

What Makes a Star Analyst? Evidence on Industry Expertise from One Million Analyst Reports

Alejandro Lopez-Lira, Yuehua Tang, Yuan Wang, Mingyin Zhu*
University of Florida

This version: September 8, 2026

Abstract

We develop direct measures of sell-side analyst industry knowledge and research quality using large language models applied to 1.2 million industry reports from 40 major brokerage houses over 2012–2025. Each report is scored along 18 dimensions capturing research quality and specific forms of industry expertise. Our measures strongly predict *Institutional Investor* All-Star status, with effects roughly five times larger than those of forecast-accuracy measures. Analysts with higher knowledge and quality scores issue more accurate earnings forecasts, and their forecast revisions and recommendation changes elicit stronger market reactions. At the firm level, coverage by more knowledgeable analysts is associated with greater liquidity, more informative prices, and a lower implied cost of equity. Together, the evidence shows that industry knowledge embedded in analyst research captures an economically important dimension of analyst ability that is not well reflected in conventional quantitative outputs.

Keywords: Large Language Models, Sell-Side Analysts, Industry Knowledge, Analyst Reports, Information Environment

JEL Classification: C45, G12, G14, G17, G24

*We have benefited from comments and suggestions from Sean Cao, Jay Ritter, Baolian Wang, and Rong Wang. We thank An-Ping Lin for sharing the All-Star analyst status data, Kelvin Law for sharing the I/B/E/S brokerage matching data, and Baolian Wang for sharing the I/B/E/S analyst code to the full name link table. Contact information: Alejandro Lopez-Lira: alejandro.lopez-lira@warrington.ufl.edu; Yuehua Tang: yuehua.tang@warrington.ufl.edu; Yuan Wang: yuan.wang@warrington.ufl.edu; Mingyin Zhu: mingyin.zhu@warrington.ufl.edu. There are no financial disclosures or conflicts of interest.

1 Introduction

Sell-side analysts are in the business of producing information, but not all information is equally valuable. Because market prices already impound publicly available information, private value accrues only to information that is costly to acquire and difficult to replicate. Practitioners and consumers of analyst research point to industry knowledge as the most important attribute: an understanding of how a sector’s supply chains, margins, technology transitions, and competitive positions transmit shocks into firm earnings, built through channel checks, proprietary surveys, expert networks, and years of sector experience. In an influential survey of 365 analysts, Brown et al. (2015) find that industry knowledge is the single most useful input to both earnings forecasts and stock recommendations. Buy-side investors rank it above stock selection and earnings estimates (Bradshaw 2011); its value is reflected in investor performance and analyst compensation (Gokkaya et al. 2023; Groysberg, Healy, and Maber 2011); and recommendations reveal industry-based expertise in ranking firms within industries (Boni and Womack 2006) and calling industries themselves (Kadan et al. 2012). Taken together, these findings carry a clear message: what distinguishes star analysts is not the precision of their numbers but the depth of their industry knowledge.

Yet the object central to this account, the industry knowledge an analyst possesses, has never been measured directly. The literature evaluates analysts almost exclusively through quantitative outputs: earnings forecast accuracy, stock recommendation profitability, and target price informativeness. This emphasis is understandable because these outcomes are available in structured form, whereas the qualitative content of analyst research is not. But they were developed to study analyst performance and general ability and provide only indirect measures of industry knowledge. A forecast error, for example, reflects not only industry knowledge but also firm-specific forecasting skill, information access, and unanticipated shocks, and thus cannot isolate sector knowledge. A few studies examine industry expertise using proxies such as prior work experience (Bradley, Gokkaya, and Liu 2017) or the ability to identify attractive industries (Kadan et al. 2012). They show that industry expertise matters, but still infer knowledge from its correlates or consequences rather than measuring it directly. The knowledge itself resides in the qualitative content of analyst research, which until recently has been difficult to measure at scale because of data constraints.

This paper aims to measure industry knowledge directly. We do so using sell-side industry research. To our knowledge, this is the first study to systematically analyze industry reports. These reports provide a novel window into analysts’ sector-level knowledge and reasoning. Unlike firm-specific reports (e.g., Asquith, Mikhail, and Au (2005) and Décaire and Graham (2024)), which are typically organized around earnings forecasts and stock recommendations, industry reports focus on industry trends, technological change, supply chains, margins, competitive forces, and the outlook for the sector. They therefore reveal not only what analysts predict, but the knowledge and reasoning underlying those predictions. Industry research is also economically important: our sample contains 1.2 million reports produced by 40 major brokerage houses from January 2012 through October 2025, of the same order of magnitude as the volume of firm-specific research.

We use a large language model, OpenAI’s open-weight `gpt-oss-120b`, to score each report on an 18-dimension rubric. Ten dimensions capture research quality, including originality, analytical depth, and quantitative rigor; eight capture the forms of industry expertise that Brown et al. (2015) identify as most valuable, such as knowledge of supply chains, margins, customers, and technological change. These two sets of dimensions define our *Quality* and *Knowledge* measures, respectively. The rubric is designed to distinguish the repackaging of public information from substantive analysis and insight. Our measures thus capture the analytical content of analyst research rather than its length, coverage, or style. Aggregating report-level scores yields time-varying measures of analyst industry knowledge and research quality, which we relate to analyst recognition, forecast performance, market reactions, and the information environment of covered firms.

We begin by examining what the measures capture. *Knowledge* and *Quality* are highly persistent over time, consistent with stable differences in analyst ability rather than idiosyncratic features of individual reports. Analysts who produce more and longer research and cover larger firms score higher, while career length is related to both measures only before report characteristics are controlled for. Of particular interest is prior work experience in a covered industry, the proxy for industry expertise in Bradley, Gokkaya, and Liu (2017). Analysts with such experience score somewhat higher on both measures, but the differences are not statistically significant. Prior industry employment is therefore one route to acquiring industry knowledge, but not the only one; our measures capture the expertise analysts actually display in their research.

Our first main result concerns buy-side recognition. We use *Institutional Investor* All-Star status, the most prominent industry-wide measure of analyst reputation and a direct reflection of buy-side voting, as an external benchmark for whether our measures capture economically meaningful analyst ability. Both *Knowledge* and *Quality* strongly predict All-Star recognition. When entered individually, their standardized logit coefficients are about five times the magnitude of that on the traditional forecast-accuracy measure. The associations remain strong after controlling for past All-Star status, forecast accuracy, prior industry experience, and other analyst and brokerage characteristics. Prior industry experience also predicts recognition, consistent with Bradley, Gokkaya, and Liu (2017). Thus, prior industry credentials and the knowledge demonstrated in analysts’ research contain distinct information about the ability valued by buy-side investors, while forecast accuracy alone provides a much narrower signal of that ability.

The measures also predict analyst performance. Analysts with higher *Knowledge* and *Quality* issue more accurate earnings forecasts, including within analyst-firm relationships. Lagged forecast accuracy, by contrast, has no incremental explanatory power once the text-based measures are included. Investors appear to recognize these differences. The market responds about 14% more strongly to a given forecast revision following a one-standard-deviation increase in *Knowledge*, with a smaller but significant effect for *Quality*. Recommendation changes display the same pattern: a one-standard-deviation increase in *Knowledge* is associated with abnormal returns 35 basis points more negative around downgrades and 16 basis points more positive around upgrades. The larger response to downgrades is consistent with negative recommendations being more credible when analysts face incentives that discourage unfavorable calls.

A natural concern with an LLM-based measure is potential look-ahead bias: the model may have learned information about outcomes occurring after the reports it evaluates. We therefore repeat the forecast-accuracy and forecast-revision tests using only observations after `gpt-oss-120b`’s July 2024 knowledge cutoff. Both results remain statistically significant and are, if anything, stronger in magnitude. This post-cutoff evidence substantially alleviates concerns that the measures predict outcomes because the model already knows what subsequently occurred.

Industry reports also allow us to examine expertise at the industry level. We extract 44,687 explicit industry recommendations from the reports, more than four times the number avail-

able through I/B/E/S. The report-based data also identify revisions substantially earlier: among matched recommendation revisions, the I/B/E/S date lags the report-based date by about 90 days at the median and 115 days on average.¹ Among more than 2,000 recommendation revisions extracted from the industry reports, those issued by analysts with greater *Knowledge* predict significantly larger subsequent industry returns in the direction of the revision. A one-standard-deviation increase in *Knowledge* is associated with a 0.34 to 0.38 percentage point higher signed abnormal industry return over the following month across alternative return benchmarks. Thus, the knowledge captured from analyst research is informative not only about individual firms but also about the industries themselves.

Finally, we examine whether analyst knowledge is associated with the information environment of covered firms. We study three sets of outcomes: market liquidity, price informativeness, and the implied cost of equity. Firms covered by more knowledgeable analysts exhibit narrower effective and realized spreads, lower Amihud illiquidity, more balanced order flow, lower price impact, and greater price efficiency as measured by variance ratio. These associations are generally stronger for *Knowledge* than for *Quality* and are particularly pronounced among smaller firms, where external information production is plausibly more valuable. The financing implications are also economically meaningful. A one-standard-deviation increase in the *Knowledge* of a firm’s top analysts is associated with an approximately 8.5-basis-point lower implied cost of capital. Taken together, the results indicate that firms followed by analysts with deeper industry knowledge are characterized not only by more liquid trading, but also by prices that incorporate information more efficiently and a lower required return on equity.

To sharpen identification, we exploit the 2023 acquisition of Credit Suisse by UBS, which led to the removal of Credit Suisse analysts from firms they previously covered. In a difference-in-differences design, the deterioration in market quality is concentrated among firms that lose analysts with high pre-acquisition *Knowledge*. Effective spreads widen by 4.2 basis points, realized spreads by 2.4 basis points, and price impact and price inefficiency increase, whereas firms that lose low-*Knowledge* analysts experience little comparable deterioration. Event-study estimates show

1. The timing gap is consistent with I/B/E/S recording industry views only when they later appear in firm-level analyst reports, rather than when those views are first formed or communicated through industry research.

no differential pre-trends, and the effects disappear among firms with extensive analyst coverage. These findings support a causal interpretation in which knowledgeable analysts improve market quality precisely where their information is difficult to replace.

Contributions and Related Literature

Our paper makes three contributions to the literature by developing a direct, text-based measure of the industry knowledge embodied in analyst research.

First, we contribute to the literature on the “black box” of sell-side analyst ability. A large literature evaluates analysts through observable outputs such as forecast accuracy, recommendation profitability, and target price informativeness (O’Brien 1988, 1990; Clement 1999; Womack 1996; Brav and Lehavy 2003), while survey evidence identifies industry knowledge as one of the attributes most valued by analysts and their clients (Brown et al. 2015; Bradshaw 2011). Existing empirical work approaches industry expertise indirectly. Boni and Womack (2006) show that analyst skill is industry specific, Kadan et al. (2012) infer expertise from the informativeness of industry recommendations, and Bradley, Gokkaya, and Liu (2017), Bradley et al. (2017), and Gokkaya et al. (2023) use prior work experience in a covered industry to identify industry expertise. These approaches observe either potential sources of expertise or outcomes through which it is revealed, but not the knowledge itself. Our paper therefore speaks more broadly to how information intermediaries acquire, process, and transmit information to capital markets: we measure the knowledge content that lies between observable inputs and observable outputs. This allows us to examine directly how underlying expertise shapes analyst performance and capital-market outcomes.

Second, we contribute to the growing literature on the textual content of analyst research. Asquith, Mikhail, and Au (2005) show that qualitative arguments in analyst reports contain information beyond forecasts and recommendations; Huang, Zang, and Zheng (2014) extract textual opinions from analyst reports; Huang et al. (2018) use topic modeling to distinguish information discovery from interpretation; and more recent work applies natural-language-processing and machine-learning methods to additional dimensions of analyst research (Décaire and Graham 2024; Bastianello, Decaire, and Guenzel 2025; K. Li et al. 2026). We extend this literature along two dimensions. First, we study industry reports rather than firm-specific research, focusing on a large

component of sell-side output in which sector-level mechanisms, competitive dynamics, technological change, and cross-firm implications are particularly visible. Second, rather than extracting sentiment, topics, or isolated textual features, we use an LLM-based evaluation framework to assess economically defined attributes such as originality, analytical depth, mechanism-based reasoning, and specific forms of industry knowledge. Our approach therefore moves textual analysis from identifying what analysts discuss toward measuring the depth and economic content in their research.

Third, we contribute to the literature on analysts as information intermediaries and firms' information environments. Prior work shows that analysts facilitate the incorporation of information into prices (Piotroski and Roulstone 2004) and links analyst coverage and information production to return synchronicity (Crawford, Roulstone, and So 2012), financing costs (Bowen, Chen, and Cheng 2008), information asymmetry (Amiram, Owens, and Rozenbaum 2016), and the value of research in periods of high uncertainty (Loh and Stulz 2018). Related work emphasizes analyst effort and coverage intensity, with Harford et al. (2019) showing that effort allocation affects liquidity and the implied cost of capital. Exogenous coverage shocks further show that analyst information production has economically important effects on prices, liquidity, financing, and investment (Kelly and Ljungqvist 2012; Derrien and Kecskés 2013). We extend this literature by shifting attention from the quantity of analyst coverage and effort to the knowledge content and quality of the research itself. This highlights a different economic margin: the value of analyst coverage depends not only on whether information is produced, but also on how informative that research is and how difficult the underlying knowledge is for other intermediaries to reproduce.

2 Institutional Background: Sell-Side Research

Sell-side research is a core output of brokerage houses and is produced to inform institutional investors' portfolio allocation, security selection, and trading decisions. Within equity research, this output takes two broad forms. The first is firm-specific research, which focuses on individual companies and often includes earnings forecasts, stock recommendations, target prices, and valuation updates tied to firm-specific news. The second is industry research, which takes a broader view and analyzes the common forces affecting a set of firms within the same sector or product mar-

ket. Industry reports typically discuss demand conditions, pricing pressure, technological change, regulation, input costs, capacity expansion, supply-chain developments, and shifts in competitive positioning. While both forms of research are important in practice, they differ in the type of information they produce and in the dimension of analyst skill they reveal.

Firm-specific research is naturally suited to studying analysts' bottom-line outputs, because it is closely linked to observable forecasts and recommendations that can be readily evaluated *ex post* (e.g., O'Brien 1988, 1990; Clement 1999; Womack 1996; Brav and Lehavy 2003). By contrast, industry research is where analysts articulate their broader economic understanding of the sectors they cover. These reports require analysts to synthesize dispersed information about industry structure, strategic interaction, customer demand, production technologies, and the transmission of macroeconomic or regulatory shocks across firms. In that sense, industry reports often reveal not merely what analysts predict, but how they think. They provide a window into the mechanisms analysts use to interpret industry developments and to map those developments into heterogeneous implications for the firms they cover.

This distinction is institutionally important because investors do not evaluate firms in isolation. The prospects of any given firm depend in part on broader industry conditions, including the evolution of competition, technological adoption, cost shocks, and changes in consumer or enterprise demand. An important function of the sell side, therefore, is to translate sector-level developments into firm-level investment implications. This role helps explain why buy-side market participants place substantial value on analysts' industry knowledge and why brokerage houses devote considerable resources to the production of sector outlooks, thematic reports, and other forms of industry research (e.g., Bradshaw 2011; Brown et al. 2015). In our setting, industry research is far from a niche product: we observe 1.2 million industry reports produced by 40 major brokerage houses over 2012 to 2025, indicating that it represents a large and economically meaningful component of sell-side activity.²

2. The scale of industry research in our sample is comparable to that of firm-specific sell-side research. Using data from LSEG Workspace, Bastianello, Decaire, and Guenzel (2025) report 2.1 million company-specific reports produced by the 43 most common brokerage houses over 2000 to 2025, after excluding industry and macro reports, M&A-related reports, and reports longer than 30 pages. By comparison, our initial sample contains 1.3 million industry reports produced by a similar list of brokerage houses over 2012 to 2025, suggesting that industry research is large in scale not only in absolute terms but also relative to firm-specific research, especially given the much shorter sample period.

Despite its practical importance, the academic literature has overwhelmingly focused on the firm-specific side of analyst output. A large body of work evaluates analysts using forecast accuracy, recommendation profitability, target price informativeness, or related career outcomes. A more recent textual literature studies the content of analyst reports, but it too has exclusively centered on firm-level research (e.g., Asquith, Mikhail, and Au 2005; Huang, Zang, and Zheng 2014; Huang et al. 2018; Décaire and Graham 2024). This emphasis, in part, is due to the fact that firm-specific outputs are easier to standardize and more directly linked to identifiable market events. At the same time, it leaves unexplored the setting in which analyst industry knowledge is most directly expressed. Industry reports are, therefore, a natural setting for studying the qualitative dimension of analyst ability. They contain rich discussions of sector-wide mechanisms that shape firm performance, thereby providing a direct basis for measuring the depth, breadth, and specificity of analysts’ industry expertise.

3 Methodology

Our goal is to construct measures of analysts’ industry knowledge and report quality directly from the textual content of sell-side industry reports. To do so, we use an LLM evaluation pipeline based on OpenAI’s open-weight `gpt-oss-120b` model (OpenAI et al. 2025) that scores each report against a structured rubric designed to capture both general quality features of research and more specific forms of industry expertise.

3.1 Scoring Dimensions

The LLM is prompted to evaluate each industry report using a predefined 18-dimension scoring rubric, categorized as 10 quality measures and 8 knowledge measures. The quality dimensions are *Originality*, *Data Richness*, *Analytical Depth*, *Actionability*, *Structure*, *Forward-Looking*, *Technical Expertise*, *Quantitative Rigor*, *Risk Awareness*, and *Timeliness*. Motivated by survey evidence by Brown et al. (2015), the industry knowledge dimensions are *Industry Trends*, *Industry Technology*, *Supply Chains*, *Distribution Models*, *Margins*, *Customers*, *Labor*, and *Management Teams*. We provide prompts and a full scoring rubric in Appendix Section A.1.

We choose these dimensions to be thorough as well as consistent with the literature. The 10 quality dimensions capture the findings of a broad literature analyzing various aspects sell-side analyst firm reports. Asquith, Mikhail, and Au (2005) and Huang, Zang, and Zheng (2014) find that the text of analysts’ firm reports provide justifications supporting analysts’ opinions, provide information beyond what is contained in earnings forecasts, recommendation levels, and price targets. They show that analyst reports provide entirely new information or interpretation of previously released information to the market. We capture these findings in the *Originality* and *Analytical Depth*. Gleason, Bruce Johnson, and Li (2013) find that there are substantial improvements in analysts’ price target performance when those analysts use a rigorous valuation technique rather than an heuristic which we capture with *Quantitative Rigor* and *Analytical Depth*. Furthermore, analyst forecast timing matters (Guttman 2010) and herding behaviors exist (Trueman 1994; Clement and Tse 2005) so we include the *Timeliness* dimension.

