts demonstrate superior earnings forecast accuracy relative to their non-expert peers.

For non-prespecified topics, we adjust our procedures as follows. First, we utilize BERT (Bidirectional Encoder Representations from Transformers, developed by Devlin, Chang, Lee, and Toutanova, 2019) and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise, developed by Campello, Moulavi, and Sander, 2013) algorithms to cluster analyst questions from earnings conference calls into distinct topical categories. Second, we identify "hot" topics based on their relative frequency during our sample period. Third, following an analogous procedure to our identification of supply-chain expert analysts, we classify topic-specific expert analysts for each topic. Fourth, for each firm–topic–year, we define the peak period as the year when the fraction of questions related to a given topic in the firm's earnings calls exceeds the sample-wide mean share of that topic by more than two standard deviations. Finally, we construct a *Coincide* variable to capture instances where at least one of the analyst's topical expertise aligns with a theme experiencing its peak information arrival for a firm the analyst covers.

When *Coincide* equals one, i.e., when the expertise "meets" its corresponding information arrival for any topic, we find that the patterns observed in our motivating study persist. Specifically, at the peak of information arrival for a particular topic: (1) expert analysts covering that topic exhibit a greater reduction in forecast errors compared to their non-expert peers; (2) investors respond more aggressively to recommendation upgrades and downgrades issued by expert analysts; and (3) expert analysts experience higher valuation in the labor market, relative to non-experts. The economic magnitudes of these effects are consistent with those in our motivating study of supply-chain-specific expertise. For example, when an expert analyst's specialization coincides with a topic of information arrival, their probability of transitioning to a buy-side institution increases by 2.7 percentage points relative to non-expert peers, conditional on job turnover; this effect represents 37.2% of the average transition probability. Similarly, conditional on moving to another sell-side brokerage firm, these experts are 6.7 percentage points more likely than their non-expert peers to join a larger brokerage firm, an increase equivalent to 15.8% of the average probability for such a move.

In the final section of this paper, we study one important economic implication of the identified topic-specific expertise. We examine the impact of analyst expertise diversity—defined as whether a firm is covered by analysts with heterogeneous expertise—on price efficiency and the real effects in capital markets. We define price efficiency as the extent to which market prices incorporate information and accurately forecast firm value, similar to the forecasting price efficiency (FPE) in Bond, Edmans, and Goldstein (2012). Prior literature, both theoretical (e.g., Goldstein and Yang, 2015) and empirical (e.g., Fang and Hope, 2021; Gerken and Painter, 2023), consistently demonstrates that information diversity, whether stemming from diverse information sources or the backgrounds of information processors, enhances price efficiency. Building on the empirical frameworks of Bai, Philippon, and Savov (2016) and Kacperczyk, Sundaresan, and Wang (2021), we provide consistent evidence that price efficiency improves with greater analyst expertise diversity.

The real effects, related to revelatory price efficiency (RPE) in Bond, Edmans, and Goldstein (2012), capture the extent to which managers can learn from prices to guide their investment decisions. Goldstein and Yang (2019) argue that whether public information disclosure improves managers' learning from prices depends on the comparative informational advantage of market participants of the disclosed information. If publicly revealed information overlaps with what managers seek to learn from traders, it may crowd out private information, thereby impairing managers' ability to learn from prices. Following their theoretical framework, we suggest that whether expertise diversity improves or hinders managers' ability to learn from prices is an empirical question, depending on the type of information that analysts can produce together as their expertise diversity increases: A positive effect prevails if increased expertise diversity increases public production of the types of information in which managers hold a comparative advantage; A negative effect may arise if the increase in analyst expertise diversity mainly impacts the types of information that managers want to learn from traders. In the latter case, the improvement of expertise diversity leads investors to put a lower weight on their own private information ex post and weakens their incentives to produce private information ex ante.

While it is empirically challenging to measure real efficiency (since researchers cannot observe all investment opportunities in front of managers), we focus on two observable dimensions of the real effects in financial markets. First, we demonstrate that analyst expertise diversity reduces buy-side participation in earnings conference calls, suggesting that higher analyst expertise diversity discourages buy-side information production activities. Second, extending the empirical approach of Chen, Goldstein, and Jiang (2007), we demonstrate that the sensitivity of corporate investment to price declines when analyst expertise diversity increases. These results are consistent with the conjecture that the increase in analyst expertise diversity is likely to improve the public production of the types of information that managers want to learn from investors. While public signals of this type of information become more accurate, investors are likely to place a lower weight on their own private information in trading ex post, which reduces their incentives to become privately informed about that information ex ante. This mechanism leads to a reduction in overall price informativeness, which hurts managers' ability to learn from prices and potentially damages the real efficiency in capital markets.

Our main contribution is to provide a generalizable methodology that provides new insight into how information intermediaries acquire and express expertise, and how it translates into performance. Under our framework, exogenous information arrivals "meet" the pre-identified skills of analysts. We show that, when the type of information coincides with the type of expertise, the outputs of information processing are more accurate. This method may generate the following implications for researchers, practitioners, and regulators. First, this method will allow practitioners and regulators to identify opinions from experts in the context of different information events, which they might have already been using in an unsystematic manner. Second, this method can be generalized and implemented by practitioners to detect all topic-specific expertise of any key information processors who raise questions or speak publicly, including corporate managers, institutional investors (e.g., fund managers), and politicians. Third, economic measures based on topic-specific expertise, such as the analyst expertise diversity measure we discussed in Section 6, help researchers and practitioners reveal economic dynamics in the capital market.

Further, our study extends the expanding literature on topical analysis of analyst outputs. Huang, Lehavy, Zang, and Zheng (2018) compare topics between those in manager narratives of earnings conference calls and those in analyst reports issued after the conference calls. They find that new topics covered in analyst reports but not in manager narratives provide incremental information to the capital market. Iwasaki, Chen, Du, and Tu (2021) investigate topic tones by a deep supervised learning approach and show that tones of qualitative topics have a significant and incremental predictive power for short-term future abnormal returns. Martineau and Zoican (2021) construct an information contribution measure based on the average cosine distance between the topic distribution for a particular analyst report and any subset of competitor reports. They find that the information content of analyst reports for high-coverage stocks is much lower than that for low-coverage stocks. All studies above focus on how topics in analyst outputs help us understand incremental information production in the capital market. Our paper contributes to this literature from a distinctive labor-market perspective, i.e., we use the frequency of topics in analysts' questions (in conference calls for *other firms*) to identify analyst topic-specific skills for a specific firm and our approach is flexible to any simple or sophisticated topical models.

Finally, previous studies consistently find that various types of analyst skills are critical determinants of forecast accuracy. These skills encompass not only general and firm-specific experience (Clement, 1999; Clement, Koonce, and Lopez, 2007) but also deep industry knowledge (Kadan, Madureira, Wang, and Zach, 2012; Bradley, Gokkaya, and Liu, 2017). Furthermore, superior performance is linked to an analyst's soft skills to acquire information through management access (Cohen, Frazzini, and Malloy, 2010; Green, Jame, Markov, and Subasi, 2014). By identifying all potential topic-specific expertise of analysts, we show that expert analysts become more accurate than non-experts when their topic-specific expertise *coincides* with the type of information arrival.

The remainder of the paper is organized as follows. Section 2 discusses several key decisions in our research design. Section 3 describes data sources and sample characteristics. Section 4 motivates this methodology by showing an in-depth study based on supply-chain-specific expertise. Section 5 generalizes our methodology to all prespecified or non-prespecified topics. Section 6 discusses an important economic implication of the identified topic-specific expertise. We conclude in Section 7.

2. Research Design

2.1. Two Assumptions

Our paper builds upon two assumptions. First, people have a natural inclination to talk about topics that interest them. Big-data firms have explored and taken advantage of this inclination in targeting specific consumers. For example, if one goes to her own Google Ads settings page, one would be surprised by how well Google knows her interests and preferences. Further, people also have a natural inclination to talk about topics that they are confident about and good at, especially in an earnings conference call setting where analysts have incentives to publicly demonstrate their expertise (i.e., to show off their expertise). Therefore, we conjecture that financial analysts' interest and expertise are manifested in the frequency of topics they cover in their questions during earnings conference calls.

The second assumption is that interests, when combined with practice, lead to expertise. There is abundant evidence in psychological studies showing that interest promotes effort, which eventually leads to expertise and superior performance. Laboratory evidence suggests that interest in a certain area stimulates attention, persistent effort (e.g., Ainley, Hidi, and Berndorff, 2002; Renninger, 1987; Reninger and Wozniak, 1985), and re-engagement (Prenzel, 1988) in the area. Furthermore, interest leads to superior performance in information inferencing (Estes and Vaughan, 1973), comprehension (Asher, 1980; Asher, Hymel, and Wigfield, 1978; Asher and Markell, 1974), and retention (Estes and Vaughan, 1973; Renninger, 1987; Renninger and Wozniak, 1985). Topic interest facilitates a high degree of cognitive organization in knowledge structures (Schiefele and Krapp, 1988).[1] Taking these two assumptions together, we argue that the frequency of an analyst's discussion of a specific topic in her conference call questions, as a joint reflection of both interest and practice, can serve as a proxy for her expertise in this topic.

[1] After reviewing more than 150 related studies, Schiefele, Krapp, and Winteler (1992) conclude that interest is correlated with both academic and laboratory performance.

2.2. The Earnings Conference Call Setting

Analysts serve as ideal subjects for our methodology because they are key information intermediaries who regularly generate insights through reports, recommendations, and questions during corporate conference calls. While our approach is theoretically applicable to any information source with sufficient data for topic identification, we focus on earnings conference call questions as the primary setting for detecting analyst expertise for the following reasons. First, unlike analyst reports, which are often coauthored or the product of team efforts, conference calls allow for clear attribution of questions to individual analysts.[2] Second, the time-constrained nature of earnings calls forces analysts to prioritize their most pressing questions, typically concentrating on a single topic. This selectivity helps us discern their areas of greatest interest when the conference call is hosted. Third, earnings calls are public forums where sell-side analysts must demonstrate their competence in front of peers, corporate managers, and buy-side institutions. The reputational risk of asking low-quality (*"cheap-talk"*) questions discourages superficial engagement. Also, such behavior is easily identifiable by researchers based on managerial responses. Together, these features make earnings conference calls an exceptionally reliable and informative setting for studying analyst topic-specific expertise. While mainly relying on the questions from earnings conference calls, we provide a full set of robustness checks based on the analyst expertise detected from analyst reports in Section 4.4 and Appendix V.

2.3. Analyst Supply-Chain-Specific Expertise

2.3.1. Why Do We Start with Supply-Chain-Specific Expertise?

We first focus on the supply-chain-related topic-specific expertise of financial analysts in the baseline tests for the following reasons. First, previous studies show that, while supply-chain-related information is important to predict earnings and other firm performance (e.g., Patatoukas, 2012), it has not received sufficient attention from market participants. For example, Cohen and Frazzini (2008) and Menzly and Ozbas (2010) show a significant delay in information diffusion along the supply chain, as exhibited by

---

[2] Even when an analyst is unable to attend, substitutes explicitly identify themselves (e.g., *"This is ABC asking on behalf of XYZ"*), ensuring transparency in sourcing.

return predictabilities along the supply chain at both the firm and the industry levels.[3] Second, the supply-chain information arrivals are likely exogenous and important to information intermediaries. For example, a supplier firm's disclosure of its principal customers, an important information event affecting the concentration risk of its future cash flows, is exogenous to economic factors determining the expertise of sell-side analysts covering the supplier firm. Major market-wide supply-chain disruptions, such as those triggered by the US-China Trade War, are beyond the control of any firms or analysts but have generated significant economic impacts on both cash flows and asset prices of firms in international supply chains (Ding, Levine, Lin, and Xie, 2021; Huang, Lin, Liu, and Tang, 2023). These features allow us to identify a causal effect of supply-chain-specific expertise on forecast accuracy at the exogenous arrival of important firm-specific or market-wide supply-chain information. Last but not least, unlike broad topics (e.g., corporate governance), supply-chain-related topics can be accurately and efficiently identified by various textual analytic methods.

2.3.2 Identification of Supply-Chain-Specific Expertise

To capture analyst supply-chain-specific expertise, we start by counting the number of questions an analyst raises in conference calls that discuss supply-chain topics. Supply-chain-related questions are defined by supply-chain-specific words. For our baseline test, we rely on a machine-learning algorithm (i.e., Word2Vec) to generate a supply-chain keywords list. We illustrate the detailed method of the machine-learning algorithm in Appendix II. In Online Appendix Table O5, we show that our methodology is robust to alternative approaches to identifying supply-chain-specific questions.

After identifying supply chain keywords, we parse all Q&As in earnings conference call transcripts, in order to construct analyst supply-chain-specific expertise variables. For each earnings forecast made by analyst i for firm j in year t ($Forecast_{i,j,t}$), we extract all questions analyst i raised in all conference calls held by *other* firms (i.e., all firms excluding firm j itself) in the period of year t-4 to the day before the earnings forecast announcement date. We *exclude* past earnings conference calls of firm j itself to address the following concern: although it is very unlikely that a firm discloses a supply-chain relationship because

[3] Cen, Hertzel, and Schiller (2024) suggest that the speed of information diffusion along the supply chain, causally affected by the attention of key information intermediaries, generates real ecomic outcomes in capital markets.

of an analyst's suggestions, it is possible that the questions raised by the analyst (before relationship disclosure) and subsequent relationship disclosure are driven by a common (but unobserved by researchers) firm-specific factor. For example, the frequency of analysts' questions on supply-chain issues in firm *j*'s conference calls and firm *j*'s ex post relationship disclosure are both driven by firm *j*'s desire to have a stable supply-chain relationship under competitive pressure. Since we exclude all questions analyst *i* raised in firm *j*'s previous earnings conference calls, the frequency of raising supply-chain-specific questions is unlikely to be affected by firm-specific supply-chain factors of firm *j*.

We parse each question into tokens so that sentences are split into a bag of words. We search for supply chain keywords in each question and count the number of questions that contain at least one supply chain keyword ($NMention_{i,j,t}$) for each analyst-firm-year observation. For each year *t*, we rank analyst-firm-year observations by the variable *NMention*. Forecasts with $NMention_{i,j,t}$ ranks within the top 10th percentile for year *t* are identified as those made by analysts with supply-chain-specific expertise and, consequently, the *Expertise* dummy has a value of 1 for these observations.[4] We refer to these analysts as "supply-chain experts" for firm *j* in year *t*.

2.3.3 The Generalization for Any Topic-Specific Expertise

For any prespecified topic, we can still employ the methodology discussed in Section 2.3.2 to identify analysts' topic-specific expertise. For non-prespecified topics, we carry out the following three-step procedures. First, we embed each analyst question into a multi-dimensional vector space using Bidirectional Encoder Representations from Transformers (BERT) developed by Devlin, Chang, Lee, and Toutanova (2019). Second, we cluster these question embeddings to form topics using HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise, developed by Campello, Moulavi, and Sander, 2013). Third, we assign each question to the topic cluster with which it has the highest cosine similarity. We provide a detailed discussion of this approach in Section 5.3.1 and Appendix IV.

[4] In Online Appendix Table O6, we demonstrate robustness to alternative cutoffs for identifying experts using conference call questions, and in Online Appendix Table O7 we show robustness to basing expertise on the fraction of firms' earnings calls in which an analyst asks a supply-chain question

3. Sample and Data Sources

3.1. Sample

Our sample includes analysts' latest one-year-ahead (FY1) earnings forecasts before the fiscal year ends from the I/B/E/S Unadjusted Detail History File for the period between 2006 and 2022. For each earnings forecast in our sample, we construct an analyst supply-chain-specific expertise variable based on the analyst's Q&A experience in earnings conference calls in a four-year period before the forecast is made.[5] The main sample starts in 2006 because the conference call transcripts data from Reuters StreetEvent, which we use to construct analysts' supply-chain expertise variables, starts in 2002. We drop observations where analysts cannot be identified, i.e., where analyst codes are missing or equal to zero. Analysts listed in I/B/E/S as covering a firm but not appearing in conference call transcripts data are also included in the sample, where their numbers of questions are assigned a value of zero. The screening criteria above yield a sample of 598,225 analyst-firm-year observations between 2006 and 2022.

3.2. Data Sources

3.2.1. Customer-Supplier Relationship Data

Customer-supplier relationships are identified from the Compustat Segment Customer File. SFAS No. 131 requires that all U.S. public firms disclose their principal customers that represent more than 10% of the firm's total sales in 10-K and 10-Q forms. In addition to mandatory disclosure, many suppliers choose to voluntarily disclose principal customers that account for less than 10% of their total sales (e.g., Intintoli, Serfling, and Shaikh, 2017). Following Cen, Chen, Hou, and Richardson (2018), we define a firm as a dependent supplier of public customers in year *t* if the firm has disclosed at least one public principal customer in year *t*. Similarly, a publicly listed firm is identified as a public principal customer in year *t* if any Compustat firm discloses this firm as a principal customer in year *t*.

3.2.2. U.S. Import and Export Data

We obtain U.S. import and export data from S&P Panjiva, which provides comprehensive records of sea-based shipments into and out of the United States. These records are sourced from mandatory filings

[5] We provide robust empirical evidence based on other horizons in Table 2.

submitted by firms to U.S. Customs and Border Protection (CBP) for physical goods crossing U.S. borders. The dataset includes detailed shipment information, such as the names and addresses of U.S. importers and exporters, their foreign counterparts, as well as shipment origin, destination, and arrival dates.

Based on the data, we identify U.S. firms' trade relationships with China. Specifically, we define $ChinaImport$ ($ChinaExport$) as a dummy variable equal to 1 if a US firm imports from (exports to) China in year t. $ChinaTrade$ is a dummy variable equal to 1 if a US firm engages in either export or import trade with China in year t.

3.2.3. Analyst Career Data

We construct analyst career histories using data from I/B/E/S and Revelio Labs. The I/B/E/S data allows us to track analysts' affiliations with brokerage firms through their earnings forecasts and recommendations. For information on job promotions and career transitions out of sell-side brokerage firms, we use Revelio Labs' comprehensive employment database. Revelio Labs collects and structures publicly available professional profiles from platforms like LinkedIn, containing detailed employment histories including job titles, ONET classifications, company names, employment durations, and job responsibilities as self-reported by individuals.

To link analysts across these datasets, we implement a three-step matching procedure: First, we match I/B/E/S analyst name abbreviations (containing last names and first initials) to full names in Revelio's database to generate potential matches. Second, we refine these matches by requiring: (1) relevant ONET occupational codes (e.g., 13-2051.00 for Financial Analysts), or (2) financial industry job titles containing terms like "Financial Analyst". Finally, we verify employer information by matching masked brokerage abbreviations in I/B/E/S to employer names in Revelio, cross-referencing these links with employer affiliations evident in earnings call transcripts where analysts participate.

We construct three labor market outcome variables using detailed analyst career data from I/B/E/S and Revelio Labs. For internal promotion, we create a dummy variable ($Internal\ Promotion$) that equals one if an analyst receives a promotion within their current employer during the subsequent two-year period, where promotions are identified through a hierarchy of job titles spanning ten ranks: analyst, associate, senior associate, vice president (VP), senior vice president (SVP), executive vice president (EVP), director, senior director, managing director (MD), and chief executive/C-suite positions. For career transition to buy-side institutions, we generate a dummy variable ($Join\ Buy - Side$) indicating whether the analyst transitions from their current sell-side position to a buy-side institution within two years. We identify buy-side institutions from employer names in Revelio Labs using the methodology developed by Jung, Wong, and Zhang (2018). Finally, $Join\ Larger\ Brokerage\ Firm$ is a dummy variable that captures whether an analyst moves to a larger sell-side brokerage within two years, determined by tracking changes in the analyst's brokerage affiliation through their I/B/E/S forecasting records and using the number of analysts employed at each brokerage (as recorded in I/B/E/S) as our measure of brokerage size.

3.2.4. Earnings Conference Call Transcripts Data

We use Q&As in earnings conference call transcripts from the Thomson Reuters StreetEvents database to construct the variable that captures analysts' expertise. Firms typically host earnings conference calls within 48 hours after earnings announcements. An earnings conference call typically has two parts: a managerial presentation and a Q&A section between analysts and managers. All questions and answers raised by analysts in earnings conference calls are included in the data. We first identify firms hosting conference calls in Compustat and CRSP. To analyze analyst earnings forecast data, analyst names in conference call transcripts are matched to the analyst codes in I/B/E/S through the I/B/E/S Detailed Recommendation File. The procedure creates a link between analyst names in conference call transcripts and analyst codes in I/B/E/S. The final conference call dataset spans 21 years (2002-2022) and comprises 6,914,485 questions in 299,166 conference calls. This conference call dataset covers 8,445 unique firms and 8,516 unique analysts after dropping observations in which the analyst names cannot be matched to I/B/E/S analyst codes by the above procedures.

3.3. Summary Statistics

Table 1 reports descriptive statistics of our sample. Panel A reports summary statistics of the main dependent and independent variables. The average absolute forecast error is 1.95% of the firm's stock price, which is comparable to that reported in Rubin, Segal, and Segal (2017). An average analyst covers 17.24 firms per year with a forecast experience of 4.6 years.[6]

[Insert Table 1 Here]

About 25% of forecasts are made for firms that are dependent suppliers, and 31.5% for principal customers. We report firm-level control variables in Panels B and C in Table 1 for dependent suppliers and principal customers, respectively. Dependent suppliers are much smaller firms relative to principal customers. Moreover, principal customers are more profitable than dependent suppliers. These patterns of differences between the two types of firms are well documented in the literature (e.g., Banerjee, Dasgupta, and Kim, 2008).

Table 1 also presents summary statistics for question-related characteristics. *# of Supply Chain Q's* for an analyst-firm-year annual forecast observation is the number of questions of an analyst that mentions supply-chain keywords in the past four years, excluding the questions from the firm's own conference call. *Total # of Questions* is the total number of questions raised by the analyst in the past four years. The average analyst raises 399 questions in conference calls in a four-year period; the average analyst participates in four earnings conference calls and raises 25 questions every quarter. On average 38 of the 399 questions, i.e., around 10% of all questions, discuss supply-chain-related issues.

4. One Motivating In-Depth Study: Analyst Supply-Chain-Specific Expertise

4.1. Supply-Chain Specific Expertise and Forecast Accuracy

Our methodology can be generalized for all topics, which we will show later in Section 5. In this section, for reasons discussed in Section 2.3.1, we first focus on a motivating in-depth study by investigating whether analysts' supply-chain-specific expertise, identified based on *ex ante* Q&As in earnings conference calls of *other firms*, reduces earnings forecast errors after firms disclose supply-chain relationships with

[6] Forecast experience and horizon reported in Table 1 are consistent with Clement and Tse (2005) and Jacob, Lys, and Neale (1999), respectively.

publicly listed customer firms. Specifically, we first identify analyst supply-chain-specific expertise based on questions they raised in conference calls hosted by *other firms* (i.e., excluding the focal firm), and then introduce a supply-chain information arrival that is very unlikely to be affected by analyst expertise or any other analyst characteristics. We examine whether forecast accuracy improves more when analyst expertise coincides with the content of the arrived information. We use the following test specification to explore this issue:

$$\begin{aligned} Abs(FE)_{i,j,t} = \quad & \beta_0 + \beta_1 \times SUP_{j,t} \times Expertise_{i,j,t} + \beta_2 \times CUS_{j,t} \times Expertise_{i,j,t} + \beta_3 \times Expertise_{i,j,t} \\ & + \beta_4 \times Analyst\ Characteristics_{i,j,t} + Analyst\ FE + Firm \times Year\ FE \\ & + \varepsilon_{i,j,t} \end{aligned} \qquad (1)$$

In Equation (1), observations are organized at the analyst-firm-year level, where i indicates the analyst, j indicates the firm, and t indicates the year of the earnings forecasts. $Expertise$ is our main measure of analyst supply-chain-specific expertise, which takes the value of 1 if an analyst is identified as a supply-chain expert analyst, as discussed in Section 2.3.2. SUP is an indicator variable that takes the value of 1 if the firm has disclosed at least one public principal customer in year t. CUS is an indicator variable that takes the value of 1 if a firm is disclosed as a principal customer by at least one public firm in year t. In addition to the key independent variables mentioned above, we also include the following control variables that have known correlations with forecast accuracy in the literature, including total number of questions raised (Mayew, Sharp, and Venkatachalam, 2013), number of firms covered (Clement, 1999), number of industries covered (Clement and Tse, 2005), forecast experience (Clement, 1999), and forecast horizon (Jacob, Lys, and Neale, 1999).[7]

We include analyst fixed effects to control for time-invariant unobservable analyst heterogeneity that may affect analyst forecast accuracy. A difference in forecast accuracy could also arise from a time-varying information environment for different firms and years. We thus include Firm×Year fixed effects in all test specifications. Therefore, the test compares the forecast accuracy of the analysts who cover the

[7] We provide a comparison of these dimensions between supply-chain expert and non-expert analysts in Online Appendix Table O1, to motivate the necessity to control these variables in our tests.

same firm in the same year. In Equation (1), β_3 captures the difference in forecast errors between supply-chain expert analysts and other analysts when a firm has neither public principal customers nor dependent suppliers. β_1 captures the incremental effect of supply-chain-specific expertise on forecast errors when the firm has at least one public principal customer, which is the primary focus of this section.

[Insert Table 2 Here]

Table 2 presents the estimates from Equation (1). We use conference call transcripts of the previous two, three, and four years to the day before the earnings forecast announcement to construct the $Expertise$ variables in columns (1), (2), and (3), respectively. The estimated coefficients of $SUP \times Expertise$ in all three columns are negative, statistically significant, and similar in economic magnitude. In the interest of brevity, we will focus on the $Expertise$ measure based on earnings conference call transcripts of the previous four years hereafter. Results in column (3) suggest that, when a firm discloses important relationships with public principal customers for the first time, supply-chain expert analysts exhibit a larger reduction in forecast errors relative to that of non-expert analysts. This difference (i.e., 0.079) is equivalent to 4.05% of the average forecast errors of the full sample and, therefore, economically meaningful. The results show a negative and significant effect of supply-chain-specific expertise on earnings forecast errors for dependent suppliers, irrespective of the number of years of conference call data used in constructing the $Expertise$ variable, suggesting that analysts with supply-chain-specific expertise outperform other analysts in forecasting earnings for dependent suppliers. The robust detectability of expertise is consistent with supply-chain-specific expertise being a relatively persistent characteristic that is detectable even with a shorter observation window, e.g., a window of two or three years.

On the other hand, the coefficient on $CUS \times Expertise$ suggests that, compared to analysts without supply-chain-specific expertise, supply-chain expert analysts do not necessarily forecast more accurately for principal customers, i.e., when a firm has dependent suppliers. This result is not surprising given the asymmetric mutual importance between principal customers and dependent suppliers. For example, on average, each dependent supplier contributes less than 2% of a public principal customer's cost of goods sold (COGS).[8] The superior performance of supply-chain expert analysts is not only attributable to their expertise in processing supply-chain information, but also to the incremental economic impact (i.e., concentration) of specific supply-chain partners on future earnings.

In our Online Appendix, we provide extensive robustness checks to validate our findings in Table 2: (1) We demonstrate that our results remain unaffected when controlling for analysts' past accuracy as an alternative measure of expertise (Online Appendix Table O2). (2) To address concerns that some questions may constitute cheap talk, we measure analyst expertise by weighing questions based on the length of the managers' answers and find consistent results (Online Appendix Table O3). (3) To clearly establish that the improvement in forecast accuracy for expert analysts relative to their peers coincides with supply-chain relationship establishment, we conduct a timing test (Online Appendix Table O4), which confirms this timing alignment. (4) Our findings are robust to using a more conservative keyword list to identify supply-chain-specific expertise (Online Appendix Table O5) and alternative cutoff levels for classifying expert analysts (Online Appendix Table O6). (5) The main empirical pattern persists when using the breadth of supply-chain-related questions to identify expert analysts, which is calculated as the percentage of firms where an analyst has raised at least one supply-chain-related question in their conference calls among all firms that the analyst covers (Online Appendix Table O7). (6) Finally, our results remain unchanged when computing the number of supply-chain-specific questions at the analyst level rather than the analyst-firm level (Online Appendix Table O8).

Previous studies suggest that the benefit of an information advantage is larger for firms with a higher level of ex ante information asymmetry (e.g., Aboody and Lev, 2002). We expect that the information processing advantage will be more pronounced if the covered firms are subject to greater information asymmetry. Further, a principal customer on average accounts for 28.4% of the total sales of its dependent supplier. The concentrated sales to a few principal customers potentially increase the risk of suppliers' future earnings, especially when principal customers face a low cost of switching to other

[8] For example, Wal-Mart has more than 120 dependent suppliers in 2017. On average, each dependent suppliers contributes 0.19% of Wal-Mart's cost of goods sold (COGS).

suppliers (e.g., Intintoli, Serfling, and Shaikh, 2017; Brown, Fee, and Thomas, 2009). Therefore, we conjecture that supply-chain-specific knowledge would be more useful in forecasting the earnings of suppliers when suppliers face a higher customer switching risk, i.e., when principal customers face a lower switching cost to replace existing suppliers.

We use analyst forecast dispersion (Diether, Malloy, and Scherbina, 2002) to proxy for information asymmetry and product market fluidity (Hoberg, Phillips, and Prabhala, 2014) to measure customer-switching risk in product market competition. We repeat our baseline test in two subsamples partitioned by the information asymmetry and customer-switching risk measures.[9] Appendix Table A1 presents results based on these two cross-sectional partitions. Consistent with our predictions, the coefficients of $SUP \times Expertise$ are more negative and statistically significant for the high information asymmetry group and high customer switching risk group, relative to the other groups in the cross-sectional partitions. The difference in the coefficients for $SUP \times Expertise$ between low and high information asymmetry (customer switching risk) groups is statistically significant at the 5% level. Overall, results in Appendix Table A1 suggest that analysts' supply-chain-specific expertise is more useful when the firms that they cover face 1) a higher level of information asymmetry and/or 2) a higher level of customer switching risk.

4.2 Market Reactions to Supply-Chain Expert Analyst's Recommendation Updates

Prior research (e.g., Ivković and Jegadeesh, 2004; Chen, Cheng, and Lo, 2010; Livnat and Zhang, 2012) has examined analysts' informational roles by assessing market reactions to their forecasts and recommendations. Building on this literature, we investigate whether supply-chain expert analysts—those who issue more accurate earnings forecasts for dependent suppliers—trigger stronger market reactions when updating stock recommendations. We choose to focus on market reactions to stock recommendations rather than earnings forecasts, as earnings forecasts are much more frequent and clustered around major information events.

[9] Specifically, forecasts made for firms that rank within the top 40 percentiles of analyst forecast dispersion are classified into the high information asymmetry group, and those ranked within the bottom 40 percentiles are classified as the low information asymmetry group. The high and low customer switching risk groups are constructed similarly.

The recommendation update dataset spans the period from 2006 to 2022, aligning with the sample period of the supply-chain-specific expertise variable. A recommendation update is computed as the difference between an analyst's current and prior recommendations for a given firm. Analyst recommendations are numerically coded on a scale from 1 ("strong buy") to 5 ("strong sell"). Consequently, a negative recommendation update signifies an upgrade (e.g., from 3 to 2), whereas a positive update indicates a downgrade (e.g., from 2 to 3). Under this definition, the change in recommendation is expected to exhibit a negative correlation with stock returns around the time of the update.

To ensure that the stock returns surrounding recommendation updates are not confounded by other information events, we adopt the filtering criteria proposed by Loh and Stulz (2011), excluding all recommendation updates that meet the following conditions: (1) recommendations issued on days with multiple analyst revisions, (2) recommendations issued during days around earnings announcements, and (3) recommendations issued during days around earnings guidance.

Following the methodology of Birru, Gokkaya, Liu, and Stulz (2022), we measure the dependent variables in our test, the abnormal return of Day 0 (AR[0]) and the cumulative abnormal returns of Days 0 and 1 (CAR[0,1]), by adjusting raw returns over two event windows using the DGTW (Daniel, Grinblatt, Titman, and Wermers, 1997) characteristic-adjusted approach.[10] Specifically, we match stocks based on size, book-to-market ratio, and momentum, employing equally weighted benchmark portfolio returns, as formalized in Equation (2) below:

$$CAR[0,T] = \prod_{t=0}^{T}(1 + Ret_{i,t}) - \prod_{t=0}^{T}\left(1 + Ret_{i,t}^{DGTW}\right) \quad (2)$$

[Insert Table 3 Here]

Following Green, Jame, Markov, and Subasi (2014), we estimate the interaction effect between supply-chain-specific expertise and supply-chain relationship establishment on market reactions to

[10] Here, Day 0 is the trading day corresponding to the recommendation update if the update is issued during trading hours or the next trade day if the update is issued after trading hours.

recommendation changes for upgrades and downgrades separately. Control variables include analyst characteristics (e.g., past accuracy, firm coverage, and industry coverage) as well as recommendation characteristics, following Green, Jame, Markov, and Subasi (2014). We include analyst and firm fixed effects to control for impacts from time-invariant determinants at the analyst and firm levels. We control for year fixed effects to control for market-wide shocks on the abnormal returns of all firms. Standard errors in this test are clustered at the firm level.

Columns (1) and (2) of Table 3 report the market response to recommendation upgrades and downgrades, respectively, on the recommendation update day. The statistically insignificant coefficients of $Expertise$ in both columns indicate that, for firms without significant supply-chain relationships, recommendations from expert and non-expert analysts elicit similar abnormal returns. However, a divergence emerges when firms announce important supply-chain relationships, as evidenced by the statistically significant coefficients on $SUP \times Expertise$ in both columns. For instance, in column (1), the coefficient on $SUP \times Expertise$ is 0.452 (significant at the 1% level), implying that recommendation upgrades from supply-chain expert analysts generate, on average, 45.2 bps higher abnormal returns than those from non-experts for firms with important supply-chain relationships. Column (2) reveals a symmetric effect for recommendation downgrades. These findings remain robust when we extend the event window to [0,1], as shown in columns (3) and (4). Overall, our evidence suggests that investors at least partially recognize analysts' supply-chain-specific expertise through increased price reaction to expert analysts' recommendation updates after the firms disclose concentrated supply-chain risks.[11]

4.3. Supply-Chain-Specific Expertise in the 2018 US-China Trade War

In July 2018, the United States imposed a 25% tariff on $34 billion worth of Chinese imports, prompting China to retaliate with tariffs on an equivalent value of US goods. The trade war intensified throughout the second half of 2018 and the first half of 2019. By the end of 2019, the U.S. had levied tariffs

[11] In Appendix Table A6, we test whether investors fully recognize analysts' supply-chain-specific expertise. Specifically, we examine whether the earnings announcement returns are correlated with the forecast difference between expert and non-expert analysts. We find that investors are more likely to be positively surprised by the earnings news (leading to more positive announced returns) when supply-chain expert analysts are more optimistic than non-experts for firms with concentrated supply-chain risks. This result is consistent with the conjecture that investors do not fully recognize analyst supply-chain-specific expertise before the earnings announcements.

on over $350 billion worth of Chinese imports, while China retaliated with tariffs on more than $100 billion in U.S. exports. The 2018 US-China Trade War represented the most significant market-wide disruption to global supply chains prior to the COVID-19 pandemic and the broader 2025 trade tensions between the US and other major economies. In this subsection, we first leverage this exogenous market-wide supply-chain shock—unrelated to the decisions of individual firms or financial analysts—to examine whether supply-chain expert analysts produced more accurate forecasts than their peers following the outbreak of the 2018 US-China Trade War.[12]

In this test, we augment our baseline specification in Equation (1) by introducing a triple interaction term, $ChinaTrade \times Post \times Expertise$, along with its corresponding double interaction terms. To measure the exposure of US firms to the trade war, we use US import and export data from the S&P Panjiva database. We define $ChinaImport$ ($ChinaExport$) as a dummy variable equal to 1 if a US firm imports from (exports to) China in year *t*. $ChinaTrade$ is a dummy variable equal to 1 if a US firm engages in either export or import trade with China in year *t*. The dummy variable, *Post*, equals 1 for observations occurring after the outbreak of the 2018 US-China trade war. The estimation results are presented in Table 4.

[Insert Table 4 Here]

In column (1) of Table 4, the coefficient on $ChinaImport \times Expertise$ is statistically indistinguishable from zero, indicating that prior to the trade war—when supply-chain risks were less salient—supply-chain expert analysts did not produce more accurate forecasts than their peers, even for firms importing from China. The statistically insignificant coefficient on $Post \times Expertise$ suggests that, for firms without trade ties to China, expert analysts do not outperform their peers after the outbreak of the trade war. However, the negative and statistically significant coefficient on $ChinaImport \times Post \times Expertise$ implies that expert analysts significantly improved their forecast accuracy relative to non-

[12] Previous studies suggest that investors are aware of public firms' exposure to the US-China Trade War. For example, Huang, Lin, Liu, and Tang (2023) demonstrate that U.S. firms that rely more heavily on exports to and imports from China experienced a greater decline in market value following tariff announcements.

experts for US firms importing from China after the 2018 US-China trade war began. Columns (2) and (3) demonstrate that this empirical pattern persists when we examine US firms exporting to China or firms engaged in either import or export trade with China.

Since the trade war has significantly increased the value of supply-chain-specific expertise in capital markets, we next examine whether the labor market—particularly employers—similarly revalues such expertise following the trade war. Job turnovers, which can result from either dismissals or headhunting, often offer limited and ambiguous insight into career advancement (Cen, Chen, Ornthanalai, and Schiller, 2021). To address this concern, we leverage detailed analyst career data from I/B/E/S and Revelio Labs to investigate three key labor-market outcomes after the onset of the 2018 US-China trade war: (1) conditional on remaining with their current employers, whether supply-chain expert analysts are more likely to receive internal promotions relative to their peers; (2) conditional on job turnover, whether expert analysts are more likely to move to buy-side institutions (generally viewed as the more lucrative career path); and (3) conditional on moving to another sell-side brokerage, whether expert analysts are more likely to join a larger brokerage firm than their previous employer. Estimation results are reported in Table 5.

[Insert Table 5 Here]

In column (1) of Table 5, we examine whether supply-chain expert analysts are more likely to receive internal promotions after the 2018 US-China Trade War if they remain with their current employers. In this test, we focus on a subsample of analysts who have no affiliation changes in the next two years. As defined in Section 3.2.3, the dependent variable, *Internal Promotion*, equals one if an analyst's job title becomes more senior in the next two years while their affiliation remains unchanged.[13] The positive and statistically significant coefficient on $Post \times Expertise$ suggests that supply-chain expert analysts experience a greater increase in internal promotion likelihood following the trade war compared to non-expert analysts. In column (2), we focus on analysts who switch employers. Prior research (e.g., Cen, Chen,

[13] For instance, when Tyler Batory of Janney Montgomery Scott LLC was promoted from Vice President to Director of Equity Research in 2019, *Internal Promotion* equals 1 for both 2017 and 2018 (the two preceding years).

Ornthanalai, and Schiller, 2021) shows that sell-side analysts moving to buy-side institutions tend to issue more accurate earnings forecasts, implying these transitions are often voluntary or driven by headhunting. We test whether supply-chain experts are more likely to move to buy side institutions after the trade war begins. The results support this conjecture: the coefficient on $Post \times Expertise$ indicates that expert analysts experienced a 6.8 percentage-point higher increase in the probability of joining buy-side institutions than non-experts after the trade war began. This effect is economically significant, given that the average probability of such a move is 7.26% among sell-side analysts undergoing job turnover.

For analysts transitioning to other brokerage firms, we define career advancement following Hong, Kubik, and Solomon (2000), i.e., whether the analyst moves to a larger brokerage firm (measured by analyst headcount). Column (3) examines this subsample and finds that, conditional on moving to other sell-side brokerage firms, expert analysts see a 14.4 percentage-point higher increase in the probability of joining larger brokerage firms post-trade war compared to non-experts. This translates to a 34% increase based on the average probability of moving to a larger brokerage firm, conditional on the analyst remaining within the sell-side industry. Overall, we find consistent evidence that supply-chain expert analysts experience more favorable career outcomes, whether through internal promotions or external job transitions, compared to non-expert peers following the 2018 US-China Trade War.

4.4 Robustness Checks Based on Alternative Information Sources

In Section 2.2, we elaborate on the main advantages of using earnings conference call questions to identify analysts' topic-specific expertise. While this method could also be applied to other recurring analyst outputs, such as research reports, each medium has distinct strengths and limitations. Compared to conference call questions, analyst reports offer certain advantages. For example, analysts typically devote more time to crafting reports than to preparing call questions. Additionally, while only a subset of analysts are selected to ask questions during earnings calls, all analysts publish their reports. However, analyst reports also present notable weaknesses. They are often produced after group discussions and may be co-authored, complicating the attribution of expertise to individual analysts. Furthermore, reports are usually unrestricted in length, allowing analysts to cover multiple topics comprehensively. This potentially makes it empirically difficult to pinpoint the key topics that reflect an analyst's expertise.

We replicate all tests presented in Tables 2–5 utilizing analyst topic-specific expertise derived from analyst reports. Specifically, to construct supply-chain-specific expertise using analyst reports, we follow an approach analogous to the methodology described in Section 2.3.2. We identify supply-chain-related analyst reports by searching for supply-chain keywords in the title and body of each report. For each analyst, we count the number of supply-chain-related reports they have issued. We then link analysts from these reports to I/B/E/S analysts by matching names and brokerage affiliations. For each year, we rank I/B/E/S analysts based on the number of supply-chain-related reports they have issued over the past four years. Analysts whose counts rank above the 90^{th} percentile in the annual distribution are identified as possessing supply-chain expertise. The estimation results, presented in Appendix Table A2-A5, confirm that the key empirical patterns remain consistent. These findings indicate that our methodology is robust to using alternative information sources, as long as they stem from repeated and credible informational events.