Several additional quality dimensions are intended to capture features of analyst research that make reports more useful to investors. We include *Data Richness* because the incorporation of richer and less conventional information sets, including alternative data, is associated with more accurate analyst forecasts and stronger market reactions to analyst output (Chi, Hwang, and Zheng 2025). We include *Technical Expertise* because analysts with prior industry work experience produce more accurate forecasts, are more likely to become Institutional Investor All-Stars, and elicit stronger market reactions, consistent with domain-specific expertise being a meaningful component of research quality (Bradley, Gokkaya, and Liu 2017). We include *Structure* because the readability, clarity and conciseness of analyst reports affect how effectively their information can be processed by investors and other market participants (Huang, Zang, and Zheng 2014; De Franco et al. 2015). We include *Risk Awareness* because high-quality financial analysis should identify thesis-specific risks, uncertainties, and downside scenarios rather than focus solely on the baseline case. Bastianello, Decaire, and Guenzel (2025) finds that in more recent years, analyst firm reports more frequently include scenario analysis both in discussion and modeling indicating the importance of this dimension for consumers of those reports. We extend this focus on risk to the industry reports. Finally, we include *Forward-Looking* and *Actionability* because one of the central functions of sell-side research is not merely to summarize current conditions, but to form a view about future

industry developments and valuation implications that lead to investable conclusions.

We derive the 8 knowledge dimensions from Brown et al. (2015) who conduct surveys finding that industry knowledge is the single most useful input to both earnings forecasts and stock recommendations. In these surveys, an analyst has industry knowledge if she has an understanding of *Industry Trends*, *Industry Technology*, *Supply Chains*, *Distribution Models*, *Margins*, *Customers*, *Labor*, and *Management Teams*. We use these survey-based dimensions directly, allowing us to measure the extent to which the text of an industry report reflects the forms of industry expertise that market participants themselves describe as most valuable.

3.2 Scoring Rubric and Procedures

For each dimension, the LLM assigns a score on a 5-point scale. A score of 5 corresponds to exceptional analysis, reflecting proprietary or especially deep industry insight and mechanism-based reasoning. A score of 4 reflects strong and differentiated analysis that is well supported by the underlying discussion. A score of 3 corresponds to solid and competent research that provides some value added beyond a basic summary. A score of 2 reflects only minimal interpretation, while a score of 1 indicates weak analysis that largely rehashes existing information without generating meaningful insight.

The scoring rubric also guides the model’s assessment of source quality and evidentiary support. Reports receive stronger evaluations in dimensions such as *Data Richness*, *Quantitative Rigor*, and *Originality* when they draw on more specialized or proprietary sources, such as channel checks, surveys, expert networks, or industry bodies, rather than relying exclusively on company filings, general news, or press releases. Similarly, reports are rewarded when they provide explicit calculations, testable claims, thesis-specific risk analysis, and a multi-dimensional analytical framework. By contrast, reports that consist largely of tables without interpretation, boilerplate discussion of risks, recycled standing views, or high page counts with limited informational density are scored less favorably.

The LLM is prompted to assign scores for each dimension based only on the content of the report and to ground its evaluation in report-specific evidence. This evidence-based design of the rubric is important because it disciplines the model to evaluate what is actually present in the text

rather than relying on stylistic cues or generic associations with the brokerage house, sector, or analyst name. In the case of chunked reports, the model first produces chunk-level evaluations, after which we apply a synthesis step to form a single report-level assessment. This synthesis stage aggregates the evidence across chunks and is intended to preserve information from the full document while avoiding over-weighting any single section. As a result, each report is ultimately mapped into one set of dimension-level scores regardless of whether it was evaluated in one pass or in multiple chunks.

Analyst reports vary substantially in length and we have to accommodate short to very long reports. For reports whose extracted text fits comfortably within the model context window, we submit the full document for evaluation in a single pass. For longer reports, however, we implement a chunking procedure. Specifically, when a report exceeds 50,000 tokens, we split the document into smaller sequential chunks and evaluate each chunk separately. This procedure ensures that the model can assess the full informational content of long reports without truncating important sections. All evaluations are conducted under deterministic model settings, with temperature set to zero and the random seed fixed at 42. These settings reduce stochastic variation in LLM outputs and ensure that a given report receives the same evaluation across repeated runs. This is particularly important in our setting because the measures are later aggregated across reports and linked to analyst-level outcomes in panel regressions. A deterministic pipeline helps ensure that the resulting variation reflects differences in report content rather than random variation in LLM responses.

The output of this process is a set of report-level scores on each rubric dimension. These report-level evaluations serve as the input for our LLM-based analyst measures. Because analysts often publish multiple reports within a month and individual reports may emphasize different dimensions depending on the topic and timing, we subsequently aggregate these report-level scores to the analyst-period level, as described in Section 4.3. This aggregation step allows us to capture persistent differences in analysts’ underlying industry knowledge and research quality while still permitting the measures to vary over time.

This methodology offers several advantages for our setting. First, it makes it feasible to evaluate a very large corpus of sell-side research reports in a consistent manner. Second, by relying on a common rubric applied uniformly across brokerage houses, sectors, and time, it reduces concerns

that analyst ability is being measured using ad hoc or institution-specific conventions. The resulting measures provide a way to quantify otherwise difficult to measure dimensions of analyst ability.

4 Data

4.1 Data Sources

Our dataset is constructed from multiple sources, including LSEG Workspace (formerly Refinitiv Eikon), Institutional Brokers’ Estimate System (I/B/E/S), *Institutional Investor (II)*, Center for Research in Security Prices (CRSP), and Compustat.

Sell-side equity research analyst reports are obtained from LSEG Workspace. We focus on English-language industry reports spanning January 2012 to October 2025 and limit our sample to the 40 major equity research firms, similar to Décaire and Graham (2024).³ This results in a total dataset of 1.3 million reports. We then exclude reports attributed to research teams or divisional entities rather than individual analysts, resulting in a final sample of 1.2 million reports. Figure 1 Panel A shows the annual number of industry reports produced by 40 brokerage houses over the 2012–2025 period. Overall report production remained relatively stable between 2012 and 2023, before experiencing a sharp increase in 2024. Table IA2 further presents the distribution of reports across brokerage houses.

We also obtain metadata for the reports directly from LSEG Workspace. Each report is identified by a unique ID and includes metadata such as the title, contributor (brokerage house), analyst name, and page count, as well as tags assigned by LSEG Workspace, including industry and region classifications. We map the LSEG industry classifications to GICS sectors.⁴ To examine how research coverage has evolved across sectors, Figure 1 Panel B plots the monthly total pages index by GICS sector. For each sector, we aggregate total report pages by month, apply a six-month

3. While Décaire and Graham (2024) identifies 42 brokerage houses, one of them, Bear Stearns, has no industry reports during our sample period. We further treat SunTrust Robinson and Truist Securities as a single entity because Truist Securities is the successor to SunTrust Robinson, following the SunTrust–BB&T merger. This results in 40 distinct brokerage houses in our sample.

4. Each report is assigned a GICS sector by mapping the TRBC economic sector of its Refinitiv-assigned industry tag to the closest GICS counterpart. The two systems are broadly comparable at the top level, as both are market-based classification schemes with similar sector definitions.

moving average, and normalize the series by the sector’s own 2012–2019 average, so that a value of two indicates coverage intensity twice its historical benchmark. Panel B reveals substantial heterogeneity across sectors: Information Technology stands out most dramatically, reaching roughly 2.5 times its benchmark level by 2025, far exceeding the modest increases seen in most other sectors.⁵

Analyst recommendations and earnings per share (EPS) forecasts are collected from the Institutional Brokers’ Estimate System (I/B/E/S). Consistent with prior literature (e.g. Bradley, Gokkaya, and Liu 2017; Harford et al. 2019; Hirshleifer et al. 2019), we focus on US common stocks and one-year-ahead EPS forecasts, yielding 1.87 million forecasts over the 2012–2025 period. Restricting to brokerages present in our report sample reduces this to 0.97 million forecasts issued by 3,794 unique analysts.⁶ Panel A of Table IA3 summarizes the sample construction steps, and Panel B compares key characteristics of brokerages within our sample and forecasts made by analysts at these brokerages to those outside it. On average, brokerages in our sample are significantly larger and cover more firms than those excluded, and EPS forecasts issued by their analysts exhibit significantly lower forecast errors.

For this sample, we further match I/B/E/S records to analyst reports using analysts’ last name, first initial, and brokerage affiliation, verifying match accuracy by cross-checking stock coverage between Refinitiv Eikon and I/B/E/S. Of these 3,794 analysts, 3,112 (over 80%) authored at least one industry report during the sample period, and their forecasts account for 0.90 million observations (more than 90% of the selected sample).⁷ Panel C of Table IA3 presents key characteristics of analysts who have authored industry reports alongside those who have not. Analysts who have authored at least one industry report tend to cover more firms and larger firms, while forecast errors are not significantly different between the two groups.

5. Information Technology consistently ranks among the top sectors in coverage growth across alternative moving average windows and benchmark periods.

6. We thank Law (2023) for generously sharing the I/B/E/S brokerage matching data.

7. We thank Wang, Li, and Yu (2025) for generously sharing the I/B/E/S analyst code (ANALYS) to full name link table for the period 2020–2023.

4.2 Industry Reports

Industry reports differ fundamentally from company-specific reports in their scope and focus.⁸ While company-specific reports concentrate on the financials and prospects of an individual firm, industry reports take a broader view: they summarize the equity analyst’s assessment across all companies covered within a sector, and more importantly, identify the overarching themes expected to shape the industry. This industry-wide perspective encompasses macroeconomic conditions, demand drivers, and competitive dynamics that cut across individual firms.

Figure 2 presents a representative example: a Software industry report produced by Morgan Stanley on January 11, 2012 (first two pages). The report opens with a broad macro assessment of the software sector, noting weak near-term conditions in the first half of 2012 while maintaining a constructive longer-term outlook. It identifies several secular demand drivers expected to shape the industry, including cloud migration, social computing, big data and analytics, security, and mobile applications.

In addition, as shown on the first page, the analyst assigns an explicit industry-level view, in this case *In-Line*, reflecting their overall stance on the sector relative to the broader market. Such industry-level views are a direct manifestation of analysts’ industry expertise. Kadan et al. (2012) were among the first to systematically study this dimension of sell-side research, documenting that analysts possess meaningful across-industry expertise as reflected in their industry recommendations, and that this industry-level information is distinct from and incremental to the information contained in firm-level recommendations. Using I/B/E/S to identify industry recommendations is subject to several potential limitations. First, I/B/E/S only records industry recommendations for brokers that include such recommendations in their firm reports, meaning that coverage is incomplete and the timing may not accurately reflect when the industry view was originally formed or expressed. Second, I/B/E/S does not directly identify the industry to which the recommendation applies; instead, the industry classification can only be inferred from the firms covered by the analyst, which introduces potential misclassification.

8. According to LSEG, brokerage houses select the report type at the time of publication on LSEG Workspace (e.g., Company, Industry, Markets, Commodity, Digest, Economics, ESG, Fixed Income, Forex, Investment Strategy, or Quant/Technical). Tagging for regions, industries, and report types is based solely on the brokerage house’s discretion.

Motivated by their evidence, and leveraging the rich textual content of industry reports, we apply large language models (LLMs) to directly extract industry views or recommendations from each industry report where such information is available, yielding a novel and more comprehensive sample of analyst industry-level recommendations. Comparing our extracted recommendations with those available in I/B/E/S, we document the following: i) Following Kadan et al. (2012), we identify 10,550 industry recommendation records from I/B/E/S Detail Recommendations for the brokerages in our sample over the 2012–2025 period. By contrast, applying LLMs to our sample of 1.2 million industry reports, we extract 44,687 industry recommendations.⁹ ii) The date on which the industry recommendation was originally issued does not necessarily coincide with the date recorded in I/B/E/S Detail Recommendations, a limitation also acknowledged by Kadan et al. (2012). In general, the true issuance date is earlier than or equal to the firm recommendation date. Figure 3 provides an illustrative example of this discrepancy. Analyst Scott Devitt (analyst code: 80523) issued a firm recommendation for WebMD Health Corp on January 11, 2012, which carries an industry recommendation of *Attractive* as recorded in I/B/E/S. However, tracing back to the original industry report, the true issuance date of this industry recommendation was January 3, 2012. iii) Inferring the industry classification from a firm’s GICS code can lead to misclassification. For instance, on January 11, 2012, another analyst issued a firm recommendation for Responsys Inc., which carries an industry recommendation of *In-Line* as recorded in I/B/E/S. Both WebMD Health and Responsys share the same GICS industry code (451010), so an inference based on GICS would assign both industry recommendations to the same industry. However, tracing back to the original industry reports, the *Attractive* recommendation applies to the Internet sector, while the *In-Line* recommendation applies to the Software sector.

Furthermore, some industry reports in our sample disclose the job title of each listed member, such as equity analyst, equity strategist, or research associate. We apply LLMs to extract this information from each report, allowing us to identify the composition of team members behind each report. This is particularly relevant in light of the distinction drawn by Kadan et al. (2012) between firm-level analysts and strategists in the issuance of industry recommendations, as the detailed disclosures in our reports allow us to directly validate which type of professional is responsible for

9. Section 5.5 presents a detailed analysis of the extracted industry recommendations.

each recommendation. Examining the subset of reports that contain an industry recommendation, we find that fewer than 5% list a strategist as a team member. Moreover, many reports explicitly state in their disclosure section that any industry view expressed reflects the analyst’s own view. Based on these findings, we attribute the industry recommendations expressed in industry reports to the individual analysts authoring them, rather than to a centralized view issued by the strategists.

4.3 LLM-Based Industry Expertise Measures

Given that brokerage houses differ in reporting conventions, such as report length and publication frequency, individual reports may only cover a subset of dimensions, and an analyst’s proficiency in each dimension is unlikely to change substantially within a short period, we aggregate to the analyst-month level by taking the maximum report-level score within each dimension for a given analyst-month. This ensures that our analyst-level measures capture consistency within a period while remaining flexible enough to vary over time. For each analyst-month, we then aggregate the dimension-level scores by taking the average across all knowledge and quality dimensions to construct our analyst-month knowledge and quality measures. Panel A of Table 1 presents the summary statistics for the analyst-month *Knowledge* and *Quality* measures and each individual dimension, as well as the total number of reports per analyst-month and average report length. The mean *Knowledge* score is 2.61 with a standard deviation of 0.74, while the mean *Quality* score is 3.37 with a standard deviation of 0.57, suggesting that analysts tend to score higher on report quality dimensions than on industry knowledge dimensions. On average, analysts publish 4 reports per month, with an average report length of 20 pages, reflecting considerable variation in publication frequency and report length across analysts.¹⁰

As the appropriate level of aggregation varies across empirical settings, we also construct LLM-based measures at the analyst-year, event-date, and firm levels, depending on the specification. The construction of these alternative measures follows the same procedure described above and is

10. The distribution of *No. of reports* is highly skewed. We manually examined several high-count observations and find that the high P99 value does not reflect a data error. Instead, it is driven by a small subset of analysts, particularly those covering sectors with frequent data releases, who issue recurring short-form monitor and update notes. This pattern is also concentrated among a few brokerage houses, reflecting their distinct publication practices. This again motivates our choice to aggregate *Knowledge* and *Quality* measures at the analyst-month level rather than the individual-report level.

detailed in the context of each respective analysis. Across all specifications, the resulting measures are standardized to have a zero mean and unit variance to facilitate the interpretation of the estimated coefficients.

4.4 Key Dependent Variables

Our first set of dependent variables comes from the *II* All-Star Ratings. Since 1972, *Institutional Investor (II)* has conducted annual surveys in which buy-side institutions rate sell-side analysts who "have been most helpful to them in researching U.S. equities over the past twelve months" (*Institutional Investor* 1996, 1997). Based on these surveys, a list of the top-three analysts and runners-up by industry is published in the magazine's October issue. These *II* All-Star ratings carry significant weight in the sell-side industry, as they are closely tied to analyst compensation and career outcomes (Groysberg, Healy, and Maber 2011). Moreover, Bradshaw (2011) mentions that the most important trait valued by institutional investors is industry knowledge. A growing body of research has since adopted All-Star recognition as a prominent measure of analyst career outcomes to study the determinants of analyst success (Bradley, Gokkaya, and Liu 2017; Harford et al. 2019; C. Li et al. 2020; Huang, Lin, and Zang 2022). We therefore use All-Star recognition as an external benchmark to validate whether our LLM-based knowledge and quality measures capture meaningful variation in analyst ability.¹¹

Specifically, we construct an *AllStar* indicator variable equal to one if an analyst is named among the top three or as a runner-up in their industry category in a given year, and zero otherwise.

Next, we test *Earnings Forecast Error*. Brown et al. (2015), in a survey-based study, documents that industry knowledge is the single most useful input to analysts' earnings forecasts and stock recommendations. Motivated by this evidence, we examine whether our LLM-based knowledge and quality measures are associated with analyst forecast accuracy.

Following prior literature, we measure forecast accuracy using *Forecast error*. For analyst i forecasting earnings per share (EPS) for company c issued at time t , $Forecast\ error_{i,c,t}$ is computed

11. We thank An-Ping Lin for generously sharing the All-Star data through 2022. Over the 2012–2022 period, there are 3,286 annual All-Star designations, of which 2,681 are affiliated with brokerage houses in our sample, corresponding to 569 unique analysts. Of these, 545 have authored at least one industry report in our sample, suggesting that our data captures the vast majority of prominent sell-side analysts recognized as All-Stars.

as the absolute difference between the actual EPS and the forecasted EPS, scaled by the stock price one year prior to the annual earnings announcement at time T , and multiplied by 100:

$$Forecast\ error_{i,c,t} = 100 \times \frac{|Actual\ EPS_{c,T} - Forecasted\ EPS_{i,c,t}|}{Price_{c,T-1}} \quad (1)$$

We next investigate whether investors recognize analysts' industry expertise and perceive their research as more informative using *Market Reaction*. If so, we would expect investors to react more strongly to earnings forecast revisions and stock recommendation changes issued by analysts with higher LLM-based knowledge and quality scores. The key dependent variable in this set of tests is the cumulative market-model-adjusted abnormal stock returns (in percentage) over a three-day event window $[-1, +1]$ around a forecast revision or recommendation change.

Additionally, we test whether analysts' industry expertise is associated with a more favorable firm-level information environment. Our first set of dependent variables captures liquidity. Specifically, we use *Amihud*, *Effective Spread*, and *Realized Spread*. *Amihud* is the average daily Amihud illiquidity (Amihud 2002) in a given year, with higher values indicating the stock is less liquid. *Effective Spread* measures the realized round-trip trading cost faced by investors relative to the prevailing midpoint quote, while *Realized Spread* captures the cost after the trade executes. Higher values of both spread measures indicate greater trading frictions and a less favorable information environment. Accordingly, lower values of *Amihud*, *Effective Spread*, and *Realized Spread* reflect lower information asymmetry and better market liquidity.