5. Generalization: How Can Question-Based Expertise Measure Be Applied to All Topics?

5.1. Conditions of Generalization

Theoretically, our approach can be generalized to assess topic-specific expertise for any topic across various market participants (e.g., corporate or fund managers). However, implementing this question-based method empirically requires three key conditions. First, the topic must be both relevant and important to investors, yet remain in an "unsaturated" information market—where analysts and investors have not yet fully incorporated (or are not fast enough to incorporate) the information into their decision-making. Second, the arrival of topic-specific information should not be entirely controlled or heavily influenced by key information processors, such as expert analysts in our context. Third, sufficient and credible textual data (e.g., questions in earnings conference calls) must be available to enable accurate identification of topic-specific expertise through machine learning or other algorithmic techniques.

5.2. Implementations for Other Pre-Specified Topics

Before generalizing our methodology to all major topics, we provide further evidence in this section that our approach can be applied to other settings where topics are defined ex ante. As discussed in Section 5.1, we first pre-specify a topic relevant to earnings and prices that analysts may develop expertise in. Following the procedure outlined in Section 2, we then identify expert analysts for this topic, as indicated by a dummy variable, *Expertise*. Next, we identify an exogenous information event—exogenous to analysts in our setting—that significantly elevates the topic's importance in the market, captured by a dummy variable, *Information Arrival*. Using the test specification from Equation (1), we examine whether analysts with topic-specific expertise exhibit a more pronounced improvement in forecast accuracy during and after the information arrival, as reflected in the coefficient of the interaction term *Expertise* × *Information Arrival*.

In this experiment, we pre-specify four increasingly important topics: environmental, product-related, labor-related, and macroeconomic-related issues. For each topic, we identify a corresponding information event that increases its relevance to earnings forecasts. For environmental issues, *Information Arrival* is set to 1 for Californian firms after the implementation of the state's cap-and-trade program in 2013. For product-related issues, *Information Arrival* is set to 1 for firm-years with any product-related announcements, including new, modified, or discontinued products. For labor-related issues, *Information Arrival* is set to 1 for firms located in states that adopt right-to-work laws. Finally, for macroeconomic-related issues, *Information Arrival* is set to 1 in years with more than 100 days on which the daily macroeconomic surprises index (Scotti, 2016) exceeds its average daily value. Control variables and fixed effects are identical to those reported in Table 2.

Results of these tests are presented in Appendix Table A7. The coefficients of *Expertise* × *Information Arrival* are all negative, ranging from -0.109 to -0.058, and statistically significant. These results show that our approach can be generalized to other pre-specified topics, including environmental, product, labor, and macroeconomic-related ones. In the next section, we will discuss the generalization of our approach for emerging topics that are not prespecified.

5.3. For Non-Prespecified Topics: When Expertise Meets Information Arrival

5.3.1. Our Approach

In this section, we propose a generalized method for identifying experts in emerging topics, relaxing the conditions outlined in Section 5.1 and setting aside endogeneity concerns (which are often secondary for practitioners). In our first step, we translate the text of analyst questions into a form that a computer can interpret and process for analyzing the content of analyst questions. In practice, this means representing each question as a vector in a multi-dimensional space, where the position reflects the question's meaning. In this space, questions with similar meanings are located closer together, and those with different meanings are farther apart. This numerical representation, known as an *embedding*, is a standard way for modern language models to capture sentence meaning so that a machine can compare and group text based on its semantics. In particular, we extract the underlying meaning of each analyst question by converting it into a multi-dimensional vector using BERT, a deep neural network-based natural language processing model developed by Devlin, Chang, Lee, and Toutanova (2019). BERT is a deep learning model for language that provides context-sensitive text representations. It considers each word in a question *in relation to all others*, rather than processing words in a fixed sequence. This bidirectional context modeling allows BERT to capture subtle differences in meaning. Compared with earlier models such as Word2Vec, which assign a single static vector to each word regardless of context, BERT generates dynamic embeddings that reflect how a word's meaning changes with surrounding words. For example, in earnings conference calls, the word *"color"* in the phrase *"give some color on the quarter"* means "additional insight," which is entirely different from its everyday use as a visual attribute. BERT distinguishes between these meanings by interpreting the surrounding words in each case, making it well-suited for our purposes. Another key advantage is that BERT comes pre-trained on massive text corpora (e.g. books and Wikipedia), so it can be applied to new tasks with relatively little data via transfer learning.[14] In our setting, we feed

[14] Note, our methodology does not rely on a particular topical model (e.g., BERT and HDBSCAN in our case). Other models that can accurately identify topics should also work in our methodology.

each analyst question through BERT's language model to obtain a vector representation (embedding) of the question. After this embedding step, each analyst question is represented by a vector in the 384-dimensional space.[15]

In our second step, we group the question embeddings into topics. Each embedding is a representation of a question, where questions with similar meaning are located closer together in this representation space. By grouping together embeddings that are close to one another, we can identify collections of questions that fall under the same underlying topic. Specifically, we use HDBSCAN (Campello, Moulavi, and Sander, 2013) to identify topics. This method looks for "pockets" of questions that are closely related in meaning, allowing the data to form topics that naturally vary in shape and size rather than conforming to a fixed, uniform pattern. HDBSCAN works by starting with the most similar questions and gradually widening the scope to include less similar ones, tracking which groupings remain stable across different levels of similarity. Stable groupings are taken as the final topics. Compared with methods such as *k*-means, which impose a uniform cluster shape and size, or the original DBSCAN, which uses a single distance threshold to decide whether two items belong in the same group, HDBSCAN can capture both compact, tightly worded topics and more diverse ones without forcing them into a rigid structure. This flexibility is important because analysts' questions vary widely: some topics are expressed in consistent ways, while others cover more specialized content or appear in a broader range of phrasing.

Third, we assign each question to the topic cluster with which it has the highest cosine similarity. After retrieving the non-prespecified topic label for each analyst question, we apply the methodology described in Section 2.3.2 to determine analyst topic-specific expertise. For each analyst–topic pair, we count the number of questions the analyst has asked on that topic over the preceding four years, excluding those posed during the focal firms' own earnings conference calls. Within each topic and year, analysts are then ranked by these question counts. Consistent with our approach to identify supply-chain-specific

[15] We reduce the 384-dimensional BERT embeddings using Uniform Manifold Approximation and Projection (UMAP) to make clustering computationally feasible (see Appendix IV for details). This reduction preserves the most important information and maintains the semantic proximity captured by BERT, ensuring that questions with similar meanings remain close together in the reduced space.

experts, we define topic-specific experts as those analysts whose number of questions on a given topic ranks above the 90th percentile.

More detailed steps are provided in Appendix IV. Next, mirroring our baseline approach, we assign topic labels to each question using the trained topic model and identify expert analysts based on the topic-specific questions they posed in conference calls over the past four years. We focus on the "hot" topics, defined as those most frequently raised in analyst questions during earnings conference calls. Specifically, we focus on the top 20 topics with the highest frequency. *Has Expertise* is a dummy variable that is equal to one if an analyst has any topic-specific expertise related to the top topics. Subsequently, we define topic-specific information arrivals by detecting the "peak" of each topic for a given firm. Specifically, for each firm–topic–year, the peak information arrival period is defined as the year when the fraction of questions related to a given topic in the firm's earnings calls exceeds the sample-wide mean share of that topic by more than two standard deviations. Finally, we construct a dummy variable, *Coincide*, which equals one whenever one of an analyst's topic-specific expertise aligns with the peak of topic-specific information arrivals.[16]

5.3.2. Topic-Specific Expertise, Information Arrival, and Forecast Accuracy

We include both *Coincide* and *Has Expertise* in the test specification of Equation (1).[17] Our focus in this test is whether *Coincide*, i.e., when topic-specific expertise aligns with the type of information arrivals, leads to an improvement in forecast accuracy relative to peers. Estimation results are reported in Table 6.

[Insert Table 6 Here]

Results in columns (1)–(3) are based on the Top 10, 15, and 20 topics identified in our sample. All coefficients for *Coincide* in Table 6 are negative and statistically significant, indicating that expert analysts achieve markedly higher forecast accuracy when their specialized knowledge aligns with periods of

[16] When an analyst has only one topic-specific expertise, *Coincide* is equivalent to the interaction term *Expertise* × *Information Arrival* in Section 5.2.
[17] In this test, *Peak* will be subsumed by the firm×year fixed effects.

heightened topic-specific information arrival. Notably, the economic magnitude of *Coincide*'s coefficients is comparable to that of *SUP×Expertise* in Table 2, reinforcing the robustness of our findings.

5.3.3. Topic-Specific Expertise, Information Arrival, and Capital Market Reaction

We next generalize our tests in Table 3 and examine whether investors would pay more attention to recommendation updates made by analysts whose topic-specific expertise coincides with the information arrival. We also augment the test specification in Table 3 by including *Coincide*, *Has Expertise*, and *Peak* into control variables. *Peak* is defined at the firm-year level and equals one if any of the top 20 topics is in a peak information arrival period for that firm in year *t*. Estimation results are reported in Table 7.

[Insert Table 7 Here]

We find that the coefficients for *Has Expertise* in Table 7 are statistically indistinguishable from zero, indicating that analysts with topic-specific expertise do not trigger stronger capital market reactions than their peers when the information arrivals are unrelated to their expertise. In contrast, the statistically significant coefficients for *Coincide* across all specifications suggest that the supply-chain-specific expertise effects documented in Table 3 are generalizable to a broader, all-topic setting. For instance, the coefficient for *Coincide* in column (4) (−0.455, statistically significant at the 5% level) implies that an analyst's recommendation downgrade triggers a two-day ([0,1]) cumulative abnormal return that is 0.455 percentage points lower than that of peers when the information aligns with her topic-specific expertise. This pattern holds robustly for both upgrades and downgrades. These results further support our conjecture that investors consider the topic-specific expertise of information providers when assigning weights to different signals.

5.3.4. Topic-Specific Expertise, Information Arrival, and Labor Market Recognition

At the end of this section, we also generalize the labor market recognition tests to a general setting. Similar to previous tests in this section, we augment the test specification in Table 5 by including *Coincide*, *Has Expertise*, and *Peak of Covered Firms* as additional independent variables. *Peak of Covered Firms* is defined at the analyst–year level. It is a dummy variable that equals one if any firm covered by the analyst experiences a peak information arrival period in any of the top 20 topics during year *t*. Estimation results are reported in Table 8.

[Insert Table 8 Here]

Mirroring the patterns in Table 5, the results in Table 8 show consistent effects of expertise-information alignment across career transitions. For analysts staying with their current employer in column (1), the significant positive coefficient on *Coincide* indicates that analysts experience an increase in internal promotion likelihood when their topic-specific expertise aligns with the arrival of new information. Likewise, the patterns observed in columns (2) and (3) of Table 5 persist when analysts experience job turnovers. Specifically, conditional on changing employers, analysts are more likely to transition to buy-side institutions when their topic-specific expertise aligns with the arrival of new information. Similarly, when moving to another sell-side brokerage firm, they are more likely to join larger firms if their expertise aligns with prevailing information trends. Notably, the economic magnitude of the coefficients for *Coincide* is smaller than those in Table 5, likely because the 2018 US-China Trade War represents a more influential and persistent information event compared to a typical "peak" of new information arrivals in an all-topic setting. Overall, these findings suggest that the labor market values topic-specific expertise when it aligns with the theme of new information arrival.

6. Implications: Analyst Expertise Diversity, Price Efficiency, and Real Effects

The identification of financial analysts' topic-specific expertise can generate important economic implications. One key application is the development of a metric for expertise diversity among analysts covering a particular stock, distinguishing between concentrated and diversified distributions of expertise. This empirically constructed measure facilitates direct testing of how diversity in analyst expertise influences capital market dynamics.

Specifically, in this section, we investigate how analyst expertise diversity affects price efficiency and whether it has real effects. Price efficiency refers to how well prices reflect future cash flow information and predict the future value of securities (which is related to forecasting price efficiency, or FPE, in Bond, Edmans, and Goldstein (2012)). The real effects refer to the fact that managers learn from prices to guide their real decision (e.g., investment) making (which is related to revelatory price efficiency, or RPE, in Bond, Edmans, and Goldstein (2012)).

We construct two measures to capture analyst expertise diversity for a given stock by examining the top 20 most frequently discussed topics in our data. Each topic represents a distinct analyst expertise. The first measure, *# of Expertise*, counts the number of distinct expertise possessed by the analysts covering the firm. For example, consider a firm *i* covered by three analysts, A, B, and C. Suppose analyst A has expertise in Topics 1, 3, and 4; analyst B in Topics 4 and 10; and analyst C in Topics 3 and 7. In this case, *# of Expertise* for firm *i* equals 5 (i.e., Topics 1, 3, 4, 7, and 10). To mitigate the concern that *# of Expertise* may be inflated by a single analyst contributing expertise across multiple topics, we introduce a second measure, *# of Rep Expertise*, which restricts each analyst to a single expertise based on their most frequently discussed topic. In the same example, if analysts A and B most frequently discuss Topic 4, and analyst C most frequently discusses Topic 3, then *# of Rep Expertise* equals 2 for firm *i*. By construction, *# of Expertise* can exceed the number of analysts covering a firm, whereas *# of Rep Expertise* cannot—it is either equal to or lower than the number of covering analysts.

6.1. Expertise Diversity and Price Efficiency

Both theoretical and empirical studies consistently support the idea that information diversity—whether arising from heterogeneous information sources or the diverse backgrounds of information processors—enhances forecast accuracy and price efficiency. Goldstein and Yang (2015) provide theoretical evidence that information diversity improves the ability of stock prices to reflect future cash flows.[18] Empirically, Gerken and Painter (2023) demonstrate that information diversity triggered by diverse geographical locations of analysts mitigates location-based forecast biases, reducing consensus forecast

[18] In Goldstein and Yang (2015), price efficiency is captured by the inverse of the variance of future cash flows conditional on market prices.

errors and improving the predictive power of stock prices for future earnings. Similarly, Fang and Hope (2021) find that analyst forecast accuracy increases with analyst team diversity in sell-side experience, educational background, and gender. Building on this established literature, we posit that information diversity stemming from expertise diversity of key information processors, i.e., analysts in our setting, should likewise enhance forecast accuracy and price efficiency in the capital market.

We begin by testing whether analyst expertise diversity is negatively associated with consensus forecast errors. Our analysis employs a firm-year panel dataset, where the dependent variable is the absolute consensus forecast error measured at the end of each fiscal year. Alongside our key independent variables—the two measures of analyst expertise diversity—we control for the number of analyst coverage, natural logarithm of total assets, leverage, tangibility, cash holdings, Tobin's Q, ROA, ROA volatility, average past returns, and returns volatility. The regression incorporates both year and firm fixed effects, with standard errors clustered at the firm level. The estimation results, presented in Appendix Table A8, indicate that an increase in analyst expertise diversity is associated with a statistically significant reduction in consensus forecast errors. This finding holds consistently across both measures of expertise diversity.

We next test the impact of analyst expertise diversity on price efficiency by augmenting the test specification in Bai, Philippon, and Savov (2016) and Kacperczyk, Sundaresan, and Wang (2021). Specifically, we estimate the following regression:

$$\begin{aligned} Earnings_{i,t+1} = \quad & \beta_0 + \beta_1 \times LogMVA_{i,t} \times Expertise\ Diversity_{i,t} + \beta_2 \times LogMVA_{i,t} \\ & + \beta_3 \times Expertise\ Diversity_{i,t} + \sum_{j=1}^{n} \gamma_j \times LogMVA_{i,t} \times Other\ Firm\ Controls_{i,t} \\ & + \sum_{j=1}^{n} \delta_j \times Other\ Firm\ Controls_{i,t} + Firm\ FE + Year\ FE + \varepsilon_{i,t} \qquad (3) \end{aligned}$$

where *i* denotes the firm and *t* denotes the year. The dependent variable, $Earnings_{i,t+1}$, represents the earnings of firm *i* in year *t+1*, scaled by the total assets for fiscal year *t*. The independent variable, $LogMVA_{i,t}$, is the natural logarithm of the ratio between the firm's market capitalization and total assets in year *t*, representing the stock prices in that year. $Expertise\ Diversity_{i,t}$ represents one of our analyst expertise diversity measures for firm *i* in year *t*. Firm-level control variables include the number of analysts covering the firm (*Num Analysts*), the natural logarithm of total assets (*Ln(Assets)*), total liability divided by total assets (*Leverage*), net property, plant, and equipment divided by total assets (*Tangibility*), cash divided by total assets (*Cash*), sales divided by total assets (*Sale*), R&D expenditure scaled by total assets (*R&D*), and current earnings level (*Current Earnings*). We include firm and year fixed effects in all test specifications to control for firm-specific characteristics that remain constant over time and for overall time trends. Standard errors are clustered at the firm level.

In this test specification, β_2 captures the average level of price informativeness for future earnings. The primary coefficient of interest is β_1, which captures the effect of analyst expertise diversity on the ability of current prices to reflect future earnings. A positive and statistically significant $\beta_1$would suggest that a higher analyst expertise diversity improves price informativeness for future earnings and vice versa.

[Insert Table 9 Here]

Estimation results of Equation (3) are reported in Table 9. β_1, the coefficient for the interaction terms of $LogMVA_{i,t} \times Expertise\ Diversity_{i,t}$, is positive and statistically significant for both analyst expertise diversity measures. For example, in column (1) of Table 9, the economic magnitude of β_1 (0.039, statistically significant at the 1% level) suggests that a one-standard-deviation increase in *# of Expertise* is associated with a 19.25% increase from the average price informativeness reflected in β_2.[19] The main empirical patterns remain robust in column (2) when analyst expertise diversity is measured by *# of Rep Expertise*. Overall, our empirical evidence is consistent with the prediction in Goldstein and Yang (2015) that information diversity, triggered by expertise diversity of key information processors, improves price efficiency of reflecting future cash flows or earnings.

[19] The standard deviation of *# of Expertise* is 4.393.

6.2. Expertise Diversity and Real Effects

While analyst expertise diversity improves price efficiency, can managers learn more from prices for their real decision making? The impact of analyst expertise diversity on the real effects in financial markets is far more complicated than its impact on price efficiency. Goldstein and Yang (2019) theoretically show that, in the presence of information from multiple dimensions, whether a public information disclosure improves managers' learning from prices depends on which dimension of information is disclosed. Specifically, they suggest that public revelation of information on the dimensions that firms already know would encourage traders to put a higher weight on their own private information that firms want to learn and, therefore, improve managers' learning from prices. However, if the publicly revealed information is closely related to what managers want to learn from traders, this may crowd out private information being incorporated into price and reduce managers' ability of learning from prices.

The theoretical framework of Goldstein and Yang (2019) suggests that whether analyst expertise diversity improves or hinders real efficiency is an empirical question. A positive effect prevails if increased expertise diversity enhances the public production of types of information in which managers hold a comparative advantage, which encourages investors to put a higher weight on their own private information (i.e., the type of information that managers want to learn) in trading. A negative effect arises if the increase in analyst expertise diversity mainly improves the public production of the type of information that managers want to learn from traders. While public signals of this type of information become more accurate, investors are likely to place a lower weight on their own private information in trading ex post, which also reduces their incentives to become privately informed about that information ex ante. This may lead to a reduction in overall price informativeness, thereby decreasing managers' ability to learn from prices for their real decision-making.

While it is empirically challenging to measure real efficiency directly (since researchers cannot observe all investment opportunities available to managers), we focus on two observable dimensions of the real effects in this section: buy-side information production activities and investment sensitivity to stock price (i.e., the price feedback effect). First, we examine whether greater diversity in sell-side analyst expertise is more likely to crowd out information production by buy-side investors. Buy-side participation in earnings conference calls provides a unique setting in which buy-side information production—and the subsequent price reactions—can be publicly observed. Cen, Ragunathan, Xiong, and Yang (2018) and Jung, Wong, and Zhang (2018) find that buy-side participation generates significant instantaneous and post-call price movements, consistent with the notion that such participation reflects instances where buy-side investors' private information is impounded into stock prices. Further, Jung, Wong, and Zhang (2018) show that while buy-side analysts are not required to ask questions during earnings calls, they are more likely to participate when sell-side coverage is sparse and earnings forecast dispersion is high.

We use buy-side participation in earnings calls as a proxy for information production of buy-side institutions and examine whether analyst expertise diversity discourages buy-side information production in this setting. Our testing sample is organized at the firm-year level, after aggregating all earnings conference calls for a given firm within a single year into one observation. To measure buy-side institutions' information production, we employ two dependent variables: *BuySideQuestions%* and *BuySideParticipants%*. *BuySideQuestions%* represents the proportion of questions raised by buy-side institutions relative to the total questions asked in earnings calls within a firm-year. Similarly, *BuySideParticipants%* captures the percentage of buy-side participants among all attendees in these calls. Our key independent variables are the two measures of analyst expertise diversity. We control for the number of analysts covering the firm and various firm characteristics, including size, leverage, asset tangibility, Tobin's Q, cash holdings, profitability, and past returns. The estimation results are presented in Table 10.

[Insert Table 10 Here]

The results in Table 10 indicate a statistically significant negative correlation (at the 1% level) between analyst expertise diversity and buy-side participation in earnings conference calls. For instance, the coefficient of *# of Expertise* in column (1) is −0.052, implying that a one-standard-deviation increase in this variable corresponds to a decrease of 0.23% in *BuySideQuestions%*, equivalent to a 14.85% decline relative to its mean value. These findings remain consistent regardless of whether buy-side participation is measured by the number of questions or the number of participants. Overall, our results reported in Table 10 suggest that analyst expertise diversity crowds out buy-side information production in earnings conference calls.

After demonstrating that analyst expertise diversity reduces buy-side information production activities, we next investigate its impact on investment sensitivity to stock prices, serving as a proxy for managers' learning from stock prices. We augment the test specification in Chen, Goldstein, and Jiang (2007) and estimate the following regression in Equation (4):

$$\begin{aligned} I_{i,t+1} = \quad & \beta_0 + \beta_1 \times Q_{i,t} \times Expertise\ Diversity_{i,t} + \beta_2 \times Q_{i,t} + \beta_3 \times Expertise\ Diversity_{i,t} \\ & + \gamma_1 \times CF_{i,t+1} \times Expertise\ Diversity_{i,t} + \gamma_2 \times CF_{i,t+1} \\ & + \textstyle\sum_{j=1}^{n} \delta_j \times Other\ Firm\ Controls_{i,t} + Firm\ FE + Year\ FE + \varepsilon_{i,j,t} \end{aligned} \quad (4)$$

where the dependent variable, $I_{i,t+1}$, is firm *i*'s investment in year *t+1*. Following Chen, Goldstein and Jiang (2007), we use three different measures to capture $I_{i,t+1}$, including asset growth rate (percentage changes in book assets), investment (capital expenditure plus R&D expense scaled by the beginning-of-year total assets), R&D expenditure (R&D expense scaled by the beginning-of-year total assets). Among key independent variables, $Q_{i,t}$ is firm *i*'s Tobin's Q in year *t*, a normalized measure for stock prices of firm *i*. $Expertise\ Diversity_{i,t}$ represents one of our analyst expertise diversity measures for firm i in year *t*. Other firm-level controls include the number of analysts covering the firm of year *t* (# of $Analysts_{i,t}$), the interaction term between Tobin's Q and the number of analysts covering the firm ($Q_{it} \times \#\ of\ Analysts_{i,t}$), the inverse of total assets of year *t* ($1/ASSETS_{i,t}$), sales growth rate (*SG*), cash (*Cash*), leverage (*Leverage*), return on assets (*ROA*), cash flows of year *t+1* ($CF_{i,t+1}$) and future returns of three years ($Ret_{i,t+3}$). We also include the interactions between analyst expertise diversity measures and $CF_{i,t+1}$ in our test specification. Year and firm-fixed effects are included in all test specifications. Standard errors are clustered at the firm level. Estimation results are reported in Table 11.

[Insert Table 11 Here]

In columns (1)–(3) of Table 11, we measure expertise diversity using *# of Expertise*. Consistent with prior literature, the coefficients of *Q* (β_2) are positive and statistically significant at the 1% level across all three dependent variables capturing future investment. This positive investment-to-price sensitivity suggests that managers incorporate information from market prices when making investment decisions. However, the coefficients of the interaction term, *Q* × *# of Expertise*, are negative and statistically significant in columns (1)–(3), indicating that investment-to-price sensitivity is muted as analyst expertise diversity increases. To illustrate, consider the results in column (1): the coefficient of *Q* is 1.003, while that of *Q* × *# of Expertise* is −0.050. These estimates imply that a one-unit increase in *# of Expertise* reduces R&D investment-to-price sensitivity by 4.99%. The economic magnitude is even more pronounced when we measure expertise diversity using *# of Rep Expertise* in columns (4)–(6). For instance, in column (4), a one-unit increase in *# of Rep Expertise* corresponds to a 13.76% decline in R&D investment-to-price sensitivity.

Our findings support the conjecture that greater analyst expertise diversity discourages investors from producing private information that managers want to learn from. As a result, increased expertise diversity weakens managers' reliance on price signals and mutes the learning effect. Overall, our results in this section reveal a divergence in the impact of analyst expertise diversity on market efficiency: while it enhances price efficiency, it simultaneously reduces buy-side information production and weakens the price feedback effect, which may potentially hurt the real efficiency in capital markets.

7. Conclusion

This paper introduces a novel methodology for evaluating the topic-specific expertise of financial analysts by measuring the frequency of topic-related questions they ask during earnings conference call Q&A sessions. We first demonstrate this methodology through an in-depth study of supply-chain-specific expertise, leveraging exogenous events such as the 2018 US-China trade war for identification. Analysts who frequently ask supply-chain-related questions are identified as supply-chain experts. The results show that these experts generate more accurate forecasts for firms exposed to concentrated supply-chain risks, and their stock recommendations elicit stronger market reactions. Furthermore, following the trade war shock, supply-chain expert analysts were more likely to be promoted, move to buy-side institutions, or join larger brokerage firms.

We generalize this approach to any topic, using machine learning algorithms (e.g., BERT and HDBSCAN) to cluster questions and identify "hot" topics. We find the same patterns persist in the generalized setting: when an analyst's expertise coincides with a peak in information arrival on that topic, she demonstrates higher forecast accuracy, greater market influence, and better labor market outcomes relative to non-expert peers.

Finally, the paper examines the economic implications of identified topic-specific expertise. We construct a measure of analyst expertise diversity based on the identified topic-specific expertise of analysts covering a stock. When a firm is covered by analysts with heterogeneous expertise (i.e., when analyst expertise diversity is higher), price efficiency improves. However, the increase in public information from analysts with diverse expertise appears to reduce buy-side participation in conference calls and weaken the sensitivity of corporate investment to stock prices.

References

ABOODY, DAVID, and BARUCH LEV. "Information Asymmetry, R&D, and Insider Gains." *Journal of Finance* 55 (2002): 2747–2766.

AINLEY, MARY, SUZANNE HIDI, and DAGMAR BERNDORFF. "Interest, Learning, and the Psychological Processes That Mediate Their Relationship." *Journal of Educational Psychology* 94 (2002): 545.

ASHER, STEVEN R. "Topic Interest and Children's Reading Comprehension." In *Theoretical Issues in Reading Comprehension*, 525–34. Routledge, 1980.

ASHER, STEVEN R, SHELLEY HYMEL, and ALLAN WIGFIELD. "Influence of Topic Interest on Children's Reading Comprehension." *Journal of Reading Behavior* 10 (1978): 35–47.

ASHER, STEVEN R, and RICHARD A MARKELL. "Sex Differences in Comprehension of High-and Low-Interest Reading Material." *Journal of Educational Psychology* 66 (1974): 680.

BAE, KEE-HONG, RENÉ M. STULZ, and HONGPING TAN. "Do Local Analysts Know More? A Cross-Country Study of the Performance of Local Analysts and Foreign Analysts." *Journal of Financial Economics* 88 (2008): 581–606.

BAI, JENNIE, THOMAS PHILIPPON, and ALEXI SAVOV. "Have Financial Markets Become More Informative?" *Journal of Financial Economics* 122 (2016): 625–654.

BANERJEE, SHANTANU, SUDIPTO DASGUPTA, and YUNGSAN KIM. "Buyer-Supplier Relationships and the Stakeholder Theory of Capital Structure." *Journal of Finance* 63 (2008): 2507–2552.

BIRRU, JUSTIN, SINAN GOKKAYA, XI LIU, and RENÉ M. STULZ. "Are Analyst Short-Term Trade Ideas Valuable?" *Journal of Finance* 77 (2022): 1829–1875.

BOND, PHILIP, ALEX EDMANS, and ITAY GOLDSTEIN. "The Real Effects of Financial Markets." *Annual Review of Financial Economics* 4 (2012): 339–360.

BRADLEY, DANIEL, SINAN GOKKAYA, and XI LIU. "Before an Analyst Becomes an Analyst: Does Industry Experience Matter?" *Journal of Finance* 72 (2017): 751–792.

BROWN, DAVID T, C EDWARD FEE, and SHAWN E THOMAS. "Financial Leverage and Bargaining Power with Suppliers: Evidence from Leveraged Buyouts." *Journal of Corporate Finance* 15 (2009): 196–211.

CAMPELLO, RICARDO J. G. B., DAVOUD MOULAVI, and JÖRG SANDER. "Density-Based Clustering Based on Hierarchical Density Estimates." In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 160–172. Springer, 2013.

CEN, LING, FENG CHEN, YU HOU, and GORDON D RICHARDSON. "Strategic Disclosures of Litigation Loss Contingencies When Customer-Supplier Relationships Are at Risk." *Accounting Review* 93 (2018): 137–159.

CEN, LING, JING CHEN, CHAYAWAT ORNTHANALAI, and CHRISTOPH M SCHILLER. "Navigating Wall Street: Career Concerns and Analyst Transitions from Sell-Side to Buy-Side." *Working Paper*. (2021)

CEN, LING, MICHAEL G. HERTZEL, and CHRISTOPH M SCHILLER. "Speed Matters: Limited Attention and Supply-Chain Information Diffusion." *Management Science* (2024) Forthcoming

CEN, LING, VANITHA RAGUNATHAN, YAN XIONG, and LIYAN YANG. "The Dynamics of Information Production and Diffusion: Evidence from Buy-Side Participation in Earnings Conference Calls." *Working Paper.* (2018).

CHEN, XIA, QIANG CHENG, and KIN LO. "On the Relationship between Analyst Reports and Corporate Disclosures: Exploring the Roles of Information Discovery and Interpretation." *Journal of Accounting and Economics* 49 (2010): 206–226.

CHEN, QI, ITAY GOLDSTEIN, and WEI JIANG. "Price Informativeness and Investment Sensitivity to Stock Price." *Review of Financial Studies* 20 (2007): 619–650.

CLEMENT, MICHAEL B. "Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter?" *Journal of Accounting and Economics* 27 (1999): 285–303.

CLEMENT, MICHAEL B., LISA KOONCE, and THOMAS J. LOPEZ. "The Roles of Task-Specific Forecasting Experience and Innate Ability in Understanding Analyst Forecasting Performance." *Journal of Accounting and Economics* 44 (2007): 378–398.

CLEMENT, MICHAEL B, and SENYO Y TSE. "Financial Analyst Characteristics and Herding Behavior in Forecasting." *Journal of Finance* 60 (2005): 307–341.

COHEN, LAUREN, and ANDREA FRAZZINI. "Economic Links and Predictable Returns." *Journal of Finance* 63 (2008): 1977–2011.

COHEN, LAUREN, ANDREA FRAZZINI, and CHRISTOPHER MALLOY. "Sell-Side School Ties." *Journal of Finance* 65 (2010): 1409–1437.

DANIEL, KENT, MARK GRINBLATT, SHERIDAN TITMAN, and RUSS WERMERS. "Measuring Mutual Fund Performance with Characteristic-based Benchmarks." *Journal of Finance* 52 (1997): 1035–1058.

DEVLIN, JACOB, MING-WEI CHANG, KENTON LEE, and KRISTINA TOUTANOVA. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. 2019.

DIETHER, KARL B., CHRISTOPHER J. MALLOY, and ANNA SCHERBINA. "Differences of Opinion and the Cross Section of Stock Returns." *Journal of Finance* 57 (2002): 2113–2141.

DING, WENZHI, ROSS LEVINE, CHEN LIN, and WENSI XIE. "Corporate Immunity to the COVID-19 Pandemic." *Journal of Financial Economics* 141 (2021): 802–830.

ESTES, THOMAS H, and JOSEPH L VAUGHAN. "Reading Interest and Comprehension: Implications." *Reading Teacher* 27 (1973): 149–153.

FAMA, EUGENE F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (1970): 383–417.

FAMA, EUGENE F., LAWRENCE FISHER, MICHAEL C. JENSEN, and RICHARD ROLL. "The Adjustment of Stock Prices to New Information." *International Economic Review* 10 (1969): 1–21.

FANG, BINGXU, and OLE-KRISTIAN HOPE. "Analyst Teams." *Review of Accounting Studies* 26 (2021): 425–467.

GERKEN, WILLIAM C., and MARCUS O. PAINTER. "The Value of Differing Points of View: Evidence from Financial Analysts' Geographic Diversity." *Review of Financial Studies* 36 (2023): 409–449.

GLEASON, CRISTI A, and CHARLES M C LEE. "Analyst Forecast Revisions and Market Price Discovery." *Accounting Review* 78 (2003): 193–225.

GOLDSTEIN, ITAY, and LIYAN YANG. "Information Diversity and Complementarities in Trading and Information Acquisition." *Journal of Finance* 70 (2015): 1723–1765.

GOLDSTEIN, ITAY, and LIYAN YANG. "Good Disclosure, Bad Disclosure." *Journal of Financial Economics* 131 (2019): 118–138.

GREEN, T CLIFTON, RUSSELL JAME, STANIMIR MARKOV, and MUSA SUBASI. "Access to Management and the Informativeness of Analyst Research." *Journal of Financial Economics* 114 (2014): 239–255.

GRINBLATT, MARK, and SHERIDAN TITMAN. "The Persistence of Mutual Fund Performance." *Journal of Finance* 47 (1992): 1977–1984.

HOBERG, GERARD, GORDON PHILLIPS, and NAGPURNANAND PRABHALA. "Product Market Threats, Payouts, and Financial Flexibility." *Journal of Finance* 69 (2014): 293–324.

HONG, HARRISON, JEFFREY D. KUBIK, and AMIT SOLOMON. "Security Analysts' Career Concerns and Herding of Earnings Forecasts." *RAND Journal of Economics* 31 (2000): 121–144.

HUANG, ALLEN H, REUVEN LEHAVY, AMY Y ZANG, and RONG ZHENG. "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach." *Management Science* 64 (2018): 2833–2855.

HUANG, YI, CHEN LIN, SIBO LIU, and HEIWAI TANG. "Trade Networks and Firm Value: Evidence from the US-China Trade War." *Journal of International Economics* 145 (2023): 103811.

INTINTOLI, VINCENT J, MATTHEW SERFLING, and SARAH SHAIKH. "CEO Turnovers and Disruptions in Customer-Supplier Relationships." *Journal of Financial and Quantitative Analysis* 52 (2017): 2565–2610.

IVKOVIĆ, ZORAN, and NARASIMHAN JEGADEESH. "The Timing and Value of Forecast and Recommendation Revisions." *Journal of Financial Economics* 73 (2004): 433–463.

IWASAKI, HITOSHI, YING CHEN, QIANQIAN DU, and JUN TU. "Topic Tones of Analyst Reports and Stock Returns: A Deep Learning Approach." *Working Paper*. (2021)

JACOB, JOHN, THOMAS Z LYS, and MARGARET A NEALE. "Expertise in Forecasting Performance of Security Analysts." *Journal of Accounting and Economics* 28 (1999): 51–82.

JUNG, MICHAEL J., M. H. FRANCO WONG, and X. FRANK ZHANG. "Buy-Side Analysts and Earnings Conference Calls." *Journal of Accounting Research* 56 (2018): 913–952.

KACPERCZYK, MARCIN, SAVITAR SUNDARESAN, and TIANYU WANG. "Do Foreign Institutional Investors Improve Price Efficiency?" *Review of Financial Studies* 34 (2021): 1317–1367.

KADAN, OHAD, LEONARDO MADUREIRA, RONG WANG, and TZACHI ZACH. "Analysts' Industry Expertise." *Journal of Accounting and Economics* 54 (2012): 95–120.

LIVNAT, JOSHUA, and YUAN ZHANG. "Information Interpretation or Information Discovery: Which Role of Analysts Do Investors Value More?" *Review of Accounting Studies* 17 (2012): 612–641.

LOH, ROGER K., and RENÉ M. STULZ. "When Are Analyst Recommendation Changes Influential?" *Review of Financial Studies* 24 (2011): 593–627.

MARTINEAU, CHARLES, and MARIUS ZOICAN. "Measuring Information in Analyst Reports: A Machine Learning Approach." *Working Paper*. (2021)

MAYEW, WILLIAM J., NATHAN Y. SHARP, and MOHAN VENKATACHALAM. "Using Earnings Conference Calls to Identify Analysts with Superior Private Information." *Review of Accounting Studies* 18 (2013): 386–413.

MENZLY, LIOR, and OGUZHAN OZBAS. "Market Segmentation and Cross-Predictability of Returns." *Journal of Finance* 65 (2010): 1555–1580.

MIKOLOV, TOMAS, ILYA SUTSKEVER, KAI CHEN, GREG S. CORRADO, and JEFF DEAN. "Distributed Representations of Words and Phrases and Their Compositionality." *Advances in Neural Information Processing Systems* 26 (2013).

PATATOUKAS, PANOS N. "Customer-Base Concentration: Implications for Firm Performance and Capital Markets: 2012 American Accounting Association Competitive Manuscript Award Winner." *Accounting Review* 87 (2012): 363–392.

PRENZEL, MANFRED. "Task Persistence and Interest." In: *U. Schiefele (Ed.), Content and interest as motivated factors in learning.* (1988). Symposium conducted at the annual meeting of the American Educational Research Association, New Orleans.

RENNINGER, ANN. "Do Individual Interests Make a Difference?" In: *Essays by the Spencer. Fellows IV*. National Academy of Education. Cambridge, MA, (1987): 228-253.

RENNINGER, K. ANN and ROBERT H WOZNIAK. "Effect of Interest on Attentional Shift, Recognition, and Recall in Young Children." *Developmental Psychology* 21 (1985): 624–632.

RUBIN, AMIR, BENJAMIN SEGAL, and DAN SEGAL. “The Interpretation of Unanticipated News Arrival and Analysts’ Skill.” *Journal of Financial and Quantitative Analysis* 52 (2017): 1491–1518.

SCHIEFELE, ULRICH, and ANDREAS KRAPP. “The Impact of Interest on Qualitative and Structural Indicators of Knowledge.” In *U. Schiefele (Chair), Content and interest as motivational factors in learning.* (1988) Symposium conducted at the annual meeting of the American Educational Research Association, New Orleans.

SCHIEFELE, ULRICH, ANDREAS KRAPP, and ADOLF WINTELER. “Interest as a Predictor of Academic Achievement: A Meta-Analysis of Research.” In *K. A. Renninger, S. Hidi, & A. Krapp (Eds.), The role of interest in learning and development* (1992): 183–212. Lawrence Erlbaum Associates, Inc.

SCOTTI, CHIARA. “Surprise and Uncertainty Indexes: Real-Time Aggregation of Real-Activity Macro-Surprises.” *Journal of Monetary Economics* 82 (2016): 1–19.

STICKEL, SCOTT E. “Reputation and Performance among Security Analysts.” *Journal of Finance* 47 (1992): 1811–1836.

STICKEL, SCOTT E. “The Anatomy of the Performance of Buy and Sell Recommendations.” *Financial Analysts Journal* 51 (1995): 25–39.

WOMACK, KENT L. “Do Brokerage Analysts’ Recommendations Have Investment Value?” *Journal of Finance* 51 (1996): 137–167.

Appendix I. Variable Definitions

Dependent Variables	
FE	The absolute difference between the analyst's earning forecast and actual reported earnings, divided by the stock price of the firm at the end of the month prior to actual earnings announcement.
Abs(FE)	|FE| (winsorized at 1%) × 100.
DGTW AR [0]	DGTW-adjusted abnormal return on day 0 of the recommendation upgrades and downgrades, multiplied by 100.
DGTW CAR [0,1]	DGTW-adjusted cumulative abnormal returns for two days after recommendation updates, multiplied by 100.
Internal Promotion	Dummy variable that is equal to one if an analyst's job title becomes more senior at the same brokerage firm over the next two years, and zero otherwise.
Join Buy-side	Dummy variable that is equal to one if the analyst joins a buy-side firm, and zero otherwise.
Join Larger Brokerage Firm	Dummy variable that is equal to one if the analyst joins a sell-side brokerage firm that is larger (in terms of the number of analysts) than their current employer, and zero otherwise.
$Earnings_{t+1}$	Corporate earnings in year *t+1*, scaled by total assets in year *t*. It is multiplied by 100 to express values in percentage terms.
BuySideQuestion%	The percentage of questions raised by buy-side analysts, calculated as the number of questions asked by buy-side analysts divided by the total number of questions asked during a firm's earnings conference calls in year *t*, multiplied by 100.
BuySideParticipant%	The percentage of participants who are buy-side analysts, calculated as the number of unique buy-side analysts who raised questions divided by the total number of unique participants who asked questions during the firm's earnings conference calls in year *t*, multiplied by 100.
RND_{t+1}	R&D expense in year *t+1* scaled by the beginning-of-year total assets.
$CHGASSET_{t+1}$	The change in assets in year *t+1* scaled by the beginning-of-year total assets.
$CAPXRND_{t+1}$	Capital expenditure plus R&D in year *t+1* scaled by the beginning-of-year total assets.