Our second set of dependent variables captures price informativeness. We use *Buy-Sell Ratio* and *Lambda*. *Buy-Sell Ratio* is the absolute value of the buy-sell imbalance measure constructed from trading volume, with higher values indicating that trading is more one-sided and therefore more likely to reflect asymmetric information or order-flow imbalances rather than balanced incorporation of information into prices. *Lambda* is a price impact measure that captures the sensitivity of prices to order flow. Higher values imply that trades move prices more strongly, consistent with a less liquid and less informative market. For both variables, lower values indicate a more favorable information environment in which firm-specific information is incorporated into prices with fewer frictions. We use *Variance Ratio* which is the absolute value of the difference between the ratio

of 60-to-300 second stock return variance and one (`var_ratio3` from WRDS Millisecond Intraday Indicators). We use this measure to assess the price efficiency of the firm’s stock. If analysts’ knowledge and quality of reports for a firm’s industry enhance the price efficiency of the firm’s stock, we should see a reduction in the variance ratio.

Next, we test if the implied cost of capital changes with analyst expertise. If analysts with greater industry knowledge and higher quality research improve the firm’s information environment, then these benefits should be reflected not only in trading-based outcomes, but also in a lower implied cost of capital.

Because the cost of capital is not directly observable, we proxy for it using implied cost of capital (*ICC*). We follow Eskildsen et al. (2026), who in turn follow Mohanram and Gode (2013), and measure *ICC* as the equal-weighted average of four standard estimates from the accounting literature.¹² Lower values of *ICC* indicate that the firm enjoys a lower implied cost of equity financing.

4.5 Other Variables

We include a set of control variables capturing analyst-, firm-, and event-specific characteristics following prior literature (e.g. Harford et al. 2019; Gibbons, Iliev, and Kalodimos 2021). *Analyst career length* is the total number of years the analyst has appeared in I/B/E/S. *Time covering company* is the total number of years since the analyst first issued an earnings forecast for the firm. *No. of industries* is the number of six-digit GICS industries covered by the analyst.¹³ *No. of firms* is the number of unique firms followed by the analyst. *Average firm size* is the average size of all firms covered by the analyst. *Average forecast error* is the average forecast error across all forecasts issued by the analyst in the year. *Brokerage size* is the total number of analysts employed at the

12. We use the data provided by Eskildsen et al. (2026) for the four ICC measures. The four estimates of implied cost of capital are based on commonly used methods found in Ohlson and Juettner-Nauroth (2005), Easton (2004), Claus and Thomas (2001), and Gebhardt, Lee, and Swaminathan (2001).

13. The use of six-digit GICS industries follows Kadan et al. (2012). GICS is widely adopted by investment banks as an industry classification system and has been shown to better explain stock comovements compared to other popular industry classifications (Bhojraj, Lee, and Oler 2003). Moreover, Boni and Womack (2006) show that the GICS classification is a good proxy for how sell-side analysts specialize by industry. Our results remain unchanged when we instead use two-digit SIC industries to define industry coverage, following Bradley, Gokkaya, and Liu (2017) and Harford et al. (2019).

brokerage house. *Market value* is the firm’s year-end market capitalization, measured as share price multiplied by total shares outstanding. *Institutional ownership* is the fraction of a firm’s equity held by institutional investors at year-end. *No. of analysts* is the number of unique analysts issuing earnings forecasts for the firm. *Forecast revision* is the difference between the analyst’s revised forecast and the prior forecast, scaled by the one-year lagged stock price and multiplied by 100. *Forecast horizon* is the number of days between the forecast issuance date and the associated earnings announcement date. We further control for firm-level stock return characteristics measured prior to the forecast revision or recommendation change. *BM* is the ratio of book value of equity to market value of equity, measured at the fiscal year-end preceding the event. *Lag monthly return* is the cumulative firm’s percentage stock return over the month before the event. *Volatility* is the standard deviation of daily percentage stock returns over the two months preceding the event. *Momentum* is the cumulative stock return from six months to one month before the event, measured as a decimal. Panel B of Table 1 presents the descriptive statistics of all key variables used in the empirical analysis.

In the liquidity and price-informativeness specifications, we control for a broader set of firm and market characteristics following the literature. *Log analyst coverage* is the logarithm of the number of analysts covering the firm. *Size* is firm size. *ROA* is return on assets. *Leverage* is the firm’s leverage ratio. *Market-to-book* is the ratio of market value to book value. *Past return* captures prior stock performance, depending on the specification. *Log price* is the logarithm of stock price. *Log volume* is the logarithm of trading volume. We also include the lagged dependent variable to capture persistence in the information environment. In the implied cost of capital tests, we also include *MAE*, *Forecast dispersion*, *Beta*, *Earnings variability*, and *LT earnings growth*.

5 Empirical Results

5.1 What Drives Analyst Industry Knowledge?

In this section, we investigate what drives analyst industry knowledge and report quality. We construct analyst-year LLM-based knowledge and quality measures by averaging the corresponding

monthly scores across all months in a given year and standardizing them to have zero mean and unit variance. We further restrict the sample to analysts who have issued at least one report tagged with a US country or region in Refinitiv Eikon in that year, ensuring our analysis focuses on analysts covering US markets. Then we regress these measures on three sets of variables. The first is the lagged dependent variable, included to test the persistence of our measures over time. The second set captures report and brokerage characteristics, including the number of reports issued (*No. of reports*), average report length (*Average report length*), and the number of analysts at the brokerage who author industry reports (*Brokerage size*).¹⁴ The third set captures analyst characteristics from I/B/E/S, including *Analyst career length*, *No. of industries*, *Average firm size*, and *Industry Experience*. *Industry Experience* is an indicator equal to one if the analyst has prior work experience in an industry overlapping with the firms they cover in a given year, in the spirit of Bradley, Gokkaya, and Liu (2017). To construct this variable, we manually collect LinkedIn profile information for analysts in our sample and assign six-digit GICS industry codes to their prior employers.¹⁵ With the exception of the lagged dependent variable, all remaining controls are measured contemporaneously. In each specification, we additionally include brokerage, industry, and year fixed effects.

Table 2 presents the results.¹⁶ For both *Knowledge* and *Quality*, the lagged dependent variable enters positively and highly significantly across all specifications. Before adding analyst fixed effects, the coefficients range from 0.42 to 0.56, suggesting that analyst industry knowledge and report quality are highly persistent over time. Once analyst fixed effects are included in Columns (4) and (8), the coefficients are still positive and highly significant, at approximately 0.11. Turning to report characteristics, both *No. of reports* and *Average report length* are positively and significantly associated with both measures across Columns (1)–(3) and (5)–(7), indicating that analysts who

14. In regression analyses, *No. of reports* and *Average report length* are each divided by 10 to facilitate interpretation of the estimated coefficients.

15. We are able to identify LinkedIn profiles for 2,932 of the 3,112 analysts in our sample. Consistent with the industry classification used throughout the paper, we define industry overlap at the six-digit GICS level. We require at least two years of prior work experience in the overlapping industry. For prior employers, we assign six-digit GICS codes using CRSP/Compustat matches when available and LLM-assisted classification otherwise, based on employer names and business descriptions.

16. The difference in sample size across specifications is due to two factors. First, not all analysts who author industry reports necessarily issue EPS forecasts or stock recommendations. Second, I/B/E/S coverage is incomplete; notably, one of the largest brokerages in our sample (UBS Equities) was removed from the I/B/E/S Detail History database.

produce more reports and longer reports tend to exhibit higher knowledge and quality scores. *Brokerage size* is positively associated with both measures when analyst characteristics are excluded (Columns (3) and (7)), but becomes insignificant once they are controlled for (Columns (4) and (8)).

Among analyst characteristics from I/B/E/S, *No. of industries* is positively and significantly associated with *Knowledge* in Column (2), consistent with the notion that broader industry exposure enhances industry knowledge. *Average firm size* is positively associated with both measures in Columns (2)–(3) and (6)–(7), suggesting that analysts covering more firms and larger firms accumulate greater industry expertise and produce higher quality research. *Analyst career length* is positively associated with both measures when report characteristics are excluded (Columns (2) and (6)), but becomes insignificant once they are controlled for in Columns (3)–(4) and (7)–(8), suggesting that the effect of experience may operate through report production. *Industry Experience* is positively associated with both *Knowledge* and *Quality* across specifications, though the coefficients are not statistically significant.

5.2 What Makes a Star Analyst?

We next examine what makes a star analyst, providing a further external validation of our LLM-based measures. Specifically, we estimate a logit regression to investigate how analyst characteristics affect the probability of receiving an *Institutional Investor* All-Star rating. Our dependent variable is *AllStar*, an indicator equal to one if the analyst is named among the top three or as a runner-up in their industry category in a given year, and zero otherwise. The independent variables include our LLM-based *Knowledge* and *Quality* measures, alongside a set of analyst and brokerage controls: *Average forecast error*, lagged All-Star status, *Analyst career length*, *Brokerage size*, *No. of industries*, *Average firm size*, and *Industry Experience*. The analysis is conducted at the analyst-year level over the period 2012–2025, with all independent variables lagged by one period. Our model is specified as follows:¹⁷

17. All-Star data for 2012–2022 are obtained from An-Ping Lin, providing a complete list of All-Star designations. For 2023–2025, we manually collect All-Star recognition from publicly available sources (e.g., analysts’ LinkedIn profiles and brokerage house websites), covering approximately 40% of All-Star analysts in 2023 and 60% in 2024–2025. For analysts with missing data in 2023–2025, we impute All-Star status

$$\mathbb{P}(AllStar_{i,t} = 1) = \Lambda\left(\beta_1 Expertise_{i,t-1} + \beta_2 Average\ forecast\ error_{i,t-1} + \gamma Control_{i,t-1}\right) \quad (2)$$

where *Expertise* is proxied by either *Knowledge* or *Quality* as defined in Section 4.3. To facilitate comparison, *Knowledge*, *Quality*, and *Average forecast error* are standardized to have zero mean and unit variance. The results are reported in Table 3. All specifications include brokerage, industry, and year fixed effects. Columns (1) and (2) show that both *Knowledge* and *Quality* are individually positively and significantly associated with the probability of receiving an All-Star rating, with coefficients of 0.636 and 0.629, respectively. Column (3) shows that *Average forecast error*, when included on its own, is negatively associated with All-Star status, indicating that forecast accuracy is also related to analyst recognition. Once analyst and brokerage controls are included in columns (4) and (5), *Average forecast error* loses its significance, while *Knowledge* and *Quality* remain positively and significantly associated with All-Star status, with coefficients of 0.509 and 0.530, respectively. These associations remain positive and significant after further controlling for *Industry Experience* and lagged All-Star status in columns (6) and (7), with coefficients of 0.295 and 0.289, respectively.

In terms of magnitude, because *Knowledge*, *Quality*, and *Average forecast error* are standardized, their coefficients are directly comparable within the same logit specification. In the univariate specifications, the coefficients on *Knowledge* and *Quality* are 0.636 and 0.629, respectively, approximately five times larger in absolute value than the coefficient on *Average forecast error* (−0.128). More importantly, after including analyst and brokerage controls, *Average forecast error* becomes statistically insignificant, whereas *Knowledge* and *Quality* remain positive and significant. This pattern suggests that the qualitative dimensions of analyst output captured by our LLM-based measures are more strongly associated with buy-side recognition than traditional forecast accuracy metrics.

based on their 2022 recognition: analysts who received All-Star status in 2022 are assumed to retain it in subsequent years, given the high persistence of All-Star recognition over time.

5.3 Industry Expertise and EPS Forecast Accuracy

Having established that our LLM-based measures serve as a valid proxy for analyst ability, we next examine whether they are associated with analyst forecast accuracy. We expect forecasts issued by analysts with greater industry *Expertise* to be more accurate. To test this prediction, we estimate the following regression model:

$$\text{Forecast error}_{i,c,t} = \beta_1 \text{Expertise}_{i,c,t-1} + \gamma \text{Control}_{t-1} + \mu_i + \delta_c + \theta_t + \varepsilon_{i,c,t} \quad (3)$$

Here, the dependent variable *Forecast error* is defined as in Section 4.4. The key independent variables are analyst-event-specific *Expertise*, proxied by *Knowledge* or *Quality* measures, each standardized to have zero mean and unit variance, constructed using reports issued within a one-month window prior to the forecast issuance date. Specifically, for each dimension, we first take the maximum score across all reports within this window, and then average across the relevant dimensions to obtain the *Knowledge* and *Quality* measures, respectively. We also control for a set of analyst-, firm-, and forecast-level characteristics, including *Average forecast error*, $\text{Ln}(\text{Market value})$, *Institutional ownership*, $\text{Ln}(\text{No. of analysts})$, $\text{Ln}(\text{Brokerage size})$, $\text{Ln}(\text{Analyst career length})$, $\text{Ln}(\text{Time covering company})$, and $\text{Ln}(\text{Forecast horizon})$. With the exception of $\text{Ln}(\text{Analyst career length})$, $\text{Ln}(\text{Time covering company})$, and $\text{Ln}(\text{Forecast horizon})$, which are measured based on the forecast date, all other control variables are lagged by one period. We estimate specifications with varying combinations of analyst, firm, and year-month fixed effects.

Table 4 presents the results. Both *Knowledge* and *Quality* are negatively and significantly associated with *Forecast error* across all specifications, indicating that analysts with higher knowledge and quality scores produce more accurate EPS forecasts. These results are robust to the inclusion of analyst-firm fixed effects in columns (2) and (4). As our LLM-based measures are constructed on a bounded scale, we focus on the sign and statistical significance of the coefficients rather than their economic magnitude. By contrast, *Average forecast error*, a proxy for past forecast accuracy, loses its significance across all of our specifications, while our LLM-based measures remain significant throughout, suggesting that *Knowledge* and *Quality* capture important dimensions of analyst ability that are not subsumed by past numerical performance.

5.4 Market Reactions to Analyst Research

Building on our evidence that LLM-based *Knowledge* and *Quality* measures are associated with both All-Star recognition and forecast accuracy, we next investigate whether investors recognize analysts' industry *Expertise* and perceive their research as more informative. The All-Star validation results suggest that buy-side clients directly value analyst knowledge and report quality through their voting behavior. Meanwhile, investors may also learn about analyst ability indirectly over time by observing the accuracy of their forecasts. If investors recognize higher-ability analysts, as proxied by our LLM-based measures, we would expect them to react more strongly to earnings forecast revisions and stock recommendations issued by those analysts.

5.4.1 Earnings Forecast Revisions

We first examine market reactions to forecast revisions, expecting more pronounced reactions to revisions issued by analysts with higher *Knowledge* and *Quality* scores. We estimate the following regression model to test our prediction:

$$\begin{aligned} CAR_{i,c,t} = & \beta_1 Forecast\ revision_{i,c,t-1} + \beta_2 Expertise_{i,t-1} \\ & + \beta_3 (Forecast\ revision \times Expertise_{i,t-1}) \\ & + \gamma Control_{t-1} + \mu_i + \delta_c + \theta_t + \varepsilon_{i,c,t} \end{aligned} \tag{4}$$

Here, the dependent variable *CAR* is the cumulative market-model-adjusted abnormal stock returns over a three-day event window $[-1, +1]$ around a forecast revision. The key independent variables are analyst-event-specific *Expertise*, proxied by *Knowledge* or *Quality* measures, following the same procedure as described in Section 5.3. The key variables of interest are the interaction terms *Forecast revision* $\times$ *Knowledge*(*Quality*), which capture whether the market reaction to forecast revisions is more pronounced for analysts with higher knowledge and quality scores. Control variables include a set of analyst- and firm-level characteristics, as well as stock return characteristics measured in the month prior to the forecast revision: *BM*, *Lag monthly return*, *Volatility*, and *Momentum*. All control variables are either lagged by one period or measured using information available prior to the forecast revision date. All specifications include analyst, firm, and year-month

fixed effects, and standard errors are clustered at the firm level.

Table 5 presents the results. The coefficient on *Forecast revision* is positive and highly significant in both specifications, indicating that the stock market responds positively to upward revisions and negatively to downward revisions, with larger revisions eliciting greater price reactions. More relevant for our purposes are the interaction terms between *Forecast revision* and the *Knowledge* and *Quality* measures. We find that the interaction coefficients are positive and significant in both specifications, indicating that conditional on the direction and magnitude of a forecast revision, the stock market reacts more strongly to revisions issued by analysts with higher knowledge and quality scores. This further suggests that investors recognize and incorporate analyst ability into their trading responses, treating research from higher-ability analysts as more informative.¹⁸

5.4.2 Stock Recommendation Changes

We next turn to market reactions to stock recommendation changes. Prior work shows that analysts with superior forecast accuracy also issue more informative stock recommendations (Loh and Mian 2006), and that accurate forecasts are primarily used as inputs for corresponding stock recommendations (Brown et al. 2015). Motivated by this evidence, we expect more pronounced market reactions to recommendation changes issued by analysts with higher *Knowledge* and *Quality* scores. We estimate the following regression model:

$$CAR_{i,c,t} = \beta_1 Expertise_{i,t-1} + \gamma Control_{t-1} + \mu_i + \delta_c + \theta_t + \varepsilon_{i,c,t} \quad (5)$$

Here, the key explanatory variables are analyst-event-specific analyst-event-specific *Expertise*, proxied by *Knowledge* or *Quality* measures, constructed following the same procedure as described in Section 5.3. All other variables and fixed effects are defined and constructed as in Section 5.4.1. Given the asymmetric market reactions to recommendation upgrades and downgrades documented in the literature (e.g., Liu, Smith, and Syed 1990; Kecskés, Michaely, and Womack 2017), we estimate separate regressions for each.

18. Our results are robust to the inclusion of analyst-firm fixed effects and to an alternative definition of *Forecast revision* following Ivković and Jegadeesh (2004), measured as the difference between the analyst’s revised forecast and the prior forecast scaled by the absolute value of the prior forecast.

Table 6 presents the results. Columns (1) and (2) report results for recommendation downgrades, and Columns (3) and (4) for upgrades. Consistent with our prediction, the coefficients on *Knowledge* and *Quality* are negative and significant for downgrades and positive and significant for upgrades. These results indicate that the market reacts more strongly to recommendation changes issued by analysts with higher knowledge and quality scores. The informativeness of stock recommendations, as perceived by the market, thus appears to be increasing in analysts’ knowledge and quality scores.

Comparing the magnitudes of the coefficients across downgrades and upgrades, the effects are notably larger for downgrades than for upgrades, with coefficients of -0.346 and -0.349 for the former and 0.158 and 0.200 for the latter. This asymmetry is consistent with prior literature suggesting that investors consider downgrades more credible and informative than upgrades, as the latter may be driven by analysts’ conflicts of interest, namely their incentives to maintain relationships with firm management and generate order flow.

Taken together, the above two sets of analyses suggest that investors recognize and incorporate analyst ability into their response to analyst research. Specifically, we find that earnings forecast revisions and stock recommendation changes issued by analysts with higher *Knowledge* and *Quality* scores elicit stronger market reactions, indicative of research produced by higher-ability analysts conveying greater information content.

Finally, we address the potential look-ahead bias arising from the possibility that the LLM was trained on data post-dating our sample period. To mitigate this concern, we re-estimate the two main sets of tests above restricting the sample to the post-tainting period (July 2024 onward). Table 7 reports the results. The findings remain qualitatively unchanged: *Knowledge* and *Quality* continue to predict lower forecast errors and stronger market reactions to forecast revisions. Taken together, these results alleviate concerns that our findings are driven by look-ahead bias.¹⁹

19. We do not conduct this test for stock recommendation changes, as the post-tainting period yields an insufficient number of observations for reliable estimation.

5.5 Industry Expertise and Industry Recommendations

The analyses presented thus far have focused primarily on analyst outputs at the company level, examining how analyst industry *Expertise* relate to firm-specific forecast accuracy, stock recommendations, and information environment outcomes. However, a defining feature of analyst industry expertise is its potential to generate insights that extend beyond individual firms to the broader industry landscape. This industry-wide perspective is directly embodied in industry reports, which take a broader view by summarizing an analyst’s assessment across all companies covered within a sector. Unlike company-specific reports, they identify the overarching themes expected to shape the industry, encompassing macroeconomic conditions, demand drivers, and competitive dynamics that cut across individual firms. Many such reports also include an explicit industry-level view, reflecting the analyst’s overall stance on the sector relative to the broader market. Kadan et al. (2012) were among the first to systematically study this dimension of sell-side research, documenting that analysts possess meaningful across-industry expertise as reflected in their industry recommendations and that this information is orthogonal to, and thus incremental to, the information contained in firm-level recommendations. Motivated by their evidence, we ask whether industry recommendations issued by analysts with greater industry Expertise are more informative.

Applying LLMs to our sample of 1.2 million industry reports, we extract 44,687 industry recommendation records spanning 73 six-digit GICS industries over the 2012–2025 period, with each record classified following the standard GICS classification hierarchy. Panel A of Table 8 presents the distribution of these LLM-extracted recommendations across brokerage houses. The records are concentrated among a small number of large brokerages, with Morgan Stanley and Barclays accounting for the majority. Optimistic recommendations account for over 30% of the total, neutral recommendations represent the largest share at more than half, and pessimistic recommendations remain the least common category at below 10%. Table IA4 further presents the distribution of our LLM-extracted industry recommendations across industries, reporting the top 20 industries ranked by total recommendation count, which together account for 77% of all industry recommendations.

For comparison, we replicate the approach of Kadan et al. (2012) to construct industry-level recommendations from I/B/E/S Detail Recommendations over the same period. Among brokerages in our sample, only 10,550 records are captured by I/B/E/S, representing less than a quarter of

the recommendations identified through our LLM-based approach.

Having documented the distribution of industry recommendations, we next turn to examining their informativeness. Changes in recommendations represent an analyst’s deliberate decision to revise her industry view and are thus more likely to convey new information than reaffirmations of prior positions. We therefore focus on revisions in industry recommendations, consistent with the firm recommendation literature (e.g., Jegadeesh et al. 2004; Barber et al. 2006) and with Kadan et al. (2012). In total, we identify 2,246 recommendation revisions, including 1,149 upgrades and 1,097 downgrades. By comparison, the I/B/E/S-based approach yields only 879 revisions (463 upgrades and 416 downgrades) among brokerages in our sample. The relatively small number of revisions reflects the persistent nature of industry recommendations, with analysts frequently maintaining their sector views across consecutive reports rather than issuing new signals.

We further compare the timing of I/B/E/S-based revisions with those extracted from industry reports. Because analyst attribution and GICS industry assignments in I/B/E/S may be noisy, we conduct a rough event-level match: for each report-based revision, we search for a nearby I/B/E/S revision issued by the same brokerage, allowing matches at either the analyst or brokerage level and using either the six-digit GICS industry or the broader four-digit GICS industry group. We require that matched revisions share the same direction and discard candidate matches occurring more than 365 days apart. Among report-based revisions matched to a subsequent I/B/E/S revision, the I/B/E/S date lags by a median of 85 to 93 days, depending on the matching criterion, with mean delays of approximately 112 to 118 days. This timing difference is consistent with I/B/E/S often capturing industry views when they subsequently appear in firm-level analyst reports, rather than when the underlying industry-level view is first formed or reported in industry research. Taken together, these results indicate that I/B/E/S-based industry recommendations are both incomplete and stale.

For each recommendation revision, we measure industry performance over the 21 trading days beginning at the start of the month following the revision month, corresponding approximately to a one-month holding period, computed using market-model, Fama–French three-factor, and Carhart four-factor adjusted returns, respectively. Upgrades are treated as long signals and downgrades as short signals. For downgrades, we multiply subsequent buy-and-hold abnormal returns by -1 ,

so that larger values of the dependent variable consistently indicate greater profitability in the direction of the analyst’s recommendation revision. We next examine whether industry recommendation revisions issued by more knowledgeable analysts generate stronger subsequent industry-level returns. Specifically, we estimate the following specification:

$$SignedBHAR_{i,ind,t} = \beta_1 Expertise_{i,t-1} + \gamma Control_{t-1} + \mu_i + \delta_{ind} + \theta_t + \varepsilon_{i,ind,t} \quad (6)$$

Here, the key explanatory variable is the analyst-month *Expertise*, proxied by the *Knowledge* measure, constructed following the same procedure as described in Section 4.3, and standardized to have zero mean and unit variance. Control variables include industry-level cumulative return, return volatility, and market capitalization, as well as an upgrade indicator equal to one for recommendation upgrades and zero for downgrades, the magnitude of the recommendation revision (measured as the absolute step size of the revision on the three-point scale), and analyst-level controls including the number of reports and average report length in the revision month. The specification includes analyst, industry, and year-month fixed effects, and standard errors are clustered by industry.

Industry-level variables are constructed at the six-digit GICS industry level. Industry market capitalization is the sum of the market capitalizations of all constituent firms within the industry. Industry daily returns are computed as the market-capitalization-weighted average of constituent firm returns, where weights are based on each firm’s market capitalization at the end of the prior month, excluding firms with a prior-month-end price below \$1 to mitigate the influence of penny stocks. These daily returns are used to construct the buy-and-hold abnormal returns, as well as the control variables: industry cumulative return and return volatility, both measured over the recommendation revision month as the compounded daily return and the standard deviation of daily returns, respectively.

Panel B of Table 8 presents the results. Consistent with our prediction, the coefficient on *Knowledge* is positive and statistically significant across all three benchmark specifications. The coefficient estimates range from 0.336 to 0.379, indicating that recommendation revisions issued by more knowledgeable analysts are associated with greater subsequent industry-level profitability, with the effect being both statistically and economically significant. These results further validate

that our LLM-based measures also capture meaningful dimensions of analyst ability at the industry level.

5.6 Do Firms Benefit from Analysts' Industry Expertise?

Having shown that our LLM-based *Knowledge* and *Quality* measures are positively associated with analyst recognition, forecast accuracy, and market reactions, we next ask whether firms themselves benefit when they are followed by analysts with greater industry knowledge and higher quality research. If analyst ability improves the firm's information environment, then stocks followed by stronger analysts should exhibit lower information asymmetry and a lower cost of capital. Similar phenomena have been shown in analyst coverage (Bowen, Chen, and Cheng 2008) and analyst effort allocation (Harford et al. 2019). These outcomes represent economically meaningful channels through which analyst industry expertise may create value beyond the immediate trading response to individual research events.

We investigate this broader question along two dimensions. First, we examine whether analyst ability is associated with lower information asymmetry as reflected in liquidity measures and price informativeness proxies. Second, we examine whether these improvements in the information environment translate into a lower implied cost of capital. Across these tests, our central prediction is that firms followed by analysts with higher *Knowledge* and *Quality* scores should face a more favorable information environment and, consequently, lower financing frictions.

5.6.1 Information Asymmetry

We begin by examining whether firms followed by more knowledgeable analysts exhibit lower information asymmetry. Analysts with deeper industry knowledge and higher quality research should produce more informative outputs, reducing uncertainty among market participants. If so, stocks followed by such analysts should display better liquidity and greater price informativeness. To test this prediction, we estimate regressions of the following form:

$$Y_{f,t+1} = \beta_1 Expertise_{f,t} + \beta_2 Size_{f,t} + \beta_3 Expertise_{f,t} \times Size_{f,t} + \gamma Control_{f,t} + \mu_f + \delta_t + \varepsilon_{f,t}, \quad (7)$$

where $Y_{f,t+1}$ denotes a next-period measure of the firm’s information environment. In Table 9, the dependent variables are *Amihud*, *Effective Spread*, and *Realized Spread*. In Table 10, the dependent variables are *Buy-Sell Ratio*, *Lambda*, and the 5-minute *Variance Ratio*. The key explanatory variables are firm-year analyst *Knowledge* and *Quality* measures. We construct these firm-year measures using the mean of the top three analyst scores within each firm-year, with the goal of capturing the influence of the firm’s strongest analysts rather than the average quality of all analysts following the firm. We also interact the analyst measure with firm size to allow the relation between analyst expertise and the firm’s information environment to vary systematically across firms.

Table 9 reports the liquidity results. The coefficients on *Knowledge* and *Quality* are negative across all three liquidity outcomes for both knowledge and quality, indicating that firms followed by stronger analysts tend to exhibit lower future trading frictions. The evidence is strongest for analyst *Knowledge*. In particular, knowledge is negatively and significantly associated with *Amihud*, *Effective Spread*, and *Realized Spread*. The corresponding quality coefficients show the same pattern. Overall, these results suggest that analyst industry expertise, especially knowledge, is associated with better future liquidity.

The interaction terms between analyst expertise and firm size are positive and statistically significant across all liquidity specifications. This pattern indicates that the beneficial association between analyst expertise and liquidity is weaker for larger firms. Economically, this is plausible. Large firms already tend to have richer information environments, broader investor attention, and greater trading activity, leaving less scope for analyst expertise to further reduce information asymmetry. The marginal informational contribution of highly knowledgeable analysts therefore appears to be more important for smaller firms.

Table 10 turns to price informativeness and market efficiency. We examine three proxies: *Buy-Sell Ratio*, *Lambda*, and the 5-minute *Variance Ratio*, all measured in the next period. Lower values of *Buy-Sell Ratio* and *Lambda* indicate a more favorable information environment, reflecting more balanced order flow and lower price sensitivity to trading pressure. We discuss the results of the Variance Ratio test in section 5.6.2.

The results again point to a favorable role for analyst expertise. For *Buy-Sell Ratio*, both

Knowledge and *Quality* show negative correlation, with the knowledge coefficient significant at the 1% level and the quality coefficient significant at the 10% level. For *Lambda*, the coefficient on *Knowledge* is negative and statistically significant, while the coefficient on *Quality* is also negative but statistically significant. These results suggest that firms followed by more knowledgeable analysts and, to a lesser extent, higher-quality analysts exhibit more informative prices in subsequent periods.

As in the liquidity regressions, the interaction between analyst expertise and firm size is positive and strongly significant across all price-informativeness measures. This again implies that the beneficial relation between analyst expertise and the firm’s information environment is concentrated among smaller firms, where analyst information production is likely to be more valuable.

5.6.2 Price Efficiency

Next, we test whether firms covered by more knowledgeable analysts have better subsequent price efficiency. For this test, we use the 5-minute Variance Ratio which is the absolute value of the absolute value of the difference between the ratio of 60-to-300 second stock return variance and one. If more knowledgeable analysts improve the firm’s information environment incorporating information into prices more efficiently, then firms covered by such analysts should have lower variance ratios.

The results reported in columns (5) and (6) of Table 10 provide support for this prediction. The coefficient on both *Knowledge* and *Quality* are negative and statistically significant, with *Knowledge* exhibiting a stronger relationship than *Quality*. Similar to the liquidity and price informativeness results, we let the main measures interact with firm size since size is a prominent determinant of price efficiency and find that the interaction between analyst expertise and firm size is positive and statistically significant for both the measures. This pattern provides additional support that the benefits of analyst expertise on price efficiency is weaker for larger firms, likely since information frictions are more severe for smaller firms.

The results in Tables 9 and 10 show that firms followed by analysts with stronger LLM-based expertise scores experience a better subsequent information environment. They exhibit lower trading frictions, more balanced trading, lower price impact, and more efficient price dynamics. The

results are especially strong for the knowledge measure and are most pronounced among smaller firms, consistent with the view that analyst industry expertise improves the firm’s broader information environment.

5.6.3 Cost of Capital

We next ask whether the improvements in firms’ information environments documented above translate into a lower cost of capital. A large literature argues that better information environments reduce required returns by lowering estimation risk, mitigating information asymmetry, and improving the precision with which investors assess firm fundamentals. If analysts with greater industry knowledge and higher quality research reduce these frictions, then firms followed by such analysts should face a lower implied cost of capital. To test this prediction, we estimate regressions in which the dependent variable is the implied cost of capital measure:

$$ICC_{f,t} = \beta_1 Expertise_{f,t} + \gamma Control_{f,t} + \mu_f + \theta_t + \varepsilon_{f,t}. \quad (8)$$

The dependent variable is the average of four implied cost of capital measures from the accounting literature, following Eskildsen et al. (2026) and Mohanram and Gode (2013). The key explanatory variables are *Expertise* measured by *Knowledge* and *Quality*. Additionally, we test the interaction between knowledge and quality with analyst coverage. In this specification, we let our expertise measures interact with the number of analysts covering the firm:

$$\begin{aligned} ICC_{f,t} = & \beta_1 Expertise_{f,t} + \beta_2 AnalystCoverage_{f,t} \\ & + \beta_3 Expertise_{f,t} \times AnalystCoverage_{f,t} \\ & + \gamma Control_{f,t} + \mu_f + \theta_t + \varepsilon_{f,t}. \end{aligned} \quad (9)$$

Control variables include analyst coverage, firm size, trading activity, institutional ownership, past returns, return volatility, leverage, market-to-book, forecast error, forecast dispersion, beta, earnings variability, and long-term earnings growth.

Table 11 reports the results. Across all specifications, the coefficients on analyst *Knowledge*

and *Quality* are negative and statistically significant. A one-standard-deviation increase in analyst knowledge is associated with roughly an 8.5 basis point decline in implied cost of capital. The interaction estimates indicate that the effect of analyst knowledge and quality are conditional on the surrounding coverage environment. The coefficients are negative and significantly significant on the interaction term, suggesting the benefits of higher knowledge and quality measures are larger when there are more analysts covering a firm. For a firm with average analyst coverage, a one-standard-deviation increase in analyst knowledge is associated with an 8.7 basis point decline in implied cost of capital. For a firm with analyst coverage one standard deviation above the mean, the corresponding decline is 15.6 basis points, whereas for a firm one standard deviation below the mean the effect is only 1.8 basis points.

These findings indicate that firms followed by analysts with greater industry knowledge and higher quality research face a lower cost of equity capital. This result is economically intuitive and complements the evidence from the previous subsection. If stronger analysts reduce information asymmetry and improve price informativeness, investors should require a lower risk premium to hold the firm’s equity. The implied cost of capital results are consistent with analyst ability generating real financing benefits for firms through an improved information environment. Overall, the evidence in this subsection suggests that analyst industry knowledge has benefits to firms. Firms followed by more capable analysts appear to enjoy a more favorable information environment, characterized by better liquidity, more informative prices, and a lower implied cost of capital.

5.7 Natural Experiment

We exploit the acquisition of Credit Suisse by UBS as a plausibly exogenous shock to the information environment. The merger led to a reduction in analyst coverage for firms previously followed by Credit Suisse analysts, creating a natural setting to examine how the loss of analyst coverage of different expertise affects market quality.

5.7.1 Empirical Design

We implement a difference-in-differences framework comparing firms covered by Credit Suisse analysts (treatment group) to firms not covered by Credit Suisse (control group), before and after the

acquisition. To study heterogeneity within the treated firms, we construct firm-level measures of analyst Industry Knowledge and Report Quality using only the pre-acquisition period. These measures are time-invariant and reflect the average characteristics of Credit Suisse analysts covering each firm prior to the merger.

We then sort treated firms into high- and low-knowledge (or quality) groups based on whether their pre-period firm-level scores lie above or below the 75th percentile. The key coefficients of interest are the interaction terms between the post-acquisition indicator and these group indicators. This specification tests whether firms covered by more knowledgeable or higher-quality analysts are more affected by the loss of coverage.

The main analysis includes firms that have relatively limited analyst coverage in the pre-period, defined as those with at most 10 analysts on average. These firms are most likely to be affected by the removal of Credit Suisse analysts. We define the pre-period as the time from January 2022 to the acquisition announcement date of March 19, 2023 and the post-acquisition period from January 2024 to December 2024. The period from post acquisition announcement to the end of 2023 was when analyst layoffs likely took place. Since we do not know the exact time firings occur, we exclude this period of time from the difference-in-difference analysis.

We drop the period from the acquisition announcement date to the end of 2023 to avoid contaminating the DiD estimates with an ambiguous transition period. Although the acquisition was announced on March 19, 2023, the transaction was not legally completed until June 12, 2023, and UBS initially operated UBS AG and Credit Suisse AG as separate parent banks pending further integration. Moreover, the analyst-coverage shock did not occur immediately at announcement: Reuters reported that Credit Suisse began terminating Securities Research coverage on September 18, 2023, as UBS absorbed Credit Suisse’s research platform and moved institutional clients to UBS Global Research. UBS’s 2023 annual report further indicates that the bank finalized its target operating model and initiated the restructuring phase during 2023, with substantial integration-related expenses recognized during the year. We therefore treat March–December 2023 as a transition window in which treatment timing is potentially staggered and partially implemented, and define the post-period beginning in January 2024, after the research-coverage transition was more likely to have been realized.

We estimate the following difference-in-differences specification:

$$Y_{(i,t)} = \beta_1(Post_t \times Treat_i \times High_i) + \beta_2(Post_t \times Treat_i \times Low_i) + \gamma X_{(i,t)} + \alpha_i + \delta_t + \epsilon_{(i,t)} \quad (10)$$

where $Y_{(i,t)}$ denotes measures of market quality for firm i in week t , including liquidity measures (effective spread and realized spread) and price informativeness measures (variance ratio and lambda). $Post_t$ is an indicator equal to one for observations in the post-acquisition period and zero otherwise. $Treat_i$ is an indicator equal to one if firm i was covered by Credit Suisse analysts prior to the acquisition and zero otherwise. To capture heterogeneity within treated firms, we construct firm-level measures of analyst Industry Knowledge and Report Quality using only the pre-acquisition period. These measures are time-invariant and reflect the average characteristics of Credit Suisse analysts covering each firm. We define indicator variables $High_i$ and Low_i based on whether a firm's pre-period score lies above or below the 75th percentile threshold. The coefficients of interest, β_1 and β_2 , capture the differential effect of the acquisition for firms covered by higher- versus lower-knowledge (or quality) analysts relative to the control group of firms not covered by Credit Suisse. All specifications include firm fixed effects α_i and week fixed effects δ_t . The vector $X_{(i,t)}$ includes standard firm-level controls, including size, institutional ownership, profitability, returns, leverage, trading activity, and valuation measures.