Independent Variables

# of Supply Chain Q's	An analyst's number of questions that mention supply-chain keywords in the past four years, excluding the questions from the focal firm's own conference calls.
Expertise	A dummy variable that equals one if the analyst's *# of Supply Chain Q's* of the forecast ranks within the top 10th percentile among all analysts' forecasts of the year.
SUP	A dummy variable set equal to one if the firm is a dependent supplier for any public firm in year *t* (discloses at least one public principal customer), and zero otherwise.
CUS	A dummy variable set equal to one if the firm is a principal customer in year *t* (disclosed as a principal customer by at least one public firm), and zero otherwise.
Total # of Questions	Total number of questions raised by an analyst in the past four years, divided by 1000.
# of Firms Covered	The number of firms an analyst follows in year *t*.
# of SIC3 Inds Covered	The number of industries an analyst covers in a year, where industries are defined by 3-digit SIC codes.
LN(Experience)	The natural logarithm of 1 plus the number of days between the date of the forecast and the date of the first forecast of the analyst (winsorized at 1% level).
LN(Horizon)	The natural logarithm of 1 plus the number of days between the date of the forecast and the date of the actual earnings released (winsorized at 1% level).
LN(Firm Size)	The natural logarithm of firm size, where firm size is defined as stock price times the total number of shares outstanding at the end of last year.
LN(BM)	The natural logarithm of the firm's book value of equity divided by its market value at the end of the last fiscal year. The variable is set to missing if the book value of equity is negative (winsorized at 1% level).
Brokerage Firm Size	Size of brokerage firm as proxied by the total number of analysts hired by the brokerage firm in year *t*.
Profitability	Operating income of the firm for fiscal year *t-1*, divided by book value of equity at the end of fiscal year t-2. The variable is set to missing if the book value of equity is negative (winsorized at 1% level).
Change in Recommendation	The difference between this recommendation and the prior recommendation.
Two Weeks Before Earnings	Dummy variable that equals one if the recommendation change is issued within the two-week window prior to the firm's earnings announcement, and zero otherwise.
Two Weeks After Earnings	Dummy variable that equals one if the recommendation change is issued within the two-week window after the firm's earnings announcement, and zero otherwise.

Away from Consensus	Dummy variable that equals one if the absolute difference between this recommendation and the consensus is larger than that of the prior recommendation.
Past Accuracy	Average forecast errors in the past year multiplied by -1.
ChinaImport	Dummy variable that is equal to one if a firm imports from China in year *t*, and zero otherwise.
ChinaExport	Dummy variable that is equal to one if a firm exports to China in year *t*, and zero otherwise.
ChinaTrade	Dummy variable equal to one if a firm either imports from or exports to China in year *t*, and zero otherwise.
# of Analysts Covering	The number of analysts that cover the firm in year *t*.
Has Expertise	Dummy variable that equals one if the analyst has expertise in at least one of the most frequently discussed topics (in terms of their frequency in earnings conference calls).
Coincide	Dummy variable that equals one if at least one of the analyst's expertise topics coincides with a peak information arrival for the firm in year *t*, and zero otherwise.
Peak	For each firm–topic–year, the peak period is defined as the year when the share of questions related to a given topic in the firm's earnings calls exceeds the sample-wide mean share of that topic by more than two standard deviations. The variable *Peak* is defined at the firm-year level and equals one if any of the top 20 topics is in a peak information arrival period for that firm in year *t*.
Peak of Covered Firms	Dummy variable that is defined at the analyst–year level and equals one if any firm covered by the analyst experiences a peak information arrival period in any of the top 20 topics during year *t*.
LogMVA	The natural logarithm of the ratio between the firm's market capitalization and total assets in year *t*.
# of Expertise	The number of distinct topic-specific expertise held by analysts covering a firm in year *t*, based on the top 20 most frequently discussed topics.
# of Rep Expertise	The number of distinct top 20 topic-specific expertise possessed by the analysts covering this firm in year *t* when we restrict each analyst to a single expertise based on her most frequently discussed topic.
Q	The firm's Tobin's Q in year *t* which is calculated as market value of equity plus book value of assets minus book value of equity, scaled by book value of assets.
Leverage	Long-term debt plus debt in current liabilities, scaled by total assets in year *t*.

Tangibility	Gross property, plant, and equipment, scaled by total assets in year *t*.
Cash	Cash, scaled by total assets in year *t*.
Sale	Sale, scaled by total assets in year *t*.
R&D	R&D expenses, scaled by total assets in year *t*.
Current Earnings	Corporate earnings in year *t*, scaled by total assets in year *t*.
LN(Total Assets)	The natural logarithm of total assets in year *t*.
ROA	Returns on assets in year *t*.
ROA Std	Standard deviation of ROA in the past five years.
Ret	Average returns in year *t*.
Ret Std	Standard deviation of daily stock returns in year *t*.
Ret_{t+3}	Future 3 years returns of the firm.
SG	Sales growth rate of the firm in year *t*.
1/Assets	One divided by total assets of the firm in year *t*.
CF	The sum of net income before extraordinary items, depreciation and amortization expenses and R&D expenses in year *t+1*, scaled by beginning-of-year book assets.

Appendix II: Constructing the Supply Chain Keywords List using Machine Learning Approach (Word2Vec)

This appendix describes the methodology used to develop an expanded list of supply-chain-related keywords using a machine learning approach. The objective is to go beyond a small set of obvious terms and identify additional vocabulary that is closely related in meaning, as it appears in earnings conference call transcripts. We use Word2Vec to generate word embeddings (numerical representations of words) that allow us to measure semantic similarity between terms and systematically expand our supply chain lexicon.

Word2Vec is a neural-network-based model that represents each word as a vector in a high-dimensional space, where the geometric proximity between vectors reflects semantic similarity. Words that appear in similar contexts in the training corpus will be located closer together in this vector space, while unrelated words will be farther apart. Each dimension of the vector captures latent linguistic and semantic attributes. For example, the vector representation of "supplier" will be close to words such as "vendor" or "manufacturer" if these terms frequently occur in similar contexts in the corpus.

We train the Word2Vec model on the complete corpus of earnings conference call transcripts, ensuring that the learned embeddings reflect the way language is actually used in this specific domain. Once trained, the model assigns each word in the vocabulary a fixed vector in a 400-dimensional space, allowing direct computation of similarity between any two words.

To identify supply-chain-related terms, we begin with nine seed words that are unambiguously associated with the supply chain topic: "supply chain," "supplier," "outsource," "inventory," "export," "import," "shipping," "procure," and "logistics." For each seed word, we calculate the cosine similarity between its vector and the vector of every other word in earnings conference call transcripts. Cosine similarity measures the cosine of the angle between two vectors, with higher values indicating greater similarity in direction and, therefore, meaning. Formally, for a word *i* and seed word *j*, cosine similarity is defined as:

$$Cosine\ Similarity_{i,j} = \frac{\vec{i} \cdot \vec{j}}{\parallel \vec{i} \parallel \parallel \vec{j} \parallel} = \frac{\sum_{k=1}^{n} i_k j_k}{\sqrt{\sum_{k=1}^{n} i_k^2}\sqrt{\sum_{k=1}^{n} j_k^2}}$$

where $\vec{i}$ is the vector of the word i, and $\vec{j}$ is the vector of the seed word. n represents the total number of dimensions of the vector, which in our case is 400. k represents the kth dimension of the vector. A high cosine similarity between a word with the seed words indicates that the word is closely related to the supply chain topic.

For each seed word, we rank all other words in earnings conference call transcripts by cosine similarity and select the top 100 as candidate supply-chain-related keywords. This produces, for each seed, a set of terms that the model considers most semantically related to the underlying supply chain concept. Because some words are relevant to multiple seeds, there is overlap across these candidate lists. In such cases, a word is retained only for the seed word with which it exhibits the highest cosine similarity.

We then manually review all candidate words to refine the list and ensure topic relevance. For each term, we examine its usage in the transcript context to determine whether it specifically relates to supply chain matters. This process removes words that are overly general or ambiguous in meaning. For example, terms like “buy” or “utilize” may appear near seed words in the vector space but lack direct supply chain specificity. By applying this manual filtering, we retain only those terms that carry clear thematic relevance. Examples of the top 10 auto-generated keywords for each seed word are reported in Appendix Table II.1.

Table II.1 The Top Keywords Most Similar to Each Seed Word

Expanded Word (Supply Chain)	Expanded Word (Supplier)	Expanded Word (Outsource)	Expanded Word (Inventory)	Expanded Word (Export)
supply chains	suppliers	in-source	inventories	seaborne
sourcing	vendor	in-house	on-hand inventory	exporters
centralized procurement	provider	internalize	finished goods	import/export
supply chain-related	producer	centralize	work-in-process inventory	exportation
indirect procurement	manufacturers	offshoring	in-transit inventory	re-exported
kitting	distributor	subcontract	buffer stocks	LPG
inbound freight	OEM	regionalize	stocking	export ban
vendor-managed inventory	integrator	co-source	WIP	transshipment
JIT	subcontractor	BPO	inventory buildup	seaborne coal
cross-docking	sole supplier	re-engineer	consigned inventory	export-driven

Expanded Word (Import)	Expanded Word (Shipping)	Expanded Word (Procure)	Expanded Word (Logistics)
importers	ship	procurement	logistical
importation	shipment	sourced	warehousing
Chinese imports	supplying	replenish	freight
import duties	air freighting	long-lead	transportation
antidumping duties	loading	decouple	fulfillment
locally manufactured	ordering	offload	pick-and-pack
countervailing duties	transporting	self-supply	transport
customs duty	container	multisource	sortation
anti-dumping investigation	rerouted	unbundle	sea freight
unfairly traded	ocean freight	consign	customs clearance

Appendix III. Top 20 Supply-Chain Expert Analysts in 2019

Analyst	Brokerage Firm
Craig Irwin	Roth Capital Partners
Michael Latimore	Northland Securities
Rod Lache	Wolfe Research
John Baugh	Stifel Financial Corp.
Keith Weiss	Morgan Stanley
Eric Stine	Craig-Hallum
Vivek Arya	BofA Securities
Deane Dray	RBC Capital Markets
Fotis Giannakoulis	Morgan Stanley
James Richutti	Needham & Company
Benjamin Swinburne	Morgan Stanley
Steve Tusa	J.P. Morgan
Brett Feldman	Goldman Sachs
Mike Carrier	BofA Securities
Ryan Brinkman	J.P. Morgan
John Murphy	BofA Securities
Matt Sheerin	Stifel Financial Corp.
Sterling Auty	J.P. Morgan
Brian Johnson	Barclays Capital
Harlan Sur	J.P. Morgan

Appendix IV: NLP and Clustering Methodology for Identifying Non-Prespecified Topics

This technical appendix provides a detailed description of the topic modeling methodology used to analyze the non-prespecified topics of analyst questions from earnings conference calls.

1. Data Cleaning

Before generating embeddings, we performed a data cleaning process to ensure that the analyst questions entering the pipeline contained enough substantive information for meaningful analysis. We began by removing common stopwords, which are high-frequency words that generally carry little semantic weight such as "I," "you," and "thanks." In earnings call transcripts, these terms often serve as conversational fillers or polite prefaces and do not contribute to identifying the underlying topic of a question. Removing them reduces textual noise and focuses the model's attention on the words most likely to indicate content-specific meaning. We also excluded questions containing fewer than 50 characters (roughly fewer than 10 words). This length filter eliminates overly brief, vague, or incomplete questions that lack the contextual richness needed for accurate semantic representation and topic clustering. These preprocessing steps ensured that the BERT embeddings would be computed from linguistically meaningful inputs, improving the reliability of the subsequent topic detection.

2. BERT for Contextual Text Embeddings

BERT (Bidirectional Encoder Representations from Transformers) is a deep neural network architecture designed to produce contextualized vector representations of text. At its core, BERT is built upon the Transformer framework, which replaces sequential processing with a self-attention mechanism that allows the model to consider every word in a sequence in relation to every other word, regardless of position. In practice, this means that the representation of a given word is computed as a weighted combination of all the other words in the sentence, where the weights reflect their contextual relevance.

What distinguishes BERT from earlier representation models, such as Word2Vec or GloVe, is its bidirectional training objective. While traditional models either assign each word a fixed, context-independent vector or process text in a single direction (left-to-right or right-to-left), BERT processes text in both directions simultaneously. It does this through masked language modeling (MLM), where a random subset of words in the input is replaced by a special masked token, and the model is trained to predict the masked words using information from both left and right contexts. This bidirectionality enables BERT to capture subtle dependencies that are essential for accurately interpreting domain-specific language, such as financial analyst questions that may contain technical terms, abbreviations, or references to prior discussion.

BERT is also pre-trained on extremely large general-purpose corpora (the Toronto BookCorpus and English Wikipedia) totaling over three billion words. During pre-training, it jointly learns two tasks: MLM and next sentence prediction (NSP), the latter requiring the model to determine whether one sentence logically follows another in the original text. Together, these tasks allow BERT to learn rich syntactic and semantic patterns in language, including relationships across sentence boundaries.

When applied to analyst questions, the model first tokenizes the text into subword units using WordPiece tokenization, ensuring that even rare financial terms are broken down into known components. These tokens are then mapped to high-dimensional embeddings, passed through multiple Transformer layers, and refined via self-attention to produce the final representation. In our implementation, each question is represented as a single dense vector in a high-dimensional space (384 dimensions with MiniLM variant). Questions with similar semantics are located close together in this vector space, while semantically distinct questions are farther apart. This embedding space forms the mathematical foundation for our subsequent dimensionality reduction and clustering steps.

While BERT embeddings preserve rich semantic information, their high dimensionality poses both computational and statistical challenges for clustering. High-dimensional spaces suffer from the curse of dimensionality, where distances between points become less meaningful and algorithms become less efficient. To address this, we apply Uniform Manifold Approximation and Projection (UMAP), a non-linear dimensionality reduction technique designed to preserve both local and global structure in the data.

UMAP models the data as a weighted k-nearest neighbor graph in the original high-dimensional space, where edges represent proximity between points. It then optimizes a low-dimensional embedding so that this graph structure is preserved as closely as possible. Local relationships (pairs of points that are very close in the original space) are prioritized to maintain semantic fidelity, while global relationships ensure that distinct clusters remain well separated.

Compared to alternatives such as Principal Component Analysis (PCA) or t-distributed Stochastic Neighbor Embedding (t-SNE), UMAP offers key advantages: it captures non-linear structures that PCA cannot, preserves global topology better than t-SNE, and scales efficiently to large datasets. Most importantly for our application, UMAP reduces the embeddings to a dimension feasible for clustering while retaining the semantic neighborhoods established by BERT. After UMAP, questions with similar meaning remain near each other, enabling accurate topic discovery without the computational burden of clustering directly in the original embedding space.

3. Clustering with HDBSCAN

After embedding, the analyst questions are represented as vectors in a multi-dimensional semantic space, where distances reflect similarity in meaning. To identify coherent topics, we use Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), a density-based clustering algorithm that is well suited to handling text embeddings, which often produce clusters of irregular shapes and varying densities.

HDBSCAN begins by constructing a mutual reachability graph from the vectors, where the distance between two vectors is adjusted to account for local density variations. This graph is then transformed into a minimum spanning tree (MST), which represents the closest connections among all vectors. The algorithm progressively increases the distance threshold for connecting vectors, effectively “relaxing” the clustering criteria step-by-step. As the threshold widens, small, tight groups of questions merge into larger, looser groups. Throughout this process, HDBSCAN records a hierarchy of cluster formations and dissolutions across different density levels.

The core of the method lies in evaluating the stability of each potential cluster across this hierarchy. Stability is computed as the persistence of a cluster over a range of distance thresholds: a cluster that remains intact across multiple levels is deemed more stable than one that quickly fragments or merges. After building the hierarchy, HDBSCAN selects the most stable clusters as the final set of topics.

This approach has important advantages for our data. Analyst questions can vary dramatically in linguistic form: some topics are expressed with highly consistent phrasing, producing dense and compact clusters, while others are discussed in more diverse language, resulting in broader and sparser clusters. Unlike methods such as *k*-means, which assume uniform, spherical clusters, or the original DBSCAN, which applies a single fixed distance threshold, HDBSCAN's hierarchical density-based process can capture both tight, jargon-heavy topics and diffuse, heterogeneous topics without imposing restrictive cluster shapes. In our application, this means the algorithm can faithfully represent both narrowly defined financial issues (e.g., a specific accounting change) and broad strategic concerns (e.g., long-term growth guidance) within the same analytical framework.

4. Labeling the Analyst Questions

Using the topic model generated in the preceding steps, we assign each analyst question to the topic with which it has the strongest semantic alignment. The assignment is determined by identifying the topic whose representation has the highest similarity to the question's embedding, and we require this similarity to exceed 80%. This threshold ensures that each question is associated only with a topic when there is a clear and meaningful connection.

5. Enhancing Topic Interpretability through TF-IDF

After generating semantic embeddings of analyst questions using BERT and identifying topic clusters via HDBSCAN, we obtain a set of machine-generated topic representations in the form of dense numerical vectors. While these embeddings capture the nuanced semantic structure of each topic, they are not inherently interpretable to human readers. To bridge this gap and present the topics in a reader-friendly manner, we employ the Term Frequency-Inverse Document Frequency (TF-IDF) method, a widely used statistical measure in information retrieval and text mining.

TF-IDF quantifies the relative importance of a word within the set of questions belonging to a given topic while adjusting for its prevalence across all topics. It consists of two components. Term frequency (TF) measures how often a word appears in the documents associated with a topic, capturing its local prominence. Inverse document frequency (IDF) down-weights words that are common across many topics by measuring how rare a term is across the entire corpus. Multiplying TF by IDF yields a score that highlights words that occur frequently within a given topic but infrequently in others, thus isolating terms that are distinctive to that topic.

In our application, we treat the set of analyst questions assigned to each HDBSCAN cluster as a topic-specific corpus. We calculate the TF-IDF score for every word in that corpus, using the set of questions from all topics to compute the IDF term. This ensures that high scores are assigned to words that are both frequent within the topic and relatively uncommon in other topics. By ranking words according to their TF-

IDF scores, we extract the top keywords that best characterize the thematic essence of each topic. These keywords provide a concise, interpretable summary of the content captured by the topic's embedding. The top 20 topics for our sample period are presented in Appendix Table IV.1.

Table IV.1 The Top 20 Non-Prespecified Topics Identified from Analyst Questions

Topic	Word Representation
1	Payout Policy and Capital Structure
2	Geographical Markets
3	Growth Drivers
4	Profitability
5	Customer Orders
6	Telecommunications Segment
7	Oil & Gas Segment
8	Market Trends
9	Input Prices
10	Pharmaceutical Segment
11	Order Backlog
12	Compliance and Litigations
13	Cost Synergy Strategies
14	Markup
15	SG&A Expenses
16	Other Operating Expense
17	Guidance of Future Performance
18	Labor Issues
19	Cash
20	Macroeconomic Environment

Table IV.2 Examples of Analyst Questions of the Top 20 Topics

Topics	Example
Payout Policy and Capital Structure	Okay. the next question, Hermann, if I have my numbers correct using your $0.46 in dividend for the quarter, if I annualize that and I look at the share price, obviously the share price will be up this morning. I'm coming out with less than 4% dividend yield, which puts you at the lower end, particularly compared to the U.S. players out there. When can we expect the different pay outs to go up, particularly if you look at the very strong balance sheet even after the stock buybacks?
Geographical Markets	Can you give us some sense of the revenue shortfall or how much of that was driven by the European market versus the U.S. market?
Growth Drivers	This is Emily Wang from Raymond James. When you're talking about your expectations for year-over-year volume growth, could you more specifically quantify how much you think will be from organic growth and how much will be from inorganic growth?
Profitability	All right. and then I think in your closing remarks, David, you said that you have a good confidence that you will be reaching your long-term targets. and especially, I mean, I would like to ask on the profitability target based over a 30% EBIT margin by 2024, what makes you so confident that you are going to reach that given that I think we've been a little bit short on the profitability compared to the market estimates?
Customer Orders	I'm basically asking for the linearity of the order patterns. If you look at orders incoming minus cancellations and look at net bookings on a monthly basis, did net bookings turn positive either in May-June or so far in July?
Telecommunications Segment	Yes, Needham & Company. You guys are holding up very, very well in the current market compared to a lot of other chip companies especially with, you know, fairly significant exposure on the telecom side. Can you elaborate a little bit about maybe why your business exposure is a little bit different, and how you are able to dodge all the bullets flying around these days?
Oil & Gas Segment	Okay. And then in your oil sands project, the next one that is sort of in the hopper is that Silvertip? And when can you -- what will be the timing of getting that project ready to a point where you can file the regulatory papers?
Market Trends	Okay. And maybe just last question, any thoughts on market trends? Any color, just the usual, on mortgage or to cars? Any changes there?
Input Prices	You had talked a little bit about raising prices based on the product of the input. Can you talk about what's going on with needle coke and some of your other inputs and whether or not you see that sort of --I don't know stabilizing at this point in terms of your cost?
Pharmaceutical Segment	Good afternoon. Thanks for taking my question. my first is on the clinical development of lorcaserin. Do you have any updated thoughts on combinations with phentermine, is that something you plan to do at all, and do you expect the FDA to want to see some preliminary safety data with phentermine before approving lorcaserin?
Order Backlog	Hi guys. I was just wondering if you could give some information on your orders and your backlog for the quarter.
Compliance and Litigations	And just one more on the lawsuit with northrop. What would you say could be the probability that that situation is settled without going to court?
Cost Synergy Strategies	Is there potential for synergy with DRA to reduce those operating costs or does it kind of mete out with your growth?

Markup	And would the margins on this product for you folks, given your high gross margins at the 80 percent level - would they be comparable to the company's gross margins or a little less?
SG&A Expenses	Right. And then also on the SG&A line, Kenny, is it safe for us to assume that if based on the forward order positions that you just outlined for us, that if business remains kind of weak or flat in some certain market areas, whether it's international or part of your U.S. apparel business, would you reduce some of your brand building expenses in order to keep your SG&A in line from deleveraging to the degree it did in the first quarter?
Other Operating Expense	Good morning, Tom and John. Thanks for the detail on the 8-K as well. I wanted to follow up on one of the cost items that's the other item. It was up 21%. seems like a big jump. and I'm wondering, number one, why that was up as much, and what we might look for in that item going forward, whether there's some costs that are going to go away.
Guidance of Future Performance	Excellent. How long would it take to ramp up to the 875 expected run rate for the Pershing business? And that 875 does not include revenue synergies, correct?
Labor Issues	Okay. Anything from the standpoint of headcount changes where it would be kind of your -- are you guys doing any hiring? Are you kind of maintaining? Give me a sense on your labor pool.
Cash	Okay. If we look at the cash, how much cash do you guys end up burning in the quarter and what are you planning on sort of exiting the year at in cash balance?
Macroeconomic Environment	So my first question is one kind of on the macro backdrop. When you think about the Taper, some fiscal policy uncertainty and what's going on with inflation and how that all relates to interest rates and volatility. I was just wondering kind of high-level what scenarios keep you up at night? and then what scenario is kind of the best case for agency MBS investors?

Appendix V: Appendix Tables

Table A1 Supply-Chain-Specific Expertise and Forecast Accuracy: Cross-Sectional Partitions Based on Information Asymmetry and Customer-Switching Risk

This table presents the effect of analyst supply-chain-specific expertise on the absolute forecast errors in subgroups of firms partitioned by the measures of information asymmetry and customer-switching risk. Information asymmetry is measured by analyst forecast dispersion, defined as the standard deviation of analyst forecasts scaled by price for a given firm in the fiscal year *t*. Based on analyst forecast dispersion, forecasts made for firms that rank within the top 40% are classified into the *High Information Asymmetry* group and those ranked within the bottom 40% are classified into the *Low Information Asymmetry* group. Customer-switching risk is measured by product market fluidity in Hoberg, Phillips and Prabhala (2014). Forecasts made for firms that have product market fluidity ranking top 40% in the cross section are classified into the *High Customer Switching Risk* group while those belonging to the bottom 40% are classified into the *Low Customer Switching Risk* group. The dependent variable in these regressions is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings divided by the stock price of the firm at the end of the month before the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to according to the analyst's supply-chain-specific questions in earnings conference calls. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (the firm discloses at least one public principal customer), and zero otherwise. All other control variables, as specified in Table 2, are included in test specifications. Definitions of control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1)	(2)	(3)	(4)
	Analyst Forecast Dispersion		Product Market Fluidity	
	High Information Asymmetry	*Low Information Asymmetry*	*High Customer Switching Risk*	*Low Customer Switching Risk*
$SUP \times Expertise$	-0.164**	0.006	-0.179**	-0.037
	(0.082)	(0.005)	(0.076)	(0.033)
Control Variables	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	234,517	240,551	209,936	213,690
Adj R-squared	0.762	0.846	0.787	0.781
Difference	-0.170**		-0.142**	
Test:	High Dispersion = Low Dispersion		High Competition = Low Competition	
	p-val = 0.041		p-val = 0.042	

Table A2 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy Using Analyst Reports

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of analyst supply-chain-specific expertise. The dependent variable in all test specifications is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-related reports. Analyst supply-chain-specific expertise in column (1) is constructed using 2 years of analyst reports data before year *t*. Similarly, this variable is constructed using 3 years and 4 years of analyst reports data in columns (2) and (3), respectively. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. *Total # of Reports* is the total number of reports by the analyst in the past four years. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.069**	-0.072**	-0.098***
	(0.029)	(0.030)	(0.032)
$CUS \times Expertise$	0.053**	0.038	0.026
	(0.022)	(0.024)	(0.025)
$Expertise$	-0.018	0.013	-0.020
	(0.025)	(0.027)	(0.028)
$Total\ \#\ of\ Reports$	-0.079	-0.105**	-0.064*
	(0.061)	(0.045)	(0.037)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006*
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.010**	0.009*	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.399***	0.403***	0.407***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,415	632,827	598,223
Adj R-squared	0.788	0.789	0.788

Table A3 Analyst Supply-Chain-Specific Expertise and the Market Reaction to Analyst Recommendation Updates Using Analyst Reports

This table reports the market reactions to analysts' recommendation updates conditional on corporate supply-chain risk exposure and analyst supply-chain-specific expertise. In columns (1) and (2), the dependent variable is *DGTW AR[0]*, which is the DGTW-adjusted abnormal return on day 0 of the recommendation upgrades and downgrades. In columns (3) and (4), the dependent variable is *DGTW CAR[0,1]*, which is the DGTW-adjusted cumulative abnormal returns for two days after recommendation updates. *DGTW AR[0]* and *DGTW CAR[0,1]* are multiplied by 100 for ease of interpretation. Columns (1) and (3) report results for recommendation upgrades, while columns (2) and (4) report results for downgrades. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-related reports. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *Change in Recommendation* is the difference between this recommendation and the prior recommendation. We include a dummy variable (*Two Weeks Before Earnings*) that is equal to one if the recommendation update is issued within the two-week window before the firm's earnings announcement, and zero otherwise, as well as a dummy variable (*Two Weeks After Earnings*) that is equal to one if the recommendation update is issued within the two-week window after the firm's earnings announcement, and zero otherwise. *Away from Consensus* is a dummy variable that is equal to one if the absolute difference between this recommendation and the consensus is larger than that of the prior recommendation. *Total # of Reports* is the total number of reports by the analyst in the past four years. Definitions of other control variables are provided in Appendix I. Firm, year, and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1) DGTW AR [0]	(2) DGTW AR [0]	(3) DGTW CAR [0,1]	(4) DGTW CAR [0,1]
	Upgrade	Downgrade	Upgrade	Downgrade
$SUP \times Expertise$	0.490***	-0.401**	0.490**	-0.676***
	(0.186)	(0.188)	(0.219)	(0.225)
$Expertise$	0.226	-0.225	0.261	-0.115
	(0.188)	(0.158)	(0.208)	(0.183)
SUP	-0.166	-0.076	-0.225	-0.030
	(0.247)	(0.149)	(0.275)	(0.183)
$Change\ in\ Recommendation$	-0.583***	-0.442***	-0.665***	-0.641***
	(0.113)	(0.080)	(0.131)	(0.095)
$Two\ Weeks\ Before\ Earnings$	0.027	-0.253	-0.067	-0.153
	(0.124)	(0.156)	(0.147)	(0.178)
$Two\ Weeks\ After\ Earnings$	-0.050	-0.076	-0.050	-0.012
	(0.107)	(0.147)	(0.124)	(0.172)
$Away\ from\ Consensus$	0.259***	-0.515***	0.314***	-0.658***
	(0.085)	(0.065)	(0.096)	(0.077)
$Past\ Accuracy$	0.052***	-0.000	0.060*	-0.000
	(0.020)	(0.000)	(0.031)	(0.001)
$\#\ of\ Firms\ Covered$	0.019*	-0.016	0.011	-0.035***
	(0.011)	(0.011)	(0.012)	(0.012)
$\#\ of\ SIC3\ Inds\ Covered$	-0.016	0.040	0.005	0.051*
	(0.022)	(0.024)	(0.027)	(0.030)
$LN(Experience)$	-0.097	0.017	-0.024	0.099
	(0.098)	(0.107)	(0.123)	(0.126)
$Total\ \#\ of\ Reports$	-0.050	0.084	-0.016	0.194
	(0.355)	(0.166)	(0.377)	(0.195)
$Brokerage\ Firm\ Size$	0.010***	-0.010***	0.008***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
$LN(Firm\ Size)$	-0.636***	0.224**	-0.820***	0.250**
	(0.131)	(0.087)	(0.144)	(0.106)
$LN(BM)$	0.295*	0.017	0.333*	0.023
	(0.151)	(0.086)	(0.170)	(0.099)
$Profitability$	0.028	0.002	0.045	0.004
	(0.042)	(0.002)	(0.046)	(0.002)
$\#\ of\ Analysts\ Covering$	-0.006	0.003	-0.019**	0.004
	(0.008)	(0.008)	(0.009)	(0.010)
Analyst FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	32,595	33,230	32,585	33,225
Adj R-squared	0.313	0.206	0.296	0.205

Table A4 The 2018 US-China Trade War, Analyst Supply-Chain-Specific Expertise, and Analyst Forecast Accuracy Using Analyst Reports

This table reports the effects of analyst supply-chain-specific expertise on forecast accuracy around the 2018 US-China trade war. The dependent variable in all columns is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month prior to the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-related reports. *ChinaImport* is a dummy variable that is equal to one if a firm imports from China in year *t*, and zero otherwise. *ChinaExport* is a dummy variable that is equal to one if a firm exports to China in year *t*, and zero otherwise. *ChinaTrade* is a dummy variable equal to one if a firm either imports from or exports to China in year *t*, and zero otherwise. *Post* is a dummy variable that is equal to one if the time is 2018 or onward, and zero otherwise. *Total # of Reports* is the total number of reports by the analyst in the past four years. Due to the data availability of the Panjiva database, the sample in column (1) starts in 2007, whereas columns (2) and (3) start in 2009. The sample period ends in 2019, prior to the outbreak of the COVID-19 pandemic. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1)	(2)	(3)
$ChinaImport \times Post \times Expertise$	-0.147***		
	(0.057)		
$ChinaImport \times Expertise$	0.044		
	(0.027)		
$ChinaExport \times Post \times Expertise$		-0.171*	
		(0.102)	
$ChinaExport \times Expertise$		0.058	
		(0.037)	
$ChinaTrade \times Post \times Expertise$			-0.166***
			(0.057)
$ChinaTrade \times Expertise$			0.053*
			(0.027)
$Post \times Expertise$	0.060	-0.017	0.038
	(0.048)	(0.040)	(0.049)
$Expertise$	-0.070**	-0.002	-0.020
	(0.032)	(0.030)	(0.033)
$Total\ \#\ of\ Reports$	-0.063	-0.050	-0.050
	(0.040)	(0.044)	(0.044)
$\#\ of\ Firms\ Covered$	-0.007***	-0.004**	-0.004**
	(0.002)	(0.002)	(0.002)
$\#\ of\ SIC3\ Inds\ Covered$	0.007*	0.004	0.004
	(0.004)	(0.004)	(0.004)
$LN(Experience)$	0.004	-0.001	-0.001
	(0.006)	(0.006)	(0.006)
$LN(Horizon)$	0.409***	0.362***	0.362***
	(0.012)	(0.012)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	469,980	397,632	397,632
Adj R-squared	0.801	0.776	0.776

Table A5 The 2018 US-China Trade War, Analyst Supply-Chain-Specific Expertise, and Labor Market Outcomes Using Analyst Reports

This table presents the effects of analyst supply-chain-specific expertise on the career outcomes of analysts during the 2018 US-China Trade War. The analysis is conducted at the analyst-year level. Column (1) examines the probability of internal promotion, conditional on an analyst staying at the same brokerage firm over the next two years. The dependent variable *Internal Promotion* is a dummy variable that is equal to one if an analyst's job title becomes more senior at the same brokerage firm over the next two years, and zero otherwise. Column (2) examines an analyst's probability of joining a buy-side firm, conditional on experiencing job turnovers over the next two years. The sample for this column includes analysts who leave their current brokerage firm over the next two years. The dependent variable, *Join Buy-side*, is a dummy variable that is equal to one if the analyst joins a buy-side firm, and zero otherwise. Column (3) examines an analyst's probability of joining a larger sell-side brokerage firm, conditional on the analyst leaving their current brokerage firm and joining another sell-side brokerage firm. The sample for this column includes all analysts who depart from their current brokerage firm and move to another sell-side brokerage firm over the next two years. The dependent variable, *Join Larger Brokerage Firm*, is a dummy variable that is equal to one if the analyst joins a sell-side brokerage firm that is larger (in terms of the number of analysts) than their current employer, and zero otherwise. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-related reports. *Post* is a dummy variable that is equal to one for years 2018 and onward, and zero otherwise. Definitions of control variables are provided in Appendix I. Standard errors, reported in parentheses, are clustered at the brokerage firm level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively

VARIABLES	(1) Analysts with Existing Employers	(2) Analysts with Job Turnovers	(3) Analysts Joining Other Brokerage Firms
	Internal Promotion	Join Buy-Side	Join Larger Brokerage Firm
$Post \times Expertise$	0.020**	0.059*	0.154**
	(0.008)	(0.033)	(0.072)
$Expertise$	-0.003	0.000	0.051
	(0.003)	(0.014)	(0.032)
$Total\ \#\ of\ Reports$	0.003	-0.006	-0.047
	(0.005)	(0.019)	(0.039)
$\#\ of\ Firms\ Covered$	-0.000	-0.003***	0.001
	(0.000)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.000	-0.003**	-0.003
	(0.001)	(0.001)	(0.003)
$LN(Experience)$	-0.000	0.004**	-0.008**
	(0.000)	(0.002)	(0.004)
Brokerage×Year FE	Yes	Yes	Yes
Analyst FE	Yes	No	No
SE Clustered at Brokerage Level	Yes	Yes	Yes
Observations	38,903	6,928	5,029
Adj R-squared	0.333	0.056	0.337

Table A6. **Market Reactions to Earnings Announcements and Relative Optimism of Supply-Chain Experts**

This table reports the market reactions to earnings announcements conditional on whether analysts with supply-chain expertise are more optimistic than the consensus forecast. In column (1), the dependent variable is *DGTW AR[0]*, which is the DGTW-adjusted abnormal return on day 0 of the earnings announcement. In column (2), the dependent variable is *DGTW CAR[0,1]*, which is the DGTW-adjusted cumulative abnormal return for the two days following the earnings announcement. *DGTW AR[0]* and *DGTW CAR[0,1]* are multiplied by 100 for ease of interpretation. *Expertise Optimistic* is a dummy variable that equals one if the average earnings forecast of analysts with supply-chain expertise is more optimistic than the average forecast of other analysts. *SUP* is a dummy variable equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. We include controls for returns volatility over the past year, institutional ownership, and Amihud illiquidity. Definitions of other control variables are provided in Appendix I. Firm and year fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1) DGTW AR [0]	(2) DGTW CAR [0,1]
$SUP \times Expert\ Optimistic$	0.354**	0.613**
	(0.154)	(0.254)
$Expert\ Optimistic$	-0.187	0.396
	(0.160)	(0.288)
SUP	0.047	0.142
	(0.082)	(0.133)
SUE	0.502**	0.938**
	(0.220)	(0.411)
$LN(Firm\ Size)$	-0.187*	-0.270
	(0.098)	(0.165)
$LN(BM)$	-0.081	-0.230
	(0.097)	(0.145)
$Profitability$	-0.005	-0.018
	(0.021)	(0.047)
$\#\ of\ Analysts\ Covering$	-0.026***	-0.068***
	(0.010)	(0.017)
$Returns\ Volatility$	2.937	16.886*
	(5.654)	(9.359)
$Insti\ Ownership$	-0.234	-0.292
	(0.288)	(0.499)
$Illiquidity$	-0.022***	-0.033***
	(0.005)	(0.010)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Cluster at firm	Yes	Yes
Observations	22,392	22,390
Adj R-squared	0.019	0.035

Table A7 Analyst Topic-Specific Expertise, Information Arrivals, and Forecast Accuracy: Other Prespecified Topics

This table presents OLS regression estimates examining how analyst expertise (environment, product, labor or macroeconomic-related expertise) and corresponding information arrivals affect absolute analyst forecast error. Information arrivals are defined as follows: (i) for environmental expertise, the implementation of California cap-and-trade program for California firms after 2013; (ii) for product expertise, the information arrival dummy variable equals one for the years during which the firm has any product-related announcements—including announcements on new, changed, or discontinued products; (iii) for labor expertise, years in which firms are located in states that adopted right-to-work laws after the start of the sample period in 2006; and (iv) for macroeconomic expertise, years with more than 100 days on which the daily macroeconomic surprises index (Scotti, 2016) exceeds the average daily value. The dependent variable in all test specifications is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. Each row reports the coefficient on *Expertise*×*Information Arrival* for the specified topic. Control variables are identical to those in Table 2. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

.

Expertise	Events	Coefficients on Expertise×Information Arrival
Environmental Issues	2013 California cap-and-trade program	-0.109*** (0.019)
Product Issues	Product-related announcements	-0.058** (0.028)
Labor Issues	State-level adoption of right-to-work laws	-0.086** (0.039)
Macroeconomic Issues	Macroeconomic surprises	-0.066*** (0.025)

Table A8 Analyst Expertise Diversity and Consensus Forecast Errors

This table examines the relationship between analyst expertise diversity and consensus forecast errors. The dependent variable *Consensus Forecast Errors* is defined as the mean absolute forecast error across all analysts covering a firm in a given fiscal year. Analyst expertise diversity is proxied by two measures: *# of Expertise* counts the number of distinct topic-specific expertise held by analysts covering a firm in year t, based on the top 20 most frequently discussed topics. When we compute the second measure, *# of Rep Expertise*, we restrict each analyst to a single expertise based on the analyst's most frequently discussed topic. Definitions of other control variables are provided in Appendix I. Firm and year fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

VARIABLES	(1) Consensus Forecast Errors	(2) Consensus Forecast Errors
$\# \, of \, Expertise$	-0.001**	
	(0.001)	
$\# \, of \, Rep \, Expertise$		-0.004**
		(0.002)
$\# \, of \, Analysts \, Covering$	-0.009***	-0.009***
	(0.001)	(0.001)
$LN(Total \, Assets)$	0.070***	0.070***
	(0.009)	(0.009)
$Leverage$	-0.066***	-0.066***
	(0.025)	(0.025)
$Tangibility$	-0.041	-0.041
	(0.026)	(0.026)
$Cash$	0.093***	0.093***
	(0.033)	(0.033)
Q	0.006**	0.006**
	(0.003)	(0.003)
ROA	-0.195***	-0.195***
	(0.027)	(0.027)
$ROA \, Std$	-0.066*	-0.067*
	(0.039)	(0.039)
Ret	-0.687***	-0.686***
	(0.088)	(0.088)
$Ret \, Std$	6.476***	6.476***
	(0.384)	(0.384)
Firm FE	Yes	Yes
Year FE	Yes	Yes
SE Clustered at Firm Level	Yes	Yes
Observations	62,915	62,915
Adj R-squared	0.296	0.296

Table 1 Summary Statistics

This table provides summary statistics of the main variables in this paper. The sample period is from 2002 to 2022. *Abs(FE)* is the absolute forecast error of an analyst's last one-year-ahead forecast before the annual earnings announcement, defined as the absolute difference between earnings forecast and actual reported earnings, divided by the stock price of the firm at the end of the month prior to actual earnings announcement (multiplied by 100). *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for public firms in year *t* (i.e., the firm discloses at least one public principal customer in year *t*), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (i.e., the firm is disclosed as a principal customer by at least one public firm in year *t*), and zero otherwise. *# of Supply Chain Q's* for an analyst-firm-year annual forecast observation is the analyst's number of questions that mention supply chain keywords in the past four years after excluding the questions from the firm's own conference calls. *Expertise* is a dummy variable that is equal to one if *# of Supply Chain Q's* of an analyst ranks within the top 10% among all analysts of the year. *Total # of Questions* is the total number of questions raised in conference calls by the analyst in the past four years. *Has Expertise (Top20 Topics)* is a dummy variable that equals one if the analyst has expertise in at least one of the top 20 topics (in terms of their frequency in earnings conference calls). *Coincide (Top20 Topics)* is a dummy variable that equals one if, among the top 20 topics, the topic of analyst expertise coincides with the peak information arrivals of this topic, and zero otherwise. *# of Firms Covered* is the number of firms the analyst follows in year *t*. *# of SIC3 Inds Covered* is the number of industries the analyst covers in the year, where industries are defined by 3-digit SIC codes. *LN(Experience)* is the natural logarithm of 1 plus the number of days between the date of the forecast and the date of the first forecast of the analyst in I/B/E/S. *LN(Horizon)* is the natural logarithm of 1 plus the number of days between the date of the forecast and the date of the actual earnings released. *LN(Firm Size)* is the natural logarithm of size, where size is defined as stock price times the total number of shares outstanding at the end of last year. *LN(BM)* is the natural logarithm of the firm's book value of equity divided by its market value at the end of the last fiscal year. *Profitability* is the operating income of the firm for fiscal year t-1, divided by the book value of equity at the end of fiscal year t-2. *LN(BM)* and *Profitability* are set to missing if the book value of equity is negative. *LN(Experience)*, *LN(Horizon)*, as well as other firm-level controls are winsorized at the 1% level.