5.7.2 Market Liquidity

Table 12 shows the results. Panel A reports the effects on liquidity. We find that the deterioration in liquidity is concentrated among firms that were covered by more knowledgeable Credit Suisse analysts prior to the acquisition. In the Knowledge-based specification, the coefficient on $Post \times Treatment \times High$ is positive and statistically significant across all three measures: effective spreads by 4.2 basis points and realized spreads by 2.4 basis points.

In contrast, the corresponding coefficients for $Post \times Treatment \times Low$ are small and statistically insignificant across all specifications. This pattern indicates that the loss of analyst coverage has economically meaningful effects only when the removed analysts provided high levels of industry

knowledge. Firms covered by lower-knowledge analysts do not experience a comparable deterioration in liquidity.

A similar pattern emerges when treatment is defined using analyst Quality. Firms covered by higher-quality analysts experience significant increases in spreads following the acquisition, while firms covered by lower-quality analysts again show little to no effect. The magnitudes are somewhat smaller than in the Knowledge specification, suggesting that industry knowledge may be the more important dimension.

5.7.3 Price Informativeness

Panel B reports the effects on price informativeness. We find that the loss of high-knowledge analysts also reduces the informational efficiency of prices. In the Knowledge specification, the $\text{Post} \times \text{Treatment} \times \text{High}$ coefficient is positive and statistically significant for both the variance ratio (0.02, $t = 3.98$) and lambda (0.52, $t = 2.84$), indicating a decline in price efficiency following the acquisition.

While $\text{Post} \times \text{Treatment} \times \text{Low}$ is statistically significant for the variance ratio, the magnitude is smaller, and there is no corresponding effect for lambda. This suggests that the deterioration in informativeness is again driven primarily by firms that were covered by more knowledgeable analysts. The Quality-based results are similar. Firms covered by higher-quality analysts exhibit significant declines in price efficiency, whereas the effects for lower-quality firms are weaker and less consistent across measures.

These results show that the effects of the acquisition operate through the loss of analyst expertise rather than simply the loss of coverage. The deterioration in liquidity and price informativeness is concentrated among firms that were covered by more knowledgeable or higher-quality Credit Suisse analysts in the pre-period. These findings provide causal evidence that analyst expertise matters for financial markets. The value of analyst coverage depends not only on its presence, but also on the underlying expertise of the analysts providing it.

5.7.4 High Coverage Firms

Furthermore, in the appendix we show the results for firms with high baseline analyst coverage, defined as those with more than 10 analysts on average during the pre-acquisition period. This sample provides a useful contrast, as these firms are less likely to be affected by the loss of any single brokerage’s analysts. Panel A of Table [IA5](#) shows that the effects of the acquisition on liquidity are substantially attenuated in this sample. The coefficients on $\text{Post} \times \text{Treatment} \times \text{High}$ are small and statistically insignificant across most specifications, and in some cases negative.

Similarly, the estimates for $\text{Post} \times \text{Treatment} \times \text{Low}$ are close to zero and not statistically distinguishable from zero. Panel B shows a similar pattern for price informativeness. The coefficients on $\text{Post} \times \text{Treatment} \times \text{High}$ are small and statistically insignificant across all specifications. While there is some evidence of a positive effect for $\text{Post} \times \text{Treatment Low}$ in one specification, the results are not consistent across measures and are economically modest.

The contrast with the main results indicates that the loss of Credit Suisse analysts does not meaningfully affect liquidity for firms with already extensive analyst coverage. For firms with limited analyst coverage, the removal of Credit Suisse analysts, particularly those with high industry knowledge or quality, leads to a significant deterioration in both liquidity and price informativeness. In these settings, analyst expertise is scarce, and its loss is not easily replaced. In contrast, for firms with extensive coverage, the same shock has little to no effect. The presence of many other analysts likely allows the market to substitute away from the lost coverage, mitigating any decline in market quality. This pattern strengthens the causal interpretation of the results. The effects are largest precisely in the settings where analyst information is most constrained and disappear when alternative sources of information are abundant.

6 Conclusion

Sell-side analysts are among the most important information intermediaries in capital markets, yet the literature has predominantly evaluated them through quantitative outputs such as earnings forecasts, stock recommendations, and target prices. Practitioners and institutional investors, by contrast, have long emphasized industry knowledge as one of the most valuable dimensions of

analyst expertise. We bridge this gap by developing large-scale, text-based measures of industry knowledge and research quality from 1.2 million industry reports using an LLM-based evaluation framework. Rather than inferring expertise from prior experience or from the outcomes it produces, our approach measures the knowledge content embedded in analyst research itself.

Our findings offer a direct answer to the question posed by the paper’s title. Industry knowledge and research quality are substantially more predictive of All-Star recognition than traditional forecast accuracy, with estimated coefficients roughly five times larger. They also predict more accurate earnings forecasts and stronger market reactions to analyst output. The same measures are informative beyond firm-level forecasting: more knowledgeable analysts issue industry calls with greater subsequent return predictability, while firms covered by such analysts exhibit better liquidity, more informative prices, and lower implied costs of capital. Taken together, these results suggest that what distinguishes top analysts is not simply the precision of their numerical outputs, but the depth of the industry knowledge and reasoning embodied in their research.

More broadly, the paper highlights an important distinction between observable inputs into expertise, observable outcomes of expertise, and the knowledge that connects the two. Measuring this intermediate object helps explain why analysts with similar credentials or numerical track records can differ substantially in the value of the information they produce. The evidence therefore points to industry knowledge as a central source of heterogeneity in analyst ability and, more generally, to the importance of the underlying knowledge held by information intermediaries in shaping capital-market outcomes.

References

- Amihud, Yakov. 2002. “Illiquidity and stock returns: cross-section and time-series effects.” *Journal of Financial Markets* 5 (1): 31–56.
- Amiram, Dan, Edward Owens, and Oded Rozenbaum. 2016. “Do information releases increase or decrease information asymmetry? New evidence from analyst forecast announcements.” *Journal of Accounting and Economics* 62 (1): 121–138.
- Asquith, Paul, Michael B Mikhail, and Andrea S Au. 2005. “Information content of equity analyst reports.” *Journal of Financial Economics* 75 (2): 245–282.
- Barber, Brad M, Reuven Lehavy, Maureen McNichols, and Brett Trueman. 2006. “Buys, holds, and sells: The distribution of investment banks’ stock ratings and the implications for the profitability of analysts’ recommendations.” *Journal of accounting and Economics* 41 (1-2): 87–117.
- Bastianello, Francesca, Paul Decaire, and Marius Guenzel. 2025. “Valuation Models.” *SSRN Electronic Journal*.
- Bhojraj, Sanjeev, Charles MC Lee, and Derek K Oler. 2003. “What’s my line? A comparison of industry classification schemes for capital market research.” *Journal of accounting research* 41 (5): 745–774.
- Boni, Leslie, and Kent L Womack. 2006. “Analysts, industries, and price momentum.” *Journal of Financial and Quantitative Analysis* 41 (1): 85–109.
- Bowen, Robert M., Xia Chen, and Qiang Cheng. 2008. “Analyst Coverage and the Cost of Raising Equity Capital: Evidence from Underpricing of Seasoned Equity Offerings.” *Contemporary Accounting Research* 25 (3): 657–700.
- Bradley, Daniel, Sinan Gokkaya, and Xi Liu. 2017. “Before an analyst becomes an analyst: Does industry experience matter?” *The Journal of Finance* 72 (2): 751–792.

- Bradley, Daniel, Sinan Gokkaya, Xi Liu, and Fei Xie. 2017. “Are All Analysts Created Equal? Industry Expertise and Monitoring Effectiveness of Financial Analysts.” *Journal of Accounting and Economics* 63 (2–3): 179–206.
- Bradshaw, Mark T. 2011. “Analysts’ forecasts: what do we know after decades of work?” *Available at SSRN 1880339*.
- Brav, Alon, and Reuven Lehavy. 2003. “An empirical analysis of analysts’ target prices: Short-term informativeness and long-term dynamics.” *The Journal of Finance* 58 (5): 1933–1967.
- Brown, Lawrence D, Andrew C Call, Michael B Clement, and Nathan Y Sharp. 2015. “Inside the “black box” of sell-side financial analysts.” *Journal of Accounting Research* 53 (1): 1–47.
- Chi, Feng, Byoung-Hyoun Hwang, and Yaping Zheng. 2025. “The Use and Usefulness of Big Data in Finance: Evidence from Financial Analysts.” *Management Science* 71 (6): 4599–4621.
- Claus, James, and Jacob Thomas. 2001. “Equity Premia as Low as Three Percent? Evidence from Analysts’ Earnings Forecasts for Domestic and International Stock Markets.” *The Journal of Finance* 56 (5): 1629–1666.
- Clement, Michael B. 1999. “Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?” *Journal of Accounting and Economics* 27 (3): 285–303.
- Clement, Michael B., and Senyo Y. Tse. 2005. “Financial Analyst Characteristics and Herding Behavior in Forecasting.” *The Journal of Finance* 60 (1): 307–341.
- Crawford, Steven S., Darren T. Roulstone, and Eric C. So. 2012. “Analyst Initiations of Coverage and Stock Return Synchronicity.” *The Accounting Review* 87, no. 5 (September): 1527–1553.

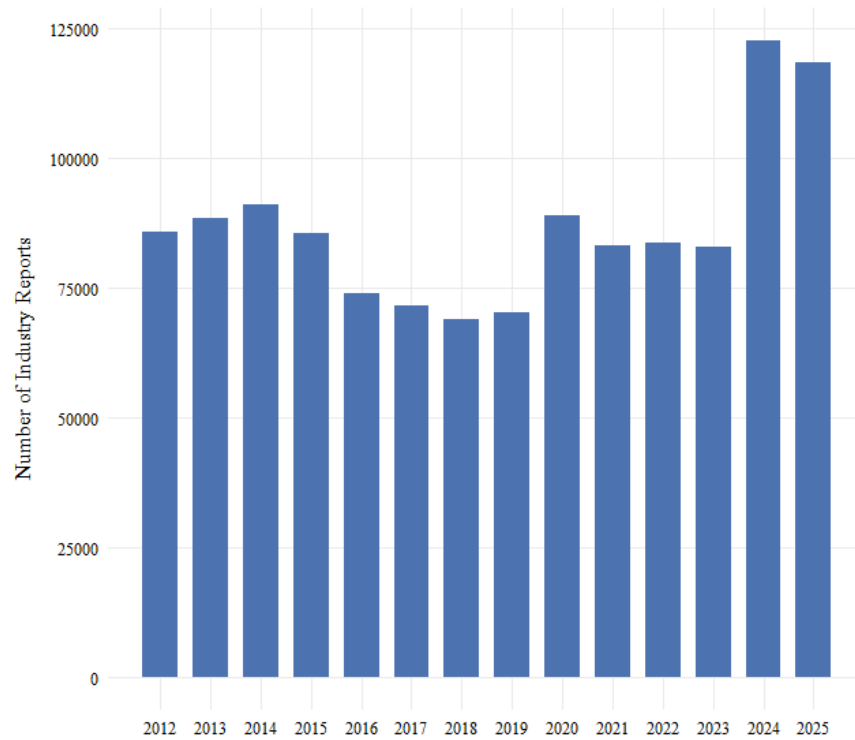
- De Franco, Gus, Ole-Kristian Hope, Dushyantkumar Vyas, and Yibin Zhou. 2015. “Analyst Report Readability.” *Contemporary Accounting Research* 32 (1): 76–104.
- Décaire, Paul, and John R Graham. 2024. “Valuation fundamentals.” *HKU Jockey Club Enterprise Sustainability Global Research Institute-Archive*.
- Derrien, François, and Ambrus Kecskés. 2013. “The Real Effects of Financial Shocks: Evidence from Exogenous Changes in Analyst Coverage.” *The Journal of Finance* 68 (4): 1407–1440.
- Easton, Peter D. 2004. “PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital.” *The Accounting Review* 79 (1): 73–95.
- Eskildsen, Marc, Markus Ibert, Theis Ingerslev Jensen, and Lasse Heje Pedersen. 2026. “In Search of the True Greenium.”
- Gebhardt, William R., Charles M. C. Lee, and Bhaskaran Swaminathan. 2001. “Toward an Implied Cost of Capital.” *Journal of Accounting Research* 39 (1): 135–176.
- Gibbons, Brian, Peter Iliev, and Jonathan Kalodimos. 2021. “Analyst information acquisition via EDGAR.” *Management Science* 67 (2): 769–793.
- Gleason, Cristi A., W. Bruce Johnson, and Haidan Li. 2013. “Valuation Model Use and the Price Target Performance of Sell-Side Equity Analysts.” *Contemporary Accounting Research* 30 (1): 80–115.
- Gokkaya, Sinan, Xi Liu, Veronika Kreply Pool, Fei Xie, and Jinfan Zhang. 2023. “Is There Investment Value in the Soft-Dollar Arrangement? Evidence from Mutual Funds.” *Review of Financial Studies* 36 (8): 3122–3162.
- Groysberg, Boris, Paul M Healy, and David A Maber. 2011. “What drives sell-side analyst compensation at high-status investment banks?” *Journal of Accounting Research* 49 (4): 969–1000.

- Guttman, Ilan. 2010. "The Timing of Analysts' Earnings Forecasts." *The Accounting Review* 85 (2): 513–545.
- Harford, Jarrad, Feng Jiang, Rong Wang, and Fei Xie. 2019. "Analyst career concerns, effort allocation, and firms' information environment." *The Review of Financial Studies* 32 (6): 2179–2224.
- Hirshleifer, David, Yaron Levi, Ben Lourie, and Siew Hong Teoh. 2019. "Decision fatigue and heuristic analyst forecasts." *Journal of Financial Economics* 133 (1): 83–98.
- Huang, Allen H, An-Ping Lin, and Amy Y Zang. 2022. "Cross-industry information sharing among colleagues and analyst research." *Journal of Accounting and Economics* 74 (1): 101496.
- Huang, Allen H, Amy Y Zang, and Rong Zheng. 2014. "Evidence on the information content of text in analyst reports." *The Accounting Review* 89 (6): 2151–2180.
- Huang, Allen H., Reuven Lehavy, Amy Y. Zang, and Rong Zheng. 2018. "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach." *Management Science* 64 (6): 2833–2855.
- Ivković, Zoran, and Narasimhan Jegadeesh. 2004. "The timing and value of forecast and recommendation revisions." *Journal of Financial Economics* 73 (3): 433–463.
- Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D Krische, and Charles MC Lee. 2004. "Analyzing the analysts: When do recommendations add value?" *The journal of finance* 59 (3): 1083–1124.
- Kadan, Ohad, Leonardo Madureira, Rong Wang, and Tzachi Zach. 2012. "Analysts' industry expertise." *Journal of Accounting and Economics* 54 (2-3): 95–120.

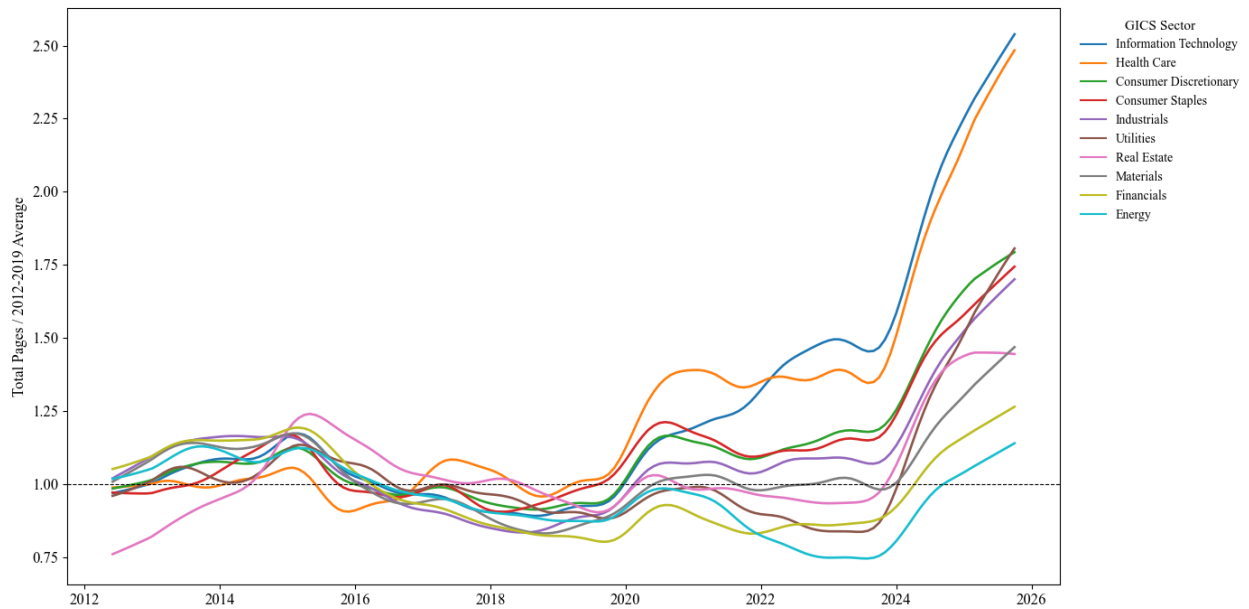
- Kecskés, Ambrus, Roni Michaely, and Kent L Womack. 2017. “Do earnings estimates add value to sell-side analysts’ investment recommendations?” *Management Science* 63 (6): 1855–1871.
- Kelly, Bryan, and Alexander Ljungqvist. 2012. “Testing Asymmetric-Information Asset Pricing Models.” *The Review of Financial Studies* 25 (5): 1366–1413.
- Law, Kelvin KF. 2023. “Good-Bye I/B/E/S (or Not?)” *Journal of Financial Reporting* 8 (1): 41–61.
- Li, Congcong, An-Ping Lin, Hai Lu, and Kevin Veenstra. 2020. “Gender and beauty in the financial analyst profession: evidence from the United States and China.” *Review of Accounting Studies* 25 (4): 1230–1262.
- Li, Kai, Feng Mai, Gabriel Wong, Chelsea Yang, and Tengfei Zhang. 2026. “Female equity analysts and corporate environmental and social performance.” *Management Science*.
- Liu, Pu, Stanley D Smith, and Azmat A Syed. 1990. “Stock price reactions to the Wall Street Journal’s securities recommendations.” *Journal of Financial and Quantitative Analysis* 25 (3): 399–410.
- Loh, Roger K, and G Mujtaba Mian. 2006. “Do accurate earnings forecasts facilitate superior investment recommendations?” *Journal of Financial Economics* 80 (2): 455–483.
- Loh, Roger K., and René M. Stulz. 2018. “Is Sell-Side Research More Valuable in Bad Times?” *The Journal of Finance* 73 (3): 959–1013.
- Mohanram, Partha, and Dan Gode. 2013. “Removing predictable analyst forecast errors to improve implied cost of equity estimates.” *Review of Accounting Studies* 18, no. 2 (June): 443–478.
- O’Brien, Patricia C. 1988. “Analysts’ forecasts as earnings expectations.” *Journal of Accounting and Economics* 10 (1): 53–83.