VARIABLES	Mean	S.D.	25%	Median	75%
		Panel A: Main Variables and Analyst-Level Controls			
Abs(FE)	1.950	6.527	0.075	0.247	0.901
SUP	0.250	0.433	0.000	0.000	1.000
CUS	0.315	0.464	0.000	0.000	1.000
# of Supply Chain Q's	37.921	70.387	0.000	3.000	49.000
Expertise	0.100	0.301	0.000	0.000	0.000
Total # of Questions	398.870	668.727	0.000	58.000	600.000
Coincide (Top20 Topics)	0.027	0.162	0.000	0.000	0.000
Has Expertise (Top20 Topics)	0.195	0.396	0.000	0.000	0.000
# of Firms Covered	17.244	11.941	11.000	16.000	22.000
# of SIC3 Inds Covered	6.103	4.747	3.000	5.000	8.000
LN(Experience)	7.422	1.572	7.004	7.814	8.331
LN(Horizon)	4.364	0.991	3.912	4.585	4.796
		Panel B: Firm-Level Controls for Dependent Suppliers			
LN(Firm Size)	6.786	1.785	5.546	6.730	7.927
LN(BM)	-0.824	0.995	-1.352	-0.787	-0.284
Profitability	-0.038	0.520	-0.088	0.069	0.165
		Panel C: Firm-Level Controls for Principal Customers			
LN(Firm Size)	8.580	1.838	7.314	8.640	9.871
LN(BM)	-0.643	1.351	-1.343	-0.761	-0.214
Profitability	0.099	0.397	0.045	0.129	0.220

Table 2 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of analyst supply-chain-specific expertise. The dependent variable in all test specifications is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-specific questions in earnings conference calls. Analyst supply-chain-specific expertise in column (1) is constructed using 2 years of conference call data before year *t*. Similarly, this variable is constructed using 3 years and 4 years of conference call data in columns (2) and (3), respectively. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.075***	-0.076***	-0.079***
	(0.027)	(0.028)	(0.030)
$CUS \times Expertise$	0.040*	0.036	0.023
	(0.022)	(0.023)	(0.024)
$Expertise$	-0.049**	-0.018	0.029
	(0.023)	(0.025)	(0.026)
$Total\ \#\ of\ Questions$	-0.107***	-0.098***	-0.094***
	(0.014)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010*	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788

Table 3 Analyst Supply-Chain-Specific Expertise and the Market Reaction to Analyst Recommendation Updates

This table reports the market reactions to analysts' recommendation updates conditional on corporate supply-chain risk exposure and analyst supply-chain-specific expertise. In columns (1) and (2), the dependent variable is *DGTW AR[0]*, which is the DGTW-adjusted abnormal return on day 0 of the recommendation upgrades and downgrades. In columns (3) and (4), the dependent variable is *DGTW CAR[0,1]*, which is the DGTW-adjusted cumulative abnormal returns for two days after recommendation updates. *DGTW AR[0]* and *DGTW CAR[0,1]* are multiplied by 100 for ease of interpretation. Columns (1) and (3) report results for recommendation upgrades, while columns (2) and (4) report results for downgrades. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-specific questions in earnings conference calls. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *Change in Recommendation* is the difference between this recommendation and the prior recommendation. We include a dummy variable (*Two Weeks Before Earnings*) that is equal to one if the recommendation update is issued within the two-week window before the firm's earnings announcement, and zero otherwise, as well as a dummy variable (*Two Weeks After Earnings*) that is equal to one if the recommendation update is issued within the two-week window after the firm's earnings announcement, and zero otherwise. *Away from Consensus* is a dummy variable that is equal to one if the absolute difference between this recommendation and the consensus is larger than that of the prior recommendation. Definitions of other control variables are provided in Appendix I. Firm, year, and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1) DGTW AR [0]	(2) DGTW AR [0]	(3) DGTW CAR [0,1]	(4) DGTW CAR [0,1]
	Upgrade	Downgrade	Upgrade	Downgrade
$SUP \times Expertise$	0.452***	-0.463***	0.447**	-0.489**
	(0.173)	(0.172)	(0.208)	(0.195)
$Expertise$	-0.095	0.114	-0.212	0.028
	(0.117)	(0.144)	(0.142)	(0.163)
SUP	-0.188	-0.031	-0.249	-0.022
	(0.247)	(0.153)	(0.276)	(0.188)
$Change\ in\ Recommendation$	-0.588***	-0.446***	-0.672***	-0.644***
	(0.113)	(0.079)	(0.131)	(0.095)
$Two\ Weeks\ Before\ Earnings$	0.017	-0.253	-0.082	-0.151
	(0.124)	(0.156)	(0.146)	(0.178)
$Two\ Weeks\ After\ Earnings$	-0.048	-0.078	-0.050	-0.015
	(0.108)	(0.147)	(0.124)	(0.172)
$Away\ from\ Consensus$	0.258***	-0.516***	0.311***	-0.657***
	(0.085)	(0.065)	(0.096)	(0.077)
$Past\ Accuracy$	0.052***	-0.000	0.061**	-0.000
	(0.020)	(0.000)	(0.031)	(0.001)
$\#\ of\ Firms\ Covered$	0.017	-0.017	0.009	-0.034***
	(0.011)	(0.011)	(0.012)	(0.012)
$\#\ of\ SIC3\ Inds\ Covered$	-0.013	0.038	0.009	0.050*
	(0.022)	(0.024)	(0.027)	(0.030)
$LN(Experience)$	-0.107	-0.001	-0.030	0.100
	(0.092)	(0.106)	(0.117)	(0.125)
$Total\ \#\ of\ Questions$	0.131**	0.046	0.185***	0.031
	(0.052)	(0.067)	(0.063)	(0.081)
$Brokerage\ Firm\ Size$	0.010***	-0.009***	0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
$LN(Firm\ Size)$	-0.641***	0.222**	-0.826***	0.249**
	(0.132)	(0.087)	(0.145)	(0.106)
$LN(BM)$	0.289*	0.017	0.326*	0.024
	(0.151)	(0.086)	(0.170)	(0.099)
$Profitability$	0.029	0.002	0.046	0.004
	(0.042)	(0.002)	(0.047)	(0.002)
$\#\ of\ Analysts\ Covering$	-0.005	0.003	-0.018**	0.004
	(0.008)	(0.008)	(0.009)	(0.010)
Analyst FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	32,601	33,240	32,591	33,235
Adj R-squared	0.313	0.206	0.296	0.204

Table 4 The 2018 US-China Trade War, Analyst Supply-Chain-Specific Expertise, and Analyst Forecast Accuracy

This table reports the effects of analyst supply-chain-specific expertise on forecast accuracy around the 2018 US-China trade war. The dependent variable in all columns is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month prior to the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-specific questions in earnings conference calls. *ChinaImport* is a dummy variable that is equal to one if a firm imports from China in year *t*, and zero otherwise. *ChinaExport* is a dummy variable that is equal to one if a firm exports to China in year *t*, and zero otherwise. *ChinaTrade* is a dummy variable equal to one if a firm either imports from or exports to China in year *t*, and zero otherwise. *Post* is a dummy variable that is equal to one if the time is 2018 or onward, and zero otherwise. Due to the data availability of the Panjiva database, the sample in column (1) starts in 2007, whereas columns (2) and (3) start in 2009. The sample period ends in 2019, prior to the outbreak of the COVID-19 pandemic. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

VARIABLES	(1)	(2)	(3)
$ChinaImport \times Post \times Expertise$	-0.192***		
	(0.071)		
$ChinaImport \times Expertise$	0.035		
	(0.031)		
$ChinaExport \times Post \times Expertise$		-0.286*	
		(0.159)	
$ChinaExport \times Expertise$		0.023	
		(0.036)	
$ChinaTrade \times Post \times Expertise$			-0.237***
			(0.075)
$ChinaTrade \times Expertise$			0.037
			(0.031)
$Post \times Expertise$	0.074	0.023	0.116
	(0.070)	(0.052)	(0.072)
$Expertise$	0.021	0.028	0.012
	(0.034)	(0.031)	(0.037)
$Total\ \#\ of\ Questions$	-0.125***	-0.157***	-0.157***
	(0.018)	(0.021)	(0.022)
$\#\ of\ Firms\ Covered$	-0.013***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)
$\#\ of\ SIC3\ Inds\ Covered$	0.006	0.004	0.004
	(0.004)	(0.004)	(0.004)
$LN(Experience)$	0.008	-0.000	-0.000
	(0.006)	(0.006)	(0.006)
$LN(Horizon)$	0.366***	0.302***	0.302***
	(0.012)	(0.011)	(0.011)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	469,980	397,632	397,632
Adj R-squared	0.824	0.802	0.802

Table 5 The 2018 US-China Trade War, Analyst Supply-Chain-Specific Expertise, and Labor Market Outcomes

This table presents the effects of analyst supply-chain-specific expertise on the career outcomes of analysts during the 2018 US-China Trade War. The analysis is conducted at the analyst-year level. Column (1) examines the probability of internal promotion, conditional on an analyst staying at the same brokerage firm over the next two years. The dependent variable *Internal Promotion* is a dummy variable that is equal to one if an analyst's job title becomes more senior at the same brokerage firm over the next two years, and zero otherwise. Column (2) examines an analyst's probability of joining a buy-side firm, conditional on experiencing job turnovers over the next two years. The sample for this column includes analysts who leave their current brokerage firm over the next two years. The dependent variable, *Join Buy-side*, is a dummy variable that is equal to one if the analyst joins a buy-side firm, and zero otherwise. Column (3) examines an analyst's probability of joining a larger sell-side brokerage firm, conditional on the analyst leaving their current brokerage firm and joining another sell-side brokerage firm. The sample for this column includes all analysts who depart from their current brokerage firm and move to another sell-side brokerage firm over the next two years. The dependent variable, *Join Larger Brokerage Firm*, is a dummy variable that is equal to one if the analyst joins a sell-side brokerage firm that is larger (in terms of the number of analysts) than their current employer, and zero otherwise. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-specific questions in earnings conference calls. *Post* is a dummy variable that is equal to one for years 2018 and onward, and zero otherwise. Definitions of control variables are provided in Appendix I. Standard errors, reported in parentheses, are clustered at the brokerage firm level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively

VARIABLES	(1) Analysts with Existing Employers	(2) Analysts with Job Turnovers	(3) Analysts Joining Other Brokerage Firms
	Internal Promotion	Join Buy-Side	Join Larger Brokerage Firm
$Post \times Expertise$	0.008**	0.068**	0.144***
	(0.004)	(0.029)	(0.054)
$Expertise$	-0.003	-0.012	0.033
	(0.003)	(0.011)	(0.028)
$Total\ \#\ of\ Questions$	0.001	-0.012	-0.009
	(0.002)	(0.008)	(0.013)
$\#\ of\ Firms\ Covered$	-0.000	-0.003***	0.000
	(0.000)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.000	-0.003**	-0.003
	(0.000)	(0.001)	(0.003)
$LN(Experience)$	-0.000	0.004***	-0.008**
	(0.001)	(0.002)	(0.004)
Brokerage×Year FE	Yes	Yes	Yes
Analyst FE	Yes	No	No
SE Clustered at Brokerage Level	Yes	Yes	Yes
Observations	38,903	6,928	5,029
Adj R-squared	0.333	0.056	0.337

Table 6 Analyst Topic-Specific Expertise, Peak of Information Arrivals, and Forecast Accuracy

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of analyst topic-specific expertise and information arrivals. The dependent variable in all test specifications is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. *Has Expertise* is a dummy variable that equals one if the analyst has expertise in at least one of the top 10, 15, and 20 topics (in terms of their frequency in earnings conference calls) in columns (1), (2), and (3), respectively. For each firm–topic–year, the peak information arrival period is defined as the year when the share of questions related to a given topic in the firm's earnings calls exceeds the sample-wide mean share of that topic by more than two standard deviations. *Coincide* is a dummy variable equal to one if at least one of the analyst's expertise topics coincides with a peak information arrival for the firm in that year, and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)
		Has Expertise in	
VARIABLES	Top 10 Topics	Top 15 Topics	Top 20 Topics
$Coincide$	-0.085***	-0.073**	-0.059**
	(0.030)	(0.029)	(0.028)
$Has\ Expertise$	-0.049***	-0.061***	-0.073***
	(0.018)	(0.018)	(0.017)
$Total\ \#\ of\ Questions$	-0.080***	-0.077***	-0.074***
	(0.013)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)
$LN(Experience)$	0.009	0.009*	0.009*
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.410***	0.411***	0.411***
	(0.012)	(0.012)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	598,225	598,225	598,225
Adj R-squared	0.788	0.788	0.788

Table 7 Analyst Topic-Specific Expertise, Peak of Information Arrivals, and Market Reaction to Analyst Recommendation Upgrades and Downgrades

This table estimates the market reaction to topic-specific expert analysts' recommendation updates. In columns (1) and (2), the dependent variable is *DGTW AR[0]*, which is the DGTW-adjusted abnormal return on day 0 of the recommendation upgrades and downgrades. In columns (3) and (4), the dependent variable is *DGTW CAR[0,1]*, which is the DGTW-adjusted cumulative abnormal returns for two days after recommendation updates. *DGTW AR[0]* and *DGTW CAR[0,1]* are multiplied by 100 for ease of interpretation. Columns (1) and (3) report results for recommendation upgrades, while columns (2) and (4) report results for downgrades. Among independent variables, *Has Expertise* is a dummy variable that equals one if the analyst has expertise in at least one of the top 20 topics. For each firm–topic–year, the peak information arrival period is defined as the year when the share of questions related to a given topic in the firm's earnings calls exceeds the sample-wide mean share of that topic by more than two standard deviations. The independent variable *Peak* is defined at the firm-year level and equals one if any of the top 20 topics is in a peak information arrival period for that firm in year *t*. *Coincide* is a dummy variable equal to one if at least one of the analyst's expertise topics coincides with a peak information arrival for the firm in that year, and zero otherwise. *Change in Recommendation* is the difference between this recommendation and the prior recommendation. We include a dummy variable (*Two Weeks Before Earnings*) that is equal to one if the recommendation update is issued within the two-week window before the firm's earnings announcement, and zero otherwise, as well as a dummy variable (*Two Weeks After Earnings*) that is equal to one if the recommendation update is issued within the two-week window after the firm's earnings announcement, and zero otherwise. *Away from Consensus* is a dummy variable that is equal to one if the absolute difference between this recommendation and the consensus is larger than that of the prior recommendation. Definitions of other control variables are provided in Appendix I. Firm, year, and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively

VARIABLES	(1) DGTW AR [0]	(2) DGTW AR [0]	(3) DGTW CAR [0,1]	(4) DGTW CAR [0,1]
	Upgrade	Downgrade	Upgrade	Downgrade
Coincide	0.322**	-0.306*	0.395**	-0.455**
	(0.157)	(0.166)	(0.178)	(0.197)
Has Expertise	-0.037	0.110	-0.078	0.028
	(0.084)	(0.092)	(0.101)	(0.107)
Peak	-0.081	0.044	-0.084	0.136
	(0.105)	(0.115)	(0.123)	(0.141)
Change in Recommendation	-0.588***	-0.447***	-0.671***	-0.644***
	(0.113)	(0.079)	(0.131)	(0.095)
2 *Weeks Before Earnings*	0.016	-0.248	-0.083	-0.146
	(0.123)	(0.156)	(0.146)	(0.178)
2 *Weeks After Earnings*	-0.050	-0.078	-0.051	-0.014
	(0.107)	(0.147)	(0.124)	(0.172)
Away from Consensus	0.258***	-0.515***	0.311***	-0.656***
	(0.085)	(0.065)	(0.096)	(0.077)
Past Accuracy	0.053***	-0.000	0.062**	-0.000
	(0.020)	(0.000)	(0.031)	(0.001)
# *of Firms Covered*	0.017	-0.017	0.009	-0.034***
	(0.011)	(0.011)	(0.012)	(0.012)
# *of SIC*3 *Inds Covered*	-0.012	0.038	0.010	0.050*
	(0.022)	(0.024)	(0.027)	(0.030)
LN(*Experience*)	-0.107	-0.008	-0.033	0.100
	(0.091)	(0.106)	(0.117)	(0.126)
Total # *of Questions*	0.127**	0.033	0.172***	0.024
	(0.054)	(0.065)	(0.064)	(0.080)
Brokerage Firm Size	0.010***	-0.010***	0.008***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
LN(*Firm Size*)	-0.643***	0.223**	-0.827***	0.250**
	(0.132)	(0.087)	(0.145)	(0.106)
LN(*BM*)	0.289*	0.017	0.325*	0.022
	(0.150)	(0.086)	(0.169)	(0.099)
Profitability	0.029	0.002	0.046	0.004
	(0.042)	(0.002)	(0.047)	(0.002)
# *of Analysts Covering*	-0.005	0.004	-0.019**	0.004
	(0.008)	(0.008)	(0.009)	(0.010)
Analyst FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	32,597	33,232	32,587	33,227
Adj R-squared	0.312	0.206	0.296	0.204

Table 8 Analyst Topic-Specific Expertise, Peak of Information Arrivals, and Analyst Labor Market Outcomes

This table evaluates the effects of analyst topic-specific expertise and peak information arrivals on analyst career outcomes. The analysis is conducted at the analyst-year level. Column (1) examines the probability of internal promotion, conditional on an analyst staying at the same brokerage firm over the next two years. The dependent variable *Internal Promotion* is a dummy variable that is equal to one if an analyst's job title becomes more senior at the same brokerage firm over the next two years, and zero otherwise. Column (2) examines an analyst's probability of joining a buy-side firm, conditional on experiencing job turnovers over the next two years. The sample for this column includes analysts who leave their current brokerage firm over the next two years. The dependent variable, *Join Buy-side*, is a dummy variable that is equal to one if the analyst joins a buy-side firm, and zero otherwise. Column (3) examines an analyst's probability of joining a larger sell-side brokerage firm, conditional on the analyst leaving their current brokerage firm and joining another sell-side brokerage firm. The sample for this column includes all analysts who depart from their current brokerage firm and move to another sell-side brokerage firm over the next two years. The dependent variable, *Join Larger Brokerage Firm*, is a dummy variable that is equal to one if the analyst joins a sell-side brokerage firm that is larger (in terms of the number of analysts) than their current employer, and zero otherwise. Among independent variables, *Has Expertise* is a dummy variable that equals one if the analyst has expertise in at least one of the top 20 topics. The independent variable *Peak of Covered Firms* is defined at the analyst–year level and equals one if any firm covered by the analyst experiences a peak information arrival in any of the top 20 topics during year *t*. *Coincide* is a dummy variable equal to one if at least one of the analyst's expertise topics coincides with a peak information arrival for the covered firms in that year, and zero otherwise. Definitions of control variables are provided in Appendix I. Standard errors, reported in parentheses, are clustered at the brokerage firm level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