- O'Brien, Patricia C. 1990. "Forecast accuracy of individual analysts in nine industries." *Journal of Accounting Research* 28 (2): 286–304.
- Ohlson, James A., and Beate E. Juettner-Nauroth. 2005. "Expected EPS and EPS Growth as Determinants of Value." *Review of Accounting Studies* 10:349–365.
- OpenAI, Sandhini Agarwal, Lama Ahmad, Jason Ai, Sam Altman, Andy Applebaum, Edwin Arbus, et al. 2025. "Gpt-Oss-120b & Gpt-Oss-20b Model Card." Pre-published, August 8, 2025. Accessed March 13, 2026.
- Piotroski, Joseph D., and Barren T. Roulstone. 2004. "The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices." *The Accounting Review* 79 (4): 1119–1151.
- Trueman, Brett. 1994. "Analyst Forecasts and Herding Behavior." *The Review of Financial Studies* 7 (1): 97–124.
- Wang, Baolian, Peixin Li, and Jiawei Yu. 2025. "The Impact of Return-to-Office Mandates on Equity Analysts." *Available at SSRN 5206058*.
- Womack, Kent L. 1996. "Do brokerage analysts' recommendations have investment value?" *The Journal of Finance* 51 (1): 137–167.

Figures



Panel A: Annual Number of Industry Reports



Panel B: Industry Report Coverage Over Time

Figure 1: Industry Report Volume and Sector Coverage

This figure displays the annual number of industry reports and sector-level coverage intensity for 40 brokerage houses from January 2012 to October 2025. Panel A shows the total number of reports issued each year across all brokerage houses in the sample. Panel B plots the monthly total pages index by GICS sector, normalized to each sector's 2012–2019 average, with a six-month moving average applied. A value of two indicates coverage intensity twice the sector's historical benchmark.

January 11, 2012

Industry View
In-Line

Software

Finding Diamonds & Avoiding Disasters in 2012

While the macro is a concern, and 1H will be weak as usual, our CIO work points to similar software spending growth in 2012 as 2011 and we are more constructive on the group than cons. with MSFT, CTXS, FTNT, RHT, MKTG and SREV our top picks.

For 2012, our top picks have exposure to a variety of the attractive elements of the sector: a) accelerating EPS or cash flow, b) low margin expectations, c) limited currency risk, d) reasonable valuations, e) exposure to secular drivers and f) potential strategic interest. We would avoid network & systems management, material EPS deceleration, high currency risk or stretched value.

2012 Likely another Tough Yr., But Not Worse than CY11 & Expectations are Low: CY11 saw ~3% IT growth YoY, which was revised down through the yr as CY11 was hard well before ORCL's Q2 (70% of our co's missed/ lowered in CY11). Our CIO Survey calls for stable IT growth of ~3% YoY in CY12, which is low, but better than some fear. Bottoms up rev. forecasts are even more conservative for CY12, with constant currency rev. growth of +16% for the group, down from 19% in CY11 (+14% down from +22% as reported), and the average co. modelling minimal margin expansion.

Valuation Reasonable: Mid & Large cap software was down 10% YoY in CY11, driving the avg. multiple across all software to be ~25% lower from the Jan11 peak. The premium to the S&P is in the lowest 10 yr. quartile and the NTM EPS multiple is below the 5 & 10 yr. avgs. Our work suggests multiples are likely within 5-10% of the bottom, yet our base case price targets yield 15% returns for the group without any multiple expansion.

The Best Secular Demand Drivers in 10 Years should enable software to remain a share gainer of IT. Key themes: 1) continued migration to public & private clouds, 2) acceleration of social computing, 3) data growth driving analytics and new data storage & mgmt., 4) security gets smart, 5) new devices and application paradigms creating new opportunities. We also believe M&A will accelerate in 2012, similar to 2005.

MORGAN STANLEY RESEARCH NORTH AMERICA

Morgan Stanley & Co. LLC

Adam Holt
Adam.Holt@morganstanley.com
+1 415 576 2320**Keith Weiss, CFA**
Keith.Weiss@morganstanley.com
+1 212 761 4149**Jennifer Swanson, CFA**
Jennifer.Swanson@morganstanley.com
+1 212 761 3665

Ticker	Rating	Price Target/Base Case		% Upside
		Current	Prior	
CRM	OW	\$150	\$180	43%
FTNT	OW	\$28	\$28	41%
CTXS	OW	\$90	\$90	40%
MKTG	OW	\$13	\$11	35%
SREV	OW	\$20	\$19	28%
RHT	OW	\$54	\$54	26%
VMW	OW	\$105	\$115	23%
ORCL	OW	\$33	\$33	22%
MSFT	OW	\$33	\$31	19%
CHKP	EW	\$65	\$65	24%
QLIK	EW	\$28	\$28	19%
ADSK	EW	\$36	\$36	11%
SYMC	EW	\$17	\$17	7%
INTU	EW	\$58	\$53	6%
AKAM	EW	\$35	\$26	6%
FIRE	EW	\$30	\$26	-1%
ADBE	EW	\$28	\$28	-4%
TLEO	EW	\$35	\$28	-6%
WBSN	UW	\$17	\$17	-8%
GLUD	UW	\$5	\$5	-25%
SWI	UW	\$22	\$22	-26%
N	UW	\$28	\$26	-29%
Median Base Case Upside:				15%

Save the date

MORGAN STANLEY

TECHNOLOGY, MEDIA & TELECOM CONFERENCE

February 27–March 1, 2012
Palace Hotel | San Francisco, California

Morgan Stanley does and seeks to do business with companies covered in Morgan Stanley Research. As a result, investors should be aware that the firm may have a conflict of interest that could affect the objectivity of Morgan Stanley Research. Investors should consider Morgan Stanley Research as only a single factor in making their investment decision.

For analyst certification and other important disclosures, refer to the Disclosure Section, located at the end of this report.

January 11, 2012
Software

Positioning for 2012

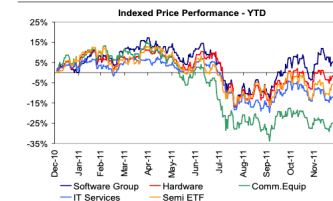
Contents

Positioning for 2012	2
Our Top Picks	4
Risk Reward Summary	5
Where We Are in the Cycle; a Look at Multiples and Performance	7
Expectations are Modest Heading into 2012	16
Earnings Should Prove Defensive in 2012	25
Secular Growth Drivers	
a) Apps and Platforms Move to the Cloud	30
b) Social Enters the Enterprise	32
c) Big Data: Finding New Opportunities in Growing Data Volumes	33
d) Security Gets Smart	34
e) Mktg Apps Adapt to Changing Media	38
f) Mobile & HTML5 Change App Ecosystem	39
g) M&A in Software Expected to Continue	40
Macro Indicators and IT Spending	43
Risk Reward Profiles	50

While EPS grew ~18% on average across software in 2011, multiples contracted 25% YoY, reversing the expansionary trend seen in 2009 and 2010, and driving mid & large cap software to decline 10% for the year, and underperform the Nasdaq which was down 2%. Growth again outperformed value and small cap meaningfully outperformed large cap—as 3 of our 27 covered companies were acquired, in turn driving up multiples across small cap. In aggregate, M&A resulted in another year of industry outperformance for software.

Exhibit 1

Software Stocks Outperformed Tech in 2011



Source: Thomson Reuters, Morgan Stanley Research

In our 15 years on the sell side, we can't recall as much uncertainty as we face heading into 2012. Even in 2001/2002 and 2008/2009, when things were terrible, at least the direction seemed clear. For 2012, we think as much can break right as can break wrong. Investors are largely bearish, and volumes reflect little conviction. Currency is creating a headwind for software in 2012, European demand is unstable, emerging market growth may be slowing, PCs and Servers remain weak, and Financial Services and Federal spending are slowing. Our U.S. strategist Adam Parker is calling for multiples to contract further in 2012, and the S&P to end at 1237, down 4% from ~1290 today. Still, our group price targets yield 15% returns even with multiples coming down given solid growth—and positive revisions could be an added lever.

At the same time, software companies are much better off in many ways than heading into last year and stocks are more attractive as a result. Our average co. now has over 50% of its revenue recurring, cash as a % of the market cap is up 250% since mid 2008, companies are still coming off cyclical lows in employment, new business starts and other key indicators, while companies have expanded sales capacity by 10-40% over the last year. We expect IT spending to remain relatively stable in CY12 at 2.9% growth vs. 2.7% in CY11, based on our latest CIO survey—which is better than most fear. Even if growth slows modestly, it shouldn't present material risk to estimates. While these spending levels are well below the 10-year average, stable growth in 2012 will result in upside opportunities in the best-positioned companies.

The secular demand drivers are as good as we have seen in 10 years, and we expect M&A to remain brisk given the big increase in cash on hand and the lack of organic growth in other tech industries. From a stock perspective, multiples appear to be within 10-15% of troughs on a number of metrics, and we have a dozen stocks with installed bases alone representing 60%+ of their values. Software expectations for many companies look low for CY12, with forecasted con. currency rev. growth of 16% for the group, down from 19% in CY11 (excluding USD reporters), and the average company modeling minimal margin expansion in CY12.

We recognize that the first half will be seasonally slow, however, and that estimates still need to come down in some areas of the group—particular around currency, while revenue and EPS growth will decelerate for many stocks. This keeps us from upgrading the group now, despite our positive leaning.

Still, we see opportunity for absolute returns with stocks that have some or all of the following characteristics: a) accelerating EPS or cash flow, b) low margin expectations, c) limited currency risk, d) reasonable valuations, e) exposure to secular drivers and f) strategic value. We would avoid network

Figure 2: Example of Software Industry Report (Morgan Stanley, January 11, 2012)

TICKER	CNAME	ESTIMID	ETEXT	AMASKCD	ANNDATS	FIRM_REC	IND_REC
WBMD	WEBMD HEALTH	MORGAN	UNDERWT/ATTRACTIVE	80523	11-Jan-12	UNDERWT	ATTRACTIVE

Panel A. I/B/E/S Record

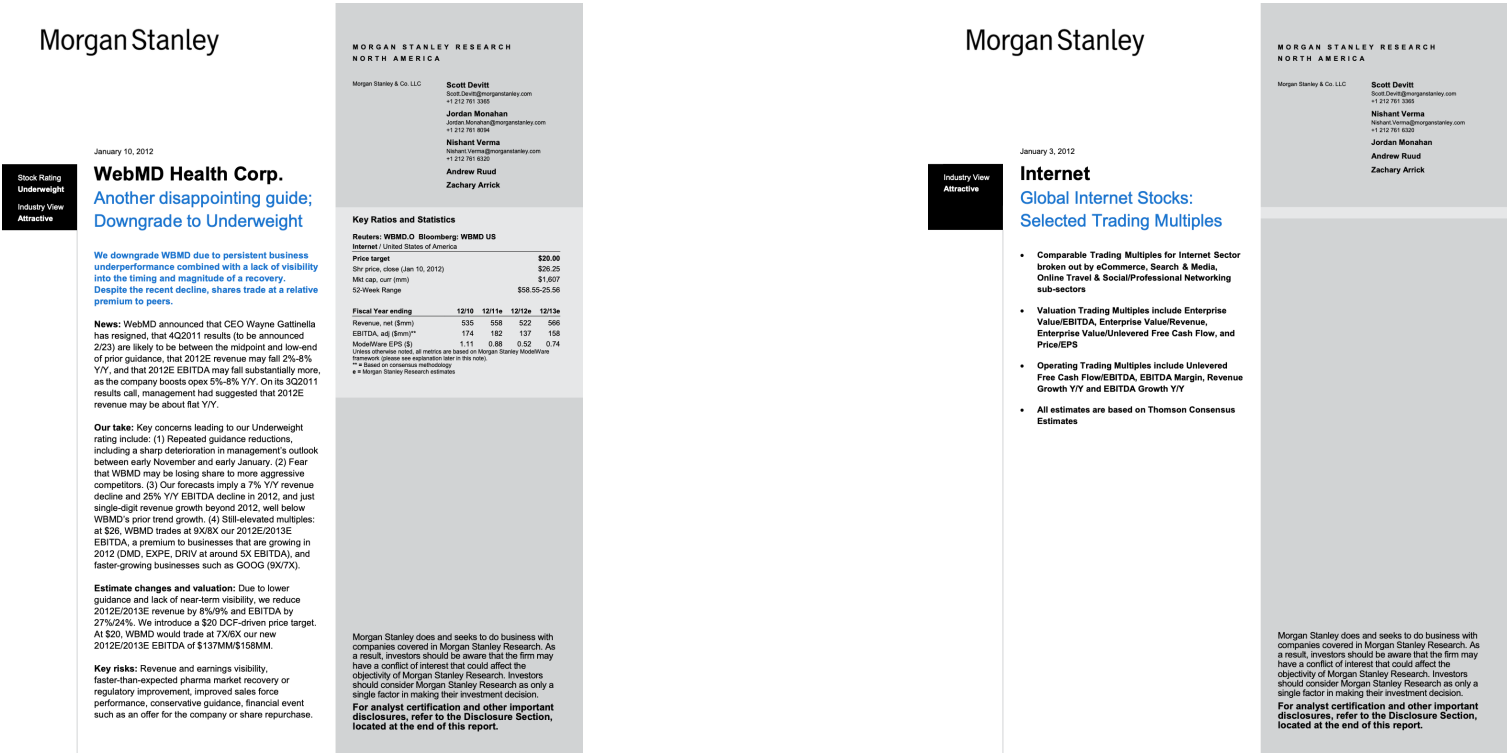
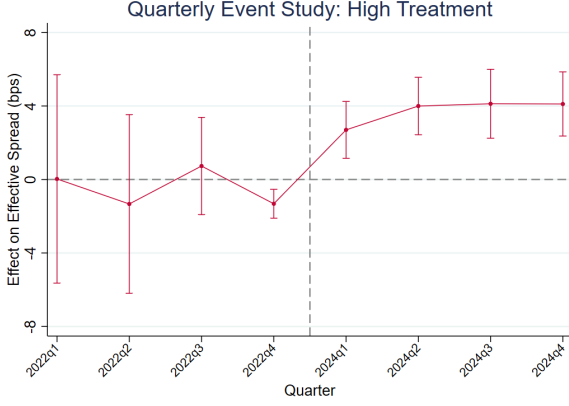


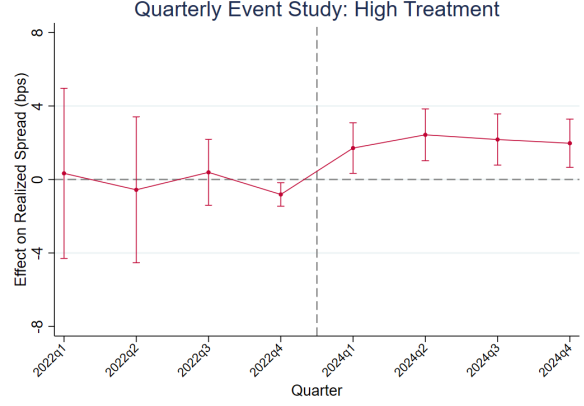
Figure 4: Dynamic Effects of Market Outcomes (Knowledge)

This figure plots the dynamic effects of Credit Suisse coverage reductions on the main market-quality outcomes. We estimate quarterly event-study specifications over a window spanning 2022Q1 through 2024Q4, excluding the transition quarters after the acquisition announcement and before the post-period. The plotted point estimates are the coefficients on event-time indicators interacted with treatment status, shown for firms with exposure to high Credit Suisse analyst knowledge. The omitted quarter is 2023Q1, which serves as the pre-event baseline because the acquisition was announced late in the quarter. The post-period coefficients correspond to 2024Q1 through 2024Q4. The regressions include firm fixed effects, quarter fixed effects, and the same controls as in the baseline difference-in-differences specifications. Standard errors are double clustered by firm and quarter. Vertical bars represent 95% confidence intervals.

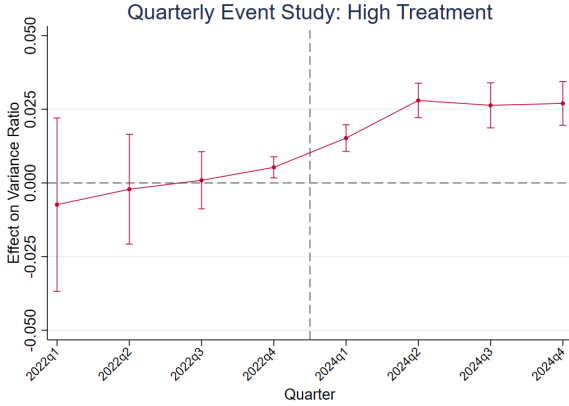
(a) Effective Spread



(b) Realized Spread



(c) Variance Ratio



(d) Lambda

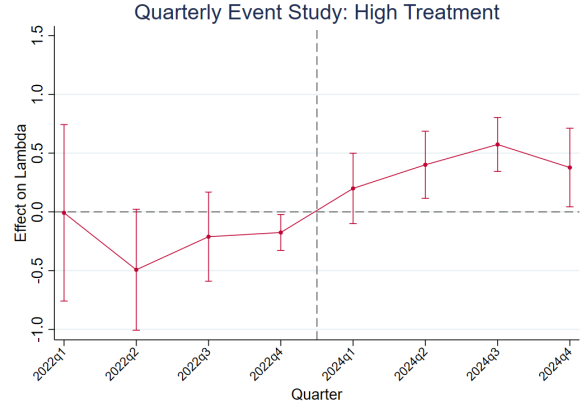
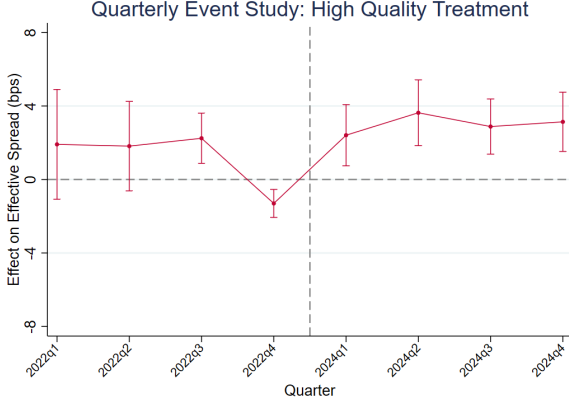


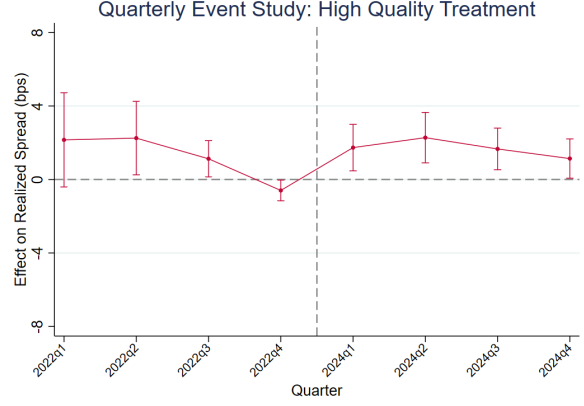
Figure 5: Dynamic Effects of Market Outcomes (Quality)

This figure plots the dynamic effects of Credit Suisse coverage reductions on the main market-quality outcomes. We estimate quarterly event-study specifications over a window spanning 2022Q1 through 2024Q4, excluding the transition quarters after the acquisition announcement and before the post-period. The plotted point estimates are the coefficients on event-time indicators interacted with treatment status, shown for firms with exposure to high Credit Suisse analyst quality. The omitted quarter is 2023Q1, which serves as the pre-event baseline because the acquisition was announced late in the quarter. The post-period coefficients correspond to 2024Q1 through 2024Q4. The regressions include firm fixed effects, quarter fixed effects, and the same controls as in the baseline difference-in-differences specifications. Standard errors are double clustered by firm and quarter. Vertical bars represent 95% confidence intervals.

(a) Effective Spread



(b) Realized Spread



(c) Variance Ratio



(d) Lambda



Tables

Table 1: Summary Statistics

This table presents summary statistics. Panel A reports statistics for analyst-monthly LLM-based measures, including the composite *Knowledge* and *Quality* scores, report characteristics (*No. of reports* and *Average report length*), and the individual knowledge and quality dimensions. Panel B reports statistics for variables used in the *Expertise* determinant and All-Star logit analyses. Panel C reports statistics for variables used in the forecast error and market reaction analyses. All variables are defined in Section 4.

Panel A. Analyst-Monthly LLM Measures						
Variable	N	Mean	SD	P1	P50	P99
<i>Knowledge</i>	303,656	2.61	0.74	1.00	2.75	3.88
<i>Quality</i>	303,656	3.37	0.57	1.50	3.50	4.30
<i>No. of reports</i>	303,656	4.00	5.00	1.00	2.00	25.00
<i>Average report length</i>	303,656	20.04	21.01	3.00	14.00	104.90
<i>Industry technology</i>	303,656	2.50	1.04	1.00	3.00	4.00
<i>Industry trends</i>	303,656	3.52	0.73	1.00	4.00	4.00
<i>Supply chain</i>	303,656	2.48	1.05	1.00	2.00	4.00
<i>Distribution model</i>	303,656	2.45	1.00	1.00	3.00	4.00
<i>Margin</i>	303,656	3.08	1.03	1.00	3.00	4.00
<i>Customer</i>	303,656	2.91	0.97	1.00	3.00	4.00
<i>Labor</i>	303,656	1.62	0.93	1.00	1.00	4.00
<i>Management team</i>	303,656	2.28	0.91	1.00	2.00	4.00
<i>Originality</i>	303,656	2.62	0.83	1.00	3.00	4.00
<i>Data</i>	303,656	3.24	0.75	1.00	3.00	5.00
<i>Analytical</i>	303,656	3.09	0.75	1.00	3.00	4.00
<i>Actionability</i>	303,656	3.28	0.89	1.00	4.00	4.00
<i>Structure</i>	303,656	3.82	0.48	2.00	4.00	5.00
<i>Forward</i>	303,656	3.48	0.73	1.00	4.00	4.00
<i>Technical</i>	303,656	3.70	0.66	1.00	4.00	5.00
<i>Quantitative</i>	303,656	3.33	0.72	1.00	3.00	4.00
<i>Risk</i>	303,656	3.18	0.90	1.00	3.00	4.00
<i>Timeliness</i>	303,656	3.96	0.54	2.00	4.00	5.00
Panel B. Expertise Determinant Analysis and All-Star Analysis						
<i>No. of reports</i>	32,373	3.15	5.01	0.10	1.30	26.10
<i>Average report length</i>	32,373	2.20	2.06	0.40	1.68	10.65
<i>Brokerage size</i>	32,373	139.20	82.39	12.00	135.00	290.00
<i>Analyst career length</i>	13,122	12.56	10.71	0.00	9.00	39.00
<i>No. of industries (GICS)</i>	13,122	3.43	2.42	1.00	3.00	11.00
<i>No. of firms</i>	13,122	12.78	9.35	1.00	13.00	36.00
<i>Average firm size (thousands \$)</i>	13,122	32,824.99	51,908.06	430.50	15,709.57	337,097.63
<i>Average forecast error</i>	13,122	2.25	4.12	0.05	0.97	29.06
Panel C. Forecast Error Analysis and Market Reaction Analysis						
<i>Forecast error</i>	602,574	1.60	3.60	0.00	0.43	25.38
<i>Forecast horizon</i>	602,574	5.12	0.78	2.08	5.31	5.90
<i>Market value (thousands \$)</i>	602,574	31,197.57	64,242.32	93.95	8,262.20	401,586.63
<i>Institutional ownership</i>	602,574	0.69	0.32	0.00	0.81	1.00
<i>No. of analysts</i>	602,574	18.77	9.69	3.00	18.00	46.00
<i>Brokerage size</i>	602,574	74.33	29.79	13.00	73.00	133.00
<i>Average forecast error</i>	602,574	2.16	3.20	0.17	1.15	21.84
<i>Analyst career length</i>	602,574	15.92	10.83	1.00	13.00	39.00
<i>Time covering company</i>	602,574	5.98	6.08	0.00	4.00	29.00
<i>Forecast revision</i>	393,662	-0.03	1.36	-6.28	0.01	5.63
<i>CAR_[-1,+1]</i>	393,662	-0.07	6.65	-22.70	-0.03	21.86
<i>BM(book to market)</i>	393,662	0.53	0.45	0.01	0.41	2.50
<i>Lag monthly return</i>	393,662	0.97	11.55	-34.68	1.04	38.73
<i>Volatility</i>	393,662	2.35	1.48	0.71	1.91	8.64
<i>Momentum</i>	393,662	4.69	25.96	-58.25	3.71	100.70

Table 2: Determinants of Analyst Industry Expertise

This table reports OLS regression results examining the determinants of analyst-year LLM-based industry expertise, proxied by *Knowledge* and *Quality*. The dependent variables are the analyst-year *Knowledge* measure in columns (1)–(4) and the *Quality* measure in columns (5)–(8). The independent variables include three sets: the lagged dependent variable, report and brokerage characteristics (*No. of reports*, *Average report length*, and *Brokerage size*), and analyst characteristics from I/B/E/S (*Analyst career length*, *No. of industries*, *Average firm size*, and *Industry Experience*). All independent variables are defined in Section 4 and measured contemporaneously, with the exception of the lagged dependent variable. All specifications include brokerage, industry, and year fixed effects, while columns (4) and (8) additionally include analyst fixed effects. Industry fixed effects are based on the analyst’s dominant mapped GICS sector in a given year, defined as the sector in which the analyst issues the largest number of reports. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the analyst and year levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	<i>Knowledge</i>				<i>Quality</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Lag(Knowledge)</i>	0.479*** (27.47)	0.564*** (30.87)	0.462*** (26.90)	0.109*** (3.88)				
<i>Lag(Quality)</i>					0.431*** (22.12)	0.540*** (26.41)	0.415*** (20.62)	0.111*** (3.32)
<i>No. of reports</i>	0.032*** (13.13)		0.031*** (11.42)	0.045*** (9.33)	0.035*** (13.32)		0.035*** (11.47)	0.044*** (9.56)
<i>Average report length</i>	0.164*** (14.52)		0.160*** (13.33)	0.191*** (9.09)	0.170*** (15.18)		0.166*** (14.06)	0.189*** (10.21)
<i>Ln(Brokerage size)</i>	0.065 (1.64)		0.085** (2.25)	0.173*** (3.39)	0.067 (1.46)		0.080 (1.73)	0.164*** (3.19)
<i>Ln(Analyst career length)</i>		0.035*** (3.65)	0.000 (0.01)	0.080* (1.87)		0.033*** (3.52)	-0.004 (-0.37)	0.092** (2.26)
<i>Ln(No. of industries)</i>		0.057** (2.23)	0.037 (1.38)	0.006 (0.17)		0.009 (0.39)	-0.024 (-0.96)	-0.001 (-0.02)
<i>Ln(Average firm size)</i>		0.047*** (7.61)	0.035*** (4.88)	0.027 (1.76)		0.047*** (8.83)	0.034*** (5.33)	0.025 (1.57)
<i>Industry Experience</i>		0.004 (0.13)	-0.015 (-0.52)	0.078 (1.18)		0.018 (0.56)	0.005 (0.16)	0.062 (1.02)
Brokerage FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Analyst FE				Y				Y
N	10,621	9,386	9,386	8,937	10,621	9,386	9,386	8,937
Adj. R^2	0.546	0.452	0.535	0.667	0.514	0.416	0.510	0.638

Table 3: What Makes a Star Analyst?

This table reports logit regression results examining the determinants of *Institutional Investor* All-Star ratings. The dependent variable *AllStar* is an indicator equal to one if the analyst is named among the top three or as a runner-up in their industry category in a given year, and zero otherwise. The independent variables include LLM-based industry expertise (proxied by *Knowledge* and *Quality*), *Average forecast error*, *Analyst career length*, *Brokerage size*, *No. of industries*, *Average firm size*, *Industry Experience*, and lagged All-Star status. All independent variables are defined in Section 4 and lagged by one period. To facilitate comparison, *Knowledge*, *Quality*, and *Average forecast error* are standardized to have zero mean and unit variance. All specifications include brokerage, industry, and year fixed effects. Industry fixed effects are based on the analyst's dominant mapped GICS sector in a given year, defined as the sector in which the analyst issues the largest number of reports. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the analyst and year levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	<i>Pr(AllStar=1)</i>						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Knowledge</i>	0.636*** (10.05)			0.509*** (8.19)		0.295*** (6.50)	
<i>Quality</i>		0.629*** (9.59)			0.530*** (8.16)		0.289*** (6.39)
<i>Average forecast error</i>			-0.128*** (-3.25)	-0.017 (-0.49)	-0.024 (-0.72)	-0.095*** (-2.74)	-0.099*** (-2.85)
<i>Ln(Analyst career length)</i>				0.919*** (11.18)	0.911*** (10.96)	0.303*** (4.24)	0.298*** (4.10)
<i>Ln(Brokerage size)</i>				0.521* (1.75)	0.524* (1.74)	0.503* (1.95)	0.505* (1.96)
<i>Ln(No. of industries)</i>				1.322*** (10.33)	1.386*** (10.84)	0.866*** (7.25)	0.904*** (7.62)
<i>Ln(Average firm size)</i>				0.251*** (5.17)	0.250*** (5.11)	0.155*** (3.73)	0.156*** (3.68)
<i>Industry Experience</i>						0.322** (2.18)	0.304** (2.04)
<i>Lag(AllStar)</i>						4.044*** (15.57)	4.039*** (15.52)
Brokerage FE	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
N	10,866	10,866	10,866	10,823	10,823	9,629	9,629

Table 4: Industry Expertise and EPS Forecast Accuracy

This table reports OLS regression results examining the relation between industry expertise and EPS forecast accuracy. The dependent variable *Forecast error* and the key independent variables, analyst-event-specific *Expertise*, proxied by *Knowledge* in Columns (1)–(2) and *Quality* in Columns (3)–(4); see Section 5.3 for details. *Knowledge* and *Quality* are standardized to have zero mean and unit variance. All control variables are defined in Section 4. With the exception of *Ln(Analyst career length)*, *Ln(Time covering company)*, and *Ln(Forecast horizon)*, which are measured based on the forecast date, all other control variables are from the prior year. Columns (1) and (3) include analyst, firm, and year-month fixed effects, while Columns (2) and (4) include analyst-firm and year-month fixed effects. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the analyst and firm levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	<i>Forecast error</i>			
	(1)	(2)	(3)	(4)
<i>Knowledge</i>	-0.018*** (-2.73)	-0.020*** (-3.05)		
<i>Quality</i>			-0.015** (-2.36)	-0.017*** (-2.78)
<i>Average forecast error</i>	0.007 (0.90)	0.005 (0.59)	0.007 (0.89)	0.005 (0.58)
<i>Ln(Market value)</i>	-1.586*** (-21.42)	-1.643*** (-19.80)	-1.586*** (-21.42)	-1.643*** (-19.80)
<i>Institutional ownership</i>	-0.944*** (-3.83)	-0.865*** (-3.56)	-0.944*** (-3.83)	-0.865*** (-3.56)
<i>Ln(No. of analysts)</i>	0.335*** (8.28)	0.240*** (6.23)	0.335*** (8.28)	0.240*** (6.23)
<i>Ln(Brokerage size)</i>	0.106 (1.54)	0.117 (1.37)	0.106 (1.53)	0.117 (1.36)
<i>Ln(Analyst career length)</i>	0.029 (0.38)	-0.205* (-1.95)	0.029 (0.38)	-0.205* (-1.95)
<i>Ln(Time covering company)</i>	0.086*** (5.88)	0.432*** (7.24)	0.086*** (5.87)	0.432*** (7.24)
<i>Ln(Forecast horizon)</i>	0.464*** (24.53)	0.471*** (24.32)	0.464*** (24.54)	0.471*** (24.33)
Analyst FE	Y		Y	
Firm FE	Y		Y	
Analyst-Firm FE		Y		Y
Year-Month FE	Y	Y	Y	Y
N	591,267	587,443	591,267	587,443
Adj. R^2	0.437	0.499	0.437	0.499

Table 5: Market Reaction to EPS Forecast Revisions

This table reports OLS regression results examining whether the market reaction to EPS forecast revisions varies with analysts' industry expertise. The dependent variable $CAR_{[-1,+1]}$ is the cumulative market-model-adjusted abnormal return over a three-day event window around the forecast revision date. The key variables of interest are the interaction terms $Forecast\ revision \times Expertise$, where *Expertise* is proxied by *Knowledge* in Column (1) and *Quality* in Column (2). *Knowledge* and *Quality* are standardized to have zero mean and unit variance. *Forecast revision* is the difference between the analyst's revised forecast and the prior forecast, scaled by the one-year lagged stock price and multiplied by 100. All control variables are defined in Section 4 and are either from the prior year or measured using information available prior to the forecast revision date. We exclude observations with stock prices below \$1. All specifications include analyst, firm, and year-month fixed effects. *t*-statistics in parentheses. Standard errors are clustered at the firm level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$CAR_{[-1,+1]}$	
	(1)	(2)
<i>Forecast revision</i> $\times$ <i>Knowledge</i>	0.104*** (7.75)	
<i>Knowledge</i>	0.001 (0.09)	
<i>Forecast revision</i> $\times$ <i>Quality</i>		0.029** (2.19)
<i>Quality</i>		0.013 (0.96)
<i>Forecast revision</i>	0.723*** (27.59)	0.710*** (27.54)
<i>Average forecast error</i>	-0.023*** (-3.51)	-0.023*** (-3.50)
<i>Ln(Market value)</i>	-1.003*** (-13.52)	-1.001*** (-13.48)
<i>Institutional ownership</i>	0.124 (0.59)	0.128 (0.61)
<i>Ln(No. of analysts)</i>	-0.043 (-0.64)	-0.044 (-0.65)
<i>Ln(Brokerage size)</i>	0.035 (0.71)	0.034 (0.70)
<i>Ln(Analyst career length)</i>	-0.000 (-0.00)	-0.003 (-0.04)
<i>Ln(Time covering company)</i>	-0.017 (-0.89)	-0.016 (-0.88)
<i>BM</i>	-0.146 (-1.53)	-0.146 (-1.53)
<i>Lag monthly return</i>	0.070*** (29.02)	0.070*** (29.08)
<i>Ln(Volatility)</i>	0.312** (2.49)	0.306** (2.44)
<i>Momentum</i>	-0.014*** (-12.09)	-0.014*** (-11.99)
Analyst FE	Y	Y
Firm FE	Y	Y
Year-Month FE	Y	Y
N	386,726	386,726
Adj. R^2	0.077	0.076

Table 6: Market Reaction to Stock Recommendation Changes

This table reports OLS regression results examining whether the market reaction to stock recommendation changes varies with analysts' industry expertise. The dependent variable $CAR_{[-1,+1]}$ is the cumulative market-model-adjusted abnormal return over a three-day event window around the recommendation change date. The key independent variables, analyst-event-specific *Expertise*, proxied by *Knowledge* in Columns (1) and (3) and *Quality* in Columns (2) and (4); see Section 5.4.2 for details. *Knowledge* and *Quality* are standardized to have zero mean and unit variance. Columns (1) and (2) report results for recommendation downgrades, and Columns (3) and (4) for upgrades. All control variables are defined in Section 4 and are either from the prior year or measured using information available prior to the forecast revision date. We exclude observations with stock prices below \$1. All specifications include analyst, firm, and year-month fixed effects. *t*-statistics in parentheses. Standard errors are clustered at the firm level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$CAR_{[-1,+1]}$			
	Rec. Downgrade		Rec. Upgrade	
	(1)	(2)	(3)	(4)
<i>Knowledge</i>	-0.346*** (-2.72)		0.158* (1.71)	
<i>Quality</i>		-0.349*** (-2.70)		0.200** (2.11)
<i>Average forecast error</i>	-0.104 (-1.48)	-0.106 (-1.51)	0.013 (0.26)	0.013 (0.27)
<i>Ln(Market value)</i>	-0.944*** (-2.76)	-0.938*** (-2.74)	-1.017*** (-4.21)	-1.017*** (-4.22)
<i>Institutional ownership</i>	-0.250 (-0.26)	-0.287 (-0.29)	-0.011 (-0.01)	0.008 (0.01)
<i>Ln(No. of analysts)</i>	0.642** (2.01)	0.646** (2.02)	-0.003 (-0.01)	-0.009 (-0.03)
<i>Ln(Brokerage size)</i>	-0.898** (-2.09)	-0.910** (-2.12)	-0.227 (-0.68)	-0.233 (-0.70)
<i>Ln(Analyst career length)</i>	-0.584 (-0.90)	-0.544 (-0.84)	-0.658 (-1.34)	-0.662 (-1.35)
<i>Ln(Time covering company)</i>	-0.257 (-1.45)	-0.260 (-1.47)	-0.123 (-0.83)	-0.125 (-0.84)
<i>BM</i>	-0.119 (-0.22)	-0.109 (-0.20)	0.061 (0.15)	0.049 (0.12)
<i>Lag monthly return</i>	0.108*** (10.15)	0.108*** (10.14)	0.080*** (9.27)	0.080*** (9.27)
<i>Ln(Volatility)</i>	0.696 (1.05)	0.712 (1.07)	1.368*** (3.02)	1.361*** (3.01)
<i>Momentum</i>	0.004 (0.71)	0.004 (0.70)	-0.011** (-2.48)	-0.010** (-2.47)
Analyst FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y
N	12,705	12,705	10,507	10,507
Adj. R^2	0.348	0.348	0.179	0.179