VARIABLES	(1) Analysts with Existing Employers	(2) Analysts with Job Turnovers	(3) Analysts Joining Other Brokerage Firms
	Internal Promotion	Join Buy-Side	Join Larger Brokerage Firm
$Coincide$	0.008**	0.027**	0.067**
	(0.003)	(0.013)	(0.030)
$Has\ Expertise$	-0.002	-0.003	0.066***
	(0.003)	(0.011)	(0.023)
$Peak\ of\ Covered\ Firms$	0.000	-0.018**	-0.001
	(0.002)	(0.008)	(0.015)
$Total\ \#\ of\ Questions$	-0.001	-0.016*	-0.033***
	(0.002)	(0.008)	(0.013)
$\#\ of\ Firms\ Covered$	-0.000	-0.003***	-0.000
	(0.000)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.000	-0.002**	-0.002
	(0.000)	(0.001)	(0.003)
$LN(Experience)$	-0.000	0.005***	-0.010**
	(0.001)	(0.001)	(0.004)
Brokerage×Year FE	Yes	Yes	Yes
Analyst FE	Yes	No	No
SE Clustered at Brokerage Level	Yes	Yes	Yes
Observations	38,903	6,928	5,029
Adj R-squared	0.342	0.057	0.342

Table 9 Analyst Expertise Diversity and Price Efficiency

This table examines the relationship between analyst expertise diversity and price efficiency at the firm-year level. The dependent variable is the corporate earnings in year *t+1*, scaled by the total assets in year *t*. It is multiplied by 100 to express values in percentage terms. *LogMVA* is the natural logarithm of the ratio between the firm's market capitalization and total assets in year *t*. Analyst expertise diversity is proxied by two measures: *# of Expertise* counts the number of distinct topic-specific expertise held by analysts covering a firm in year *t*, based on the top 20 most frequently discussed topics. When we compute the second measure, *# of Rep Expertise*, we restrict each analyst to a single expertise based on the analyst's most frequently discussed topic. Definitions of other control variables are provided in Appendix I. Firm and year fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

VARIABLES	(1) $Earnings_{t+1}$	(2) $Earnings_{t+1}$
$LogMVA$	0.890***	0.900***
	(0.240)	(0.240)
$LogMVA \times \#\ of\ Expertise$	0.039***	
	(0.013)	
$\#\ of\ Expertise$	-0.032**	
	(0.015)	
$LogMVA \times \#\ of\ Rep\ Expertise$		0.094***
		(0.035)
$\#\ of\ Rep\ Expertise$		-0.073*
		(0.044)
$LogMVA \times \#\ of\ Analysts\ Covering$	0.140***	0.139***
	(0.010)	(0.010)
$\#\ of\ Analysts\ Covering$	-0.053***	-0.053***
	(0.015)	(0.016)
$LogMVA \times Leverage$	-1.329***	-1.333***
	(0.373)	(0.374)
$LogMVA \times Tangibility$	0.535**	0.523**
	(0.219)	(0.218)
$LogMVA \times Cash$	-5.185***	-5.191***
	(0.558)	(0.558)
$LogMVA \times Sale$	1.608***	1.616***
	(0.141)	(0.141)
$LogMVA \times R\&D$	-13.326***	-13.348***
	(1.166)	(1.166)
$LN(Total\ Assets)$	1.788***	1.784***
	(0.204)	(0.204)
$Leverage$	5.366***	5.374***
	(0.690)	(0.690)
$Tangibility$	1.502**	1.509**
	(0.616)	(0.616)
$Cash$	0.779	0.787
	(0.773)	(0.773)
$Sale$	3.681***	3.675***
	(0.378)	(0.378)
$R\&D$	8.549***	8.566***
	(1.768)	(1.769)
$Current\ Earnings$	44.325***	44.336***
	(1.174)	(1.174)
Firm FE	Yes	Yes
Year FE	Yes	Yes
SE Clustered at Firm Level	Yes	Yes
Observations	78,243	78,243
Adj R-squared	0.840	0.840

Table 10 Analyst Expertise Diversity and Buy-Side Participation in Earnings Conference Calls

This table examines the relationship between analyst expertise diversity and buy-side participation in earnings conference calls. The sample includes firm-year observations in which at least one earnings conference call is hosted by the firm. Among all dependent variables, *BuySideQuestion%* is defined as the percentage of questions raised by buy-side analysts, calculated as the number of questions asked by buy-side analysts divided by the total number of questions asked during a firm's earnings conference calls in year *t*, multiplied by 100. *BuySideParticipant%* is defined as the percentage of participants who are buy-side analysts, calculated as the number of unique buy-side analysts who raised questions divided by the total number of unique participants who asked questions during the firm's earnings conference calls in year *t*, multiplied by 100. Analyst expertise diversity is proxied by two measures: *# of Expertise* counts the number of distinct topic-specific expertise held by analysts covering a firm in year *t*, based on the top 20 most frequently discussed topics. When we compute the second measure, *# of Rep Expertise*, we restrict each analyst to a single expertise based on her most frequently discussed topic. Definitions of other control variables are provided in Appendix I. Firm and year fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

VARIABLES	(1) BuySideQuestion%	(2) BuySideParticipant%	(3) BuySideQuestion%	(4) BuySideParticipant%
# of Expertise	-0.052***	-0.051***		
	(0.011)	(0.010)		
# of Rep Expertise			-0.064***	-0.062***
			(0.022)	(0.021)
# of Analysts Covering	-0.050***	-0.055***	-0.053***	-0.058***
	(0.007)	(0.007)	(0.007)	(0.007)
LN(Total Assets)	0.161	0.094	0.155	0.088
	(0.104)	(0.097)	(0.104)	(0.097)
Leverage	0.233	0.103	0.243	0.113
	(0.248)	(0.231)	(0.248)	(0.231)
Tangibility	0.197	0.237	0.211	0.251
	(0.252)	(0.226)	(0.252)	(0.227)
Cash	0.189	0.205	0.194	0.209
	(0.268)	(0.258)	(0.268)	(0.259)
Q	0.022	0.021	0.019	0.018
	(0.033)	(0.030)	(0.033)	(0.030)
ROA	0.047	0.109	0.053	0.114
	(0.210)	(0.190)	(0.210)	(0.190)
ROA Std	0.547	0.566*	0.561	0.579*
	(0.399)	(0.339)	(0.399)	(0.339)
Ret	1.580**	1.066	1.627**	1.111*
	(0.704)	(0.667)	(0.705)	(0.668)
Ret Std	0.839	0.356	1.046	0.558
	(2.772)	(2.479)	(2.774)	(2.479)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	54,544	54,544	54,544	54,544
Adj R-squared	0.270	0.274	0.270	0.273

Table 11 Analyst Expertise Diversity and Investment Sensitivity to Stock Prices

This table examines the relationship between analyst expertise diversity and investment-to-price sensitivity at the firm-year level. RND_{t+1} is calculated by the R&D expense in year *t+1* scaled by the beginning-of-year total assets. $CHGASSET_{t+1}$ is defined as the change in assets in year *t+1* scaled by the beginning-of-year total assets. $CAPXRND_{t+1}$ is defined as capital expenditure plus R&D in year *t+1* scaled by the beginning-of-year total assets. Q is the firm's Tobin's Q in year *t* which is calculated as market value of equity plus book value of assets minus book value of equity, scaled by book value of assets. Analyst expertise diversity is proxied by two measures: *# of Expertise* counts the number of distinct topic-specific expertise held by analysts covering a firm in year *t*, based on the top 20 topics. When we compute the second measure, *# of Rep Expertise*, we restrict each analyst to a single expertise based on the analyst's most frequently discussed topic. Other control variables, such as firm's future returns (Ret_{t+3}), sales growth rate (*SG*), cash (*Cash*), leverage (*Leverage*), return on assets (*ROA*), and inverse book assets (*1/Assets*) are also included in all test specifications. Firm and year fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

VARIABLES	(1) RND_{t+1}	(2) $CHGASSET_{t+1}$	(3) $CAPXRND_{t+1}$	(4) RND_{t+1}	(5) $CHGASSET_{t+1}$	(6) $CAPXRND_{t+1}$
Q	1.003***	9.105***	1.554***	0.981***	9.040***	1.529***
	(0.081)	(0.413)	(0.094)	(0.079)	(0.410)	(0.091)
$Q \times \#\ of\ Expertise$	-0.050***	-0.159***	-0.064***			
	(0.009)	(0.047)	(0.010)			
$Q \times \#\ of\ Rep\ Expertise$				-0.135***	-0.358***	-0.185***
				(0.025)	(0.137)	(0.028)
$Q \times \#\ of\ Analysts\ Covering$	-0.003	-0.110***	-0.003	0.002	-0.102***	0.006
	(0.004)	(0.026)	(0.005)	(0.005)	(0.029)	(0.006)
$\#\ of\ Expertise$	0.063***	0.041	0.113***			
	(0.013)	(0.086)	(0.017)			
$\#\ of\ Rep\ Expertise$				0.154***	-0.186	0.314***
				(0.039)	(0.245)	(0.049)
$\#\ of\ Analysts\ Covering$	-0.023**	-0.444***	-0.087***	-0.028***	-0.420***	-0.099***
	(0.010)	(0.058)	(0.013)	(0.010)	(0.063)	(0.014)
$CF \times \#\ of\ Expertise$	0.242**	2.630***	0.206*			
	(0.099)	(0.651)	(0.114)			
$CF \times \#\ of\ Rep\ Expertise$				0.797**	7.017***	0.646*
				(0.316)	(1.899)	(0.363)
$CF \times \#\ of\ Analysts\ Covering$	0.081*	1.709***	0.330***	0.039	1.478***	0.295***
	(0.044)	(0.330)	(0.055)	(0.053)	(0.358)	(0.065)
CF	1.628***	35.438***	3.816***	1.634***	35.889***	3.830***
	(0.555)	(3.841)	(0.667)	(0.552)	(3.835)	(0.663)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes	Yes	Yes
Observations	70,789	70,244	70,789	70,789	70,244	70,789
Adj R-squared	0.863	0.225	0.797	0.863	0.225	0.797

Online Appendix Tables

Table O1 Summary Statistics of Analysts With and Without Supply-Chain Expertise

This table reports summary statistics for analysts with and without supply-chain expertise over the period 2002–2022. *# of Firms Covered* is the number of firms the analyst follows in year *t*. *# of SIC3 Inds Covered* is the number of industries the analyst covers in the year, where industries are defined by 3-digit SIC codes. *LN(Experience)* is the natural logarithm of 1 plus the number of days between the date of the forecast and the date of the first forecast of the analyst in I/B/E/S. *LN(Horizon)* is the natural logarithm of 1 plus the number of days between the date of the forecast and the date of the actual earnings released. The last column reports the mean difference between experts and non-experts, with statistical significance based on t-tests of equality in means. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	Expert		Non-Expert		Difference
VARIABLES	Mean	S.D.	Mean	S.D.	
# of Firms Covered	19.889	6.851	16.948	12.345	2.941***
# of SIC3 Inds Covered	8.426	4.203	5.844	4.733	2.581***
LN(Experience)	8.111	0.757	7.345	1.619	0.767***
LN(Horizon)	4.369	0.906	4.363	1.000	0.006
Brokerage Size	49.510	39.159	56.160	46.959	-6.650***

Table O2 Past Forecast Accuracy vs. Supply-Chain-Specific Expertise

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of analyst past forecast accuracy ranking and analyst supply-chain-specific skill. The dependent variable in all test specifications is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month prior to the actual earnings announcement. *High Accuracy* is a dummy variable that is equal to one if the forecast accuracy of the analyst for the SIC3 industry ranks above 90th percentile in the past two, three, and four years for columns (1), (2), and (3), respectively. Analyst supply-chain-specific skill in column (1) is constructed using 2 years of conference call data before year *t*. Similarly, this variable is constructed using 3 years and 4 years of conference call data in columns (2) and (3), respectively. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year t (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.075***	-0.076***	-0.080***
	(0.027)	(0.028)	(0.030)
$CUS \times Expertise$	0.040*	0.036	0.023
	(0.022)	(0.023)	(0.024)
$Expertise$	-0.049**	-0.018	0.029
	(0.023)	(0.025)	(0.026)
$SUP \times High\ Accuracy$	-0.004	-0.011	-0.016
	(0.037)	(0.041)	(0.044)
$CUS \times High\ Accuracy$	0.022	0.017	0.033
	(0.024)	(0.026)	(0.027)
$High\ Accuracy$	-0.036**	-0.038**	-0.037*
	(0.018)	(0.020)	(0.021)
$Total\ \#\ of\ Questions$	-0.107***	-0.098***	-0.094***
	(0.014)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010**	0.009*
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788

Table O3 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy: Using a New Weight of Questions Based on the Length of Replies

This table estimates the baseline tests based on the expertise measure using a new weight for each supply-chain-related question, i.e., the number of words in the reply of each supply-chain-related question scaled by the average number of words in the reply based on all replies for questions raised in earnings conference calls of this firm in the evaluation period. The dependent variable in these regressions is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings divided by stock price of the firm at the end of the month prior to actual earnings announcement. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.060**	-0.077***	-0.075**
	(0.028)	(0.028)	(0.029)
$CUS \times Expertise$	0.039*	0.038*	0.024
	(0.022)	(0.022)	(0.024)
$Expertise$	-0.028	-0.019	0.022
	(0.023)	(0.024)	(0.025)
$Total\ \#\ of\ Questions$	-0.111***	-0.098***	-0.093***
	(0.014)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010**	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788

Table O4 The Timing of Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy

This table presents coefficients from OLS regressions estimating how quickly supply-chain-specific expertise is incorporated into earnings forecasts around the firms' first-time disclosure of their principal customers. The dependent variable in all test specifications is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. SUP^{-1} equals one in year t if the firm discloses its public principal customer for the first time in year *t+1* and zero otherwise. SUP^{0} equals one in year *t* if the firm discloses its public principal customer for the first time in year t and zero otherwise. SUP^{1} and SUP^{2+} are defined similarly. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP^{-1} \times Expertise$	-0.012	-0.008	0.022
	(0.096)	(0.094)	(0.095)
$SUP^{0} \times Expertise$	-0.132**	-0.118*	-0.127**
	(0.051)	(0.060)	(0.063)
$SUP^{1} \times Expertise$	-0.067	-0.086	-0.077
	(0.080)	(0.085)	(0.090)
$SUP^{2+} \times Expertise$	-0.065**	-0.066**	-0.069**
	(0.030)	(0.031)	(0.032)
$CUS \times Expertise$	0.041*	0.036	0.023
	(0.022)	(0.023)	(0.024)
$Expertise$	-0.049**	-0.018	0.028
	(0.023)	(0.025)	(0.026)
$Total\ \#\ of\ Questions$	-0.107***	-0.098***	-0.094***
	(0.014)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010*	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788

Table O5 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy: Using A Conservative Keyword List

This table estimates the effect of analyst supply-chain-specific expertise, constructed using a conservative list, and analyst forecast error. The conservative list includes only nine keywords: "supply chain", "supplier", "outsource", "inventory", "export", "import", "shipping", "procure", and "logistics". The dependent variable in these regressions is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings divided by stock price of the firm at the end of the month prior to actual earnings announcement. SUP is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year t (the firm discloses at least one public principal customer), and zero otherwise. CUS is a dummy variable that is equal to one if the firm is a principal customer in year t (the firm is disclosed as principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.064**	-0.067**	-0.072**
	(0.031)	(0.033)	(0.035)
$CUS \times Expertise$	0.021	0.036	0.003
	(0.025)	(0.028)	(0.029)
$Expertise$	-0.028	-0.004	0.020
	(0.025)	(0.030)	(0.031)
$Total\ \#\ of\ Questions$	-0.112***	-0.101***	-0.092***
	(0.013)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010*	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788

Table O6 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy: Alternative Cutoffs

This table reports the effect of analyst supply-chain-specific expertise, constructed using alternative cutoff levels of number of keywords, on analyst forecast errors. The dependent variable in all test specifications is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. *Expertise* is a dummy variable that is equal to one if an analyst is classified as an expert analyst according to the analyst's supply-chain-specific questions in earnings conference calls over the past four years. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)	(4)
	Expertise Measured with the Cutoff of			
VARIABLES	Top8%	Top3%	Top2%	Top1%
$SUP \times Expertise$	-0.052*	-0.108**	-0.101*	-0.160**
	(0.031)	(0.048)	(0.054)	(0.065)
$CUS \times Expertise$	0.026	0.027	0.013	0.050
	(0.025)	(0.037)	(0.041)	(0.054)
$Expertise$	0.017	0.027	0.046	0.034
	(0.028)	(0.040)	(0.044)	(0.050)
$Total\ \#\ of\ Questions$	-0.094***	-0.092***	-0.093***	-0.092***
	(0.013)	(0.013)	(0.013)	(0.012)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.006	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)	(0.004)
$LN(Experience)$	0.008	0.008	0.008	0.008
	(0.005)	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.409***	0.409***	0.409***	0.409***
	(0.012)	(0.012)	(0.012)	(0.012)
Analyst FE	Yes	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes	Yes
Observations	598,225	598,225	598,225	598,225
Adj R-squared	0.788	0.788	0.788	0.788

Table O7 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy: Using the Breadth Instead of the Number of Supply-Chain Questions

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of the breadth of analyst supply-chain questions, measured as the fraction of covered firms in which the analyst raises at least one supply-chain-related question. The sample for column (2) additionally excludes analysts with fewer than five total questions raised in the prior four years. The dependent variable is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month prior to the actual earnings announcement. *TopBreadth* equals one if the fraction of covered firms in which the analyst asks supply-chain questions ranks within the top 10% among all analysts of the year. *SUP* is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year *t* (i.e., the firm discloses at least one public principal customer), and zero otherwise. *CUS* is a dummy variable that is equal to one if the firm is a principal customer in year *t* (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications and standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

VARIABLES	(1)	(2)
$SUP \times TopBreadth$	-0.085***	-0.102***
	(0.033)	(0.036)
$CUS \times TopBreadth$	0.024	0.041
	(0.025)	(0.028)
$TopBreadth$	-0.074***	-0.028
	(0.022)	(0.025)
$Total\ \#\ of\ Questions$	-0.090***	-0.076***
	(0.012)	(0.012)
$\#\ of\ Firms\ Covered$	-0.007***	-0.013***
	(0.001)	(0.002)
$\#\ of\ SIC3\ Inds\ Covered$	0.006	0.000
	(0.004)	(0.005)
$LN(Experience)$	-0.005	0.007
	(0.006)	(0.010)
$LN(Horizon)$	0.409***	0.436***
	(0.012)	(0.014)
Analyst FE	Yes	Yes
Firm×Year FE	Yes	Yes
SE Clustered at Firm Level	Yes	Yes
Observations	598,225	330,626
Adj R-squared	0.788	0.786

Table O8 Analyst Supply-Chain-Specific Expertise and Analyst Forecast Accuracy: Supply-Chain-Specific Expertise Measures Computed at the Analyst Level

This table reports the estimates from OLS regressions of the absolute analyst forecast error as a function of analyst supply-chain-specific expertise, measured at the analyst level. The dependent variable in all test specifications is *Abs(FE)*, which is 100 times the absolute difference between forecasted earnings and announced actual earnings, divided by the stock price of the firm at the end of the month before the actual earnings announcement. Analyst supply-chain-specific skill in column (1) is constructed using 2 years of conference call data before year *t*. The supply-chain-specific skill is constructed using all supply-chain-related questions raised by the analyst during the window. Similarly, this variable is constructed using 3 years and 4 years of conference call data in columns (2) and (3), respectively. SUP is a dummy variable that is equal to one if the firm is a dependent supplier for at least one public firm in year t (i.e., the firm discloses at least one public principal customer), and zero otherwise. CUS is a dummy variable that is equal to one if the firm is a principal customer in year t (the firm is disclosed as the principal customer by at least one supplier firm), and zero otherwise. Definitions of other control variables are provided in Appendix I. Firm×Year and analyst fixed effects are included in all test specifications. Standard errors, reported in parentheses, are clustered at the firm level. ***, **, and * indicate the 1%, 5%, and 10% levels of statistical significance, respectively.

	(1)	(2)	(3)
		Expertise Measured Over	
VARIABLES	2-years	3-years	4-years
$SUP \times Expertise$	-0.065**	-0.070**	-0.072**
	(0.028)	(0.028)	(0.030)
$CUS \times Expertise$	0.033	0.037	0.029
	(0.023)	(0.023)	(0.024)
$Expertise$	-0.047**	-0.008	0.019
	(0.023)	(0.025)	(0.026)
$Total\ \#\ of\ Questions$	-0.108***	-0.100***	-0.093***
	(0.013)	(0.013)	(0.013)
$\#\ of\ Firms\ Covered$	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
$\#\ of\ SIC3\ Inds\ Covered$	0.005	0.005	0.006
	(0.003)	(0.003)	(0.004)
$LN(Experience)$	0.011**	0.010*	0.008
	(0.005)	(0.005)	(0.005)
$LN(Horizon)$	0.401***	0.404***	0.409***
	(0.011)	(0.011)	(0.012)
Analyst FE	Yes	Yes	Yes
Firm×Year FE	Yes	Yes	Yes
SE Clustered at Firm Level	Yes	Yes	Yes
Observations	665,417	632,828	598,225
Adj R-squared	0.788	0.789	0.788