Table 7: Industry Expertise and EPS Forecasts: Post-Knowledge Cutoff Evidence

This table re-estimates the main analyses in Tables 4 and 5 using observations after the LLM’s knowledge cutoff (July 2024 onward) to address potential look-ahead bias. Panel A re-estimates Table 4 (Industry Expertise and EPS Forecast Accuracy), and Panel B re-estimates Table 5 (Market Reaction to EPS Forecast Revisions). *Knowledge* and *Quality* are standardized to have zero mean and unit variance. All regressions include the same control variables and fixed effects as the corresponding main specifications. *t*-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Industry Expertise and EPS Forecast Accuracy				
Dependent Variable	<i>Forecast error</i>			
	(1)	(2)	(3)	(4)
<i>Knowledge</i>	-0.043*** (-2.76)	-0.062*** (-3.59)		
<i>Quality</i>			-0.047*** (-2.96)	-0.056*** (-3.17)
<i>Controls</i>	Y	Y	Y	Y
Analyst FE	Y		Y	
Firm FE	Y		Y	
Analyst-Firm FE		Y		Y
Year-Month FE	Y	Y	Y	Y
<i>N</i>	38,059	35,960	38,059	35,960
Adj. <i>R</i> ²	0.762	0.790	0.762	0.790

Panel B: Market Reaction to EPS Forecast Revisions		
Dependent Variable	<i>CAR</i> _[-1,+1]	
	(1)	(2)
<i>Forecast revision</i> × <i>Knowledge</i>	0.254*** (4.51)	
<i>Knowledge</i>	-0.088 (-1.26)	
<i>Forecast revision</i> × <i>Quality</i>		0.136*** (2.64)
<i>Quality</i>		-0.024 (-0.37)
<i>Forecast revision</i>	0.769*** (8.73)	0.715*** (8.46)
<i>Controls</i>	Y	Y
Analyst FE	Y	Y
Firm FE	Y	Y
Year-Month FE	Y	Y
<i>N</i>	28,303	28,303
Adj. <i>R</i> ²	0.255	0.254

Table 8: Industry Expertise and Industry Recommendations

This table reports the distribution of LLM-extracted industry recommendations across brokerage houses and OLS regression results examining whether recommendation revisions issued by more knowledgeable analysts generate stronger subsequent industry-level returns. Panel A presents the distribution of industry recommendations over the 2012–2025 period. For each brokerage, we report the percentage of optimistic, neutral, and pessimistic recommendations; the final column reports the total number of recommendations issued during the sample period. Panel B presents OLS regression results for return predictability. The dependent variable *SignedBHAR* is the signed buy-and-hold abnormal return for the recommended industry over a 21-trading-day window beginning at the start of the month following the recommendation revision, computed using market-model, Fama–French three-factor, and Carhart four-factor adjusted returns, respectively. The key independent variable is *Knowledge*, standardized to have zero mean and unit variance. All control variables are defined in Section 5.5 and include industry-level, recommendation-level, and analyst-level measures, all computed in the month prior to the revision. All specifications include analyst, industry, and year-month fixed effects. *t*-statistics are reported in parentheses. Standard errors are clustered at the industry level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Distribution of Industry Recommendations				
Brokerage House	Optimistic (%)	Neutral (%)	Pessimistic (%)	<i>N</i>
Morgan Stanley	31.67	61.87	6.45	24,139
Barclays	55.58	43.58	0.83	12,009
Wells Fargo	33.97	54.01	12.03	1,422
Wolfe Research	54.22	31.45	14.33	1,326
Bank of America Global Research	41.17	29.93	28.90	1,166
BMO Capital Markets	20.23	75.76	4.01	1,147
Credit Suisse	39.28	51.62	9.10	802
JP Morgan	44.77	34.05	21.19	793
SunTrust Robinson / Truist Securities	32.16	60.37	7.47	656
Deutsche Bank	35.12	46.03	18.85	541
Scotiabank Gmb	43.45	40.05	16.50	412
Macquarie Research	48.26	27.91	23.84	172
Piper Sandler	53.92	29.41	16.67	102
Total	39.40	54.00	6.53	44,687

Panel B. Return Predictability of Industry Recommendation Revisions			
Dependent Variable	<i>SignedBHAR</i>		
Benchmark	Market-Model (1)	Three-Factor-Model (2)	Four-Factor-Model (3)
<i>Knowledge</i>	0.379** (2.50)	0.376** (2.53)	0.336** (2.30)
<i>Controls</i>	Y	Y	Y
Analyst FE	Y	Y	Y
Industry FE	Y	Y	Y
Year-Month FE	Y	Y	Y
<i>N</i>	2,098	2,098	2,098
Adj. <i>R</i> ²	0.006	-0.016	-0.008

Table 9: Analyst Expertise and Stock Liquidity

This table reports OLS regression results examining whether analysts' industry expertise predicts future stock liquidity. The dependent variables are next-period measures of Amihud illiquidity, effective spread, and realized spread. Columns (1) and (2) report results for Amihud illiquidity, Columns (3) and (4) for effective spread, and Columns (5) and (6) for realized spread. *Expertise* is proxied by *Knowledge* in Columns (1), (3), and (5), and by *Quality* in Columns (2), (4), and (6). *Knowledge* and *Quality* are standardized to have zero mean and unit variance. The key variables of interest are the interactions between the respective expertise measure and firm size. All specifications include firm and year fixed effects. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the firm and year levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	<i>Amihud (t+1)</i>		<i>Effective Spread (t+1)</i>		<i>Realized Spread (t+1)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Knowledge</i>	-0.206*** (-3.49)		-0.285*** (-5.90)		-0.274*** (-4.94)	
<i>Knowledge</i> $\times$ <i>Ln(Size)</i>	0.026*** (3.74)		0.035*** (6.23)		0.034*** (5.19)	
<i>Quality</i>		-0.143** (-2.65)		-0.201*** (-4.76)		-0.196*** (-4.09)
<i>Quality</i> $\times$ <i>Ln(Size)</i>		0.019*** (3.03)		0.025*** (5.12)		0.025*** (4.41)
<i>Ln(Size)</i>	-0.107*** (-3.23)	-0.107*** (-3.20)	-0.127*** (-4.26)	-0.127*** (-4.20)	-0.114*** (-3.28)	-0.114*** (-3.26)
<i>Ln(No. of analysts)</i>	0.040 (1.28)	0.037 (1.20)	0.023 (0.99)	0.021 (0.92)	0.035 (1.34)	0.033 (1.26)
<i>Institutional Ownership</i>	0.013 (0.13)	0.018 (0.17)	-0.078 (-0.75)	-0.072 (-0.66)	-0.032 (-0.27)	-0.026 (-0.22)
<i>Institutional Ownership</i> ²	-0.029 (-0.35)	-0.039 (-0.46)	-0.012 (-0.16)	-0.026 (-0.33)	-0.037 (-0.46)	-0.049 (-0.60)
<i>Return Volatility</i>	-0.485 (-0.65)	-0.478 (-0.64)	3.665*** (3.88)	3.671*** (3.82)	0.985 (1.32)	0.986 (1.28)
<i>ROA</i>	0.115** (2.62)	0.116** (2.60)	0.007 (0.24)	0.009 (0.31)	0.058 (1.41)	0.060 (1.45)
<i>Leverage</i>	-0.018 (-1.42)	-0.019 (-1.50)	0.011 (0.64)	0.009 (0.56)	-0.006 (-0.42)	-0.008 (-0.49)
<i>Market-to-Book</i>	0.001 (1.62)	0.001 (1.67)	0.002** (2.84)	0.002** (2.84)	0.001 (1.44)	0.001 (1.50)
<i>Annual Return</i>	-0.054** (-2.31)	-0.055** (-2.38)	-0.070*** (-3.32)	-0.072*** (-3.41)	-0.032 (-1.69)	-0.034* (-1.80)
<i>Ln(Price)</i>	-0.018 (-0.69)	-0.018 (-0.68)	-0.051 (-1.63)	-0.051 (-1.61)	-0.017 (-0.50)	-0.017 (-0.49)
<i>Ln(Volume)</i>	-1.092** (-2.26)	-1.117** (-2.30)	-2.032*** (-4.42)	-2.066*** (-4.50)	-1.756*** (-3.37)	-1.787*** (-3.44)
<i>Ln(Volume)</i> ²	0.027** (2.25)	0.028** (2.29)	0.050*** (4.31)	0.050*** (4.39)	0.043*** (3.32)	0.044*** (3.38)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	16,284	16,284	16,257	16,257	16,257	16,257
Adj. <i>R</i> ²	0.721	0.720	0.819	0.818	0.732	0.731

Table 10: Analyst Expertise, Price Informativeness, and Market Efficiency

This table reports OLS regression results examining whether analysts' industry expertise predicts future price informativeness and market efficiency. The dependent variables are next-period measures of absolute buy-sell ratio volume, lambda, and the 5-minute variance ratio. Columns (1) and (2) report results for buy-sell ratio volume, Columns (3) and (4) for lambda, and Columns (5) and (6) for the variance ratio. *Expertise* is proxied by *Knowledge* in Columns (1), (3), and (5), and by *Quality* in Columns (2), (4), and (6). *Knowledge* and *Quality* are standardized to have zero mean and unit variance. The key variables of interest are the interactions between the respective expertise measure and firm size. All specifications include firm and year fixed effects. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the firm and year levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	<i>Buy-Sell Ratio (t+1)</i>		<i>Lambda (t+1)</i>		<i>Variance Ratio (t+1)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Knowledge</i>	-0.271*** (-5.78)		-0.278*** (-4.35)		-0.327*** (-5.99)	
<i>Knowledge</i> $\times$ <i>Ln(Size)</i>	0.033*** (5.99)		0.033*** (4.64)		0.041*** (6.22)	
<i>Quality</i>		-0.228*** (-6.14)		-0.186*** (-3.11)		-0.272*** (-6.25)
<i>Quality</i> $\times$ <i>Ln(Size)</i>		0.029*** (6.24)		0.023*** (3.34)		0.035*** (6.51)
<i>Ln(Size)</i>	-0.119*** (-4.10)	-0.120*** (-4.07)	-0.087*** (-3.11)	-0.088*** (-3.10)	-0.123*** (-4.00)	-0.123*** (-3.98)
<i>Ln(No. of analysts)</i>	0.039 (1.70)	0.039 (1.65)	0.063** (2.41)	0.060** (2.37)	0.016 (0.67)	0.016 (0.69)
<i>Institutional Ownership</i>	-0.126 (-1.03)	-0.124 (-1.00)	-0.006 (-0.06)	0.002 (0.02)	-0.009 (-0.08)	-0.007 (-0.06)
<i>Institutional Ownership</i> ²	0.061 (0.73)	0.054 (0.64)	-0.066 (-0.72)	-0.082 (-0.87)	-0.075 (-0.83)	-0.083 (-0.91)
<i>Return Volatility</i>	-0.105 (-0.14)	-0.125 (-0.16)	6.330*** (4.60)	6.328*** (4.55)	-0.450 (-0.55)	-0.462 (-0.55)
<i>ROA</i>	0.122** (2.91)	0.124** (2.95)	0.121* (2.15)	0.123* (2.18)	0.102* (2.04)	0.104* (2.03)
<i>Leverage</i>	-0.001 (-0.07)	-0.001 (-0.08)	0.040* (2.02)	0.038* (1.91)	0.019 (0.82)	0.019 (0.79)
<i>Market-to-Book</i>	0.000 (0.26)	0.000 (0.38)	0.003*** (5.83)	0.003*** (5.64)	0.001* (2.11)	0.001** (2.20)
<i>Annual Return</i>	-0.028 (-1.52)	-0.030 (-1.60)	-0.096*** (-3.46)	-0.098*** (-3.52)	-0.081*** (-3.68)	-0.083*** (-3.76)
<i>Ln(Price)</i>	-0.067* (-2.08)	-0.067* (-2.07)	-0.238*** (-5.28)	-0.238*** (-5.28)	-0.002 (-0.06)	-0.002 (-0.05)
<i>Ln(Volume)</i>	-1.852*** (-6.41)	-1.873*** (-6.53)	1.014*** (4.53)	0.979*** (4.46)	-2.114*** (-5.37)	-2.141*** (-5.44)
<i>Ln(Volume)</i> ²	0.042*** (5.83)	0.042*** (5.94)	-0.029*** (-5.17)	-0.028*** (-5.14)	0.052*** (5.27)	0.053*** (5.34)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	16,257	16,257	16,257	16,257	16,257	16,257
Adj. <i>R</i> ²	0.826	0.826	0.830	0.829	0.698	0.697

Table 11: Analyst Expertise and Implied Cost of Capital

This table reports OLS regression results examining whether implied cost of capital varies with analysts' industry expertise. The dependent variable *ICC* is the average of four implied cost of capital measures following the literature. The key variables of interest are the interaction terms between analyst expertise and the number of analysts covering the firm, where *Expertise* is proxied by *Knowledge* in Column (1) and *Quality* in Column (2). *Knowledge*, *Quality*, and *No. of analysts* are standardized to zero mean and unit variance. All specifications include firm and year fixed effects. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the firm and year levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	<i>ICC</i>	
	(1)	(2)
<i>Knowledge</i>	-0.030*** (-3.27)	
<i>Knowledge</i> $\times$ <i>No. of analysts</i>	-0.022* (-1.94)	
<i>Quality</i>		-0.027*** (-3.11)
<i>Quality</i> $\times$ <i>No. of analysts</i>		-0.018* (-1.98)
<i>No. of analysts</i>	-0.007 (-0.25)	-0.009 (-0.30)
<i>Ln(Size)</i>	-0.300** (-2.69)	-0.294** (-2.64)
<i>Ln(Size)</i> ²	-0.007 (-1.14)	-0.008 (-1.20)
<i>Annual Return</i>	0.049 (1.14)	0.050 (1.17)
<i>Return Volatility</i>	7.270* (2.11)	7.193* (2.07)
<i>Leverage</i>	0.176** (2.23)	0.179** (2.28)
<i>Market-to-Book</i>	0.002 (0.99)	0.002 (1.00)
<i>MAE</i>	0.057* (1.88)	0.057* (1.87)
<i>Ln(Volume)</i>	-0.443 (-0.99)	-0.441 (-0.99)
<i>Ln(Volume)</i> ²	0.016 (1.37)	0.016 (1.37)
<i>Forecast Dispersion</i>	0.053 (1.21)	0.053 (1.20)
<i>Beta</i>	0.092** (2.99)	0.092** (2.99)
<i>Earnings Variability</i>	0.032*** (3.63)	0.032*** (3.63)
<i>LT Earnings Growth</i>	0.131*** (3.88)	0.130*** (3.85)
<i>Institutional Ownership</i>	0.667** (2.51)	0.665** (2.50)
<i>Institutional Ownership</i> ²	-0.745*** (-3.41)	-0.743*** (-3.40)
Firm FE	Y	Y
Year FE	Y	Y
N	10,643	10,643
Adj. <i>R</i> ²	0.652	0.652

Table 12: Analyst Expertise on Market Quality — Evidence from Brokerage Merger

This table reports changes in market quality for firms covered by Credit Suisse analysts, the treatment group, and firms not covered by Credit Suisse, the control group, from the pre-acquisition period January 2022 to the acquisition announcement date of March 19, 2023 and the post-acquisition period from January 2024 to December 2024 due to reductions in Credit Suisse analyst coverage after the acquisition. Panel A reports effects on liquidity using *Effective Spread* and *Realized Spread*. Spread measures are in percentage points. Panel B reports effects on price informativeness using *Variance Ratio* (15-second) and *Lambda* (TSigNSqrtDVol1) multiplied by one million. These measures are from WRDS Millisecond Intraday Indicators. Treatment intensity is constructed using firm-level Credit Suisse analysts' Knowledge and Quality scores for the treatment firms. *Treatment* $\times$ *High* and *Treatment* $\times$ *Low* denote above- and below-threshold exposure based on analyst-month-level scores averaged within firm. We use the 75th percentile of firm-level Knowledge and Quality scores as the threshold. Only firms with at most 10 analysts on average covering during the pre-acquisition period are included. All specifications include the full set of controls, firm fixed effects, and week fixed effects. *t*-statistics are reported in parentheses. Standard errors are two-way clustered at the firm and week levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Market Liquidity				
Dependent Variable	Knowledge		Quality	
	Effective Spread (1)	Realized Spread (2)	Effective Spread (3)	Realized Spread (4)
<i>Post</i> $\times$ <i>Treatment</i> $\times$ <i>High</i>	0.042*** (3.40)	0.024** (2.35)	0.034*** (3.26)	0.019** (2.45)
<i>Post</i> $\times$ <i>Treatment</i> $\times$ <i>Low</i>	0.006 (0.39)	0.000 (0.02)	0.012 (0.72)	0.004 (0.31)
<i>Controls</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
High - Low	0.036	0.024	0.023	0.015
t-stat: High - Low	2.01	1.78	1.40	1.26
p-value: High - Low	0.046	0.078	0.164	0.210
N	81,625	81,623	81,625	81,623
R^2	0.857	0.775	0.857	0.775
Adj. R^2	0.855	0.771	0.855	0.771

Panel B. Price Informativeness and Efficiency				
Dependent Variable	Knowledge		Quality	
	Variance Ratio (1)	Lambda (2)	Variance Ratio (3)	Lambda (4)
<i>Post</i> $\times$ <i>Treatment</i> $\times$ <i>High</i>	0.020*** (3.98)	0.524*** (2.84)	0.021*** (3.83)	0.617*** (3.30)
<i>Post</i> $\times$ <i>Treatment</i> $\times$ <i>Low</i>	0.012*** (2.67)	0.056 (0.22)	0.012*** (2.67)	-0.003 (-0.01)
<i>Controls</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
High - Low	0.008	0.468	0.010	0.646
t-stat: High - Low	1.39	1.70	1.66	2.38
p-value: High - Low	0.166	0.092	0.099	0.019
N	81,625	81,588	81,625	81,588
R^2	0.860	0.681	0.861	0.681
Adj. R^2	0.858	0.675	0.858	0.675

A Appendix

Table IA1: Summary of Institutional Investor Ranking Surveys, 1998–2011

Institutional Investor collected rankings of the importance they assign to sell-side equity research attributes from participants in the All-America Research Team survey. This table shows the results from 1998 to 2011. Special services includes company visits, conferences, etc. Research delivery includes entitlement, technology and customization of buy-side needs. The year 2011 was the final year that we could find the rankings in the print magazines.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Industry knowledge	1	1	1	1	1	1	1	1	1	1	1	1	1	2
Integrity/professionalism					2	2	2	2	3	3	3	2	2	1
Accessibility/responsiveness				2	3	3	3	3	2	2	2	3	3	3
Special services	4	3	2	5	7	6	5	5	5	4	5	5	4	4
Management access				7	5	5	4	4	4	5	4	4	7	6
Servicing	6	4	6											
Written reports	3	2	4	6	8	7	7	6	6	6	6	6	5	5
Timely calls and visits				4	4	4	6	7	7	7	8	8	8	7
Local market knowledge/country knowledge													6	8
Management of conflicts of interest				3	6	8	9	10	10	12				
Financial models			3	8	9	10	10	9	9	9	9	7	9	9
Communication skills					10	9	8	8	8	8				
Idea generation											7	9	10	10
Stock selection	2	5	7	10	11	11	11	11	11	10	12	12		
Earnings estimates	5	6	5	9	12	12	12	12	12	11	10	11	12	12
Quality of sales force	7	7	8	11	13	13	13							
Research delivery											11	10	11	11
Market making	8	8	9	12	14	14	14							
Primary market services			10		15	15	15							

Table IA2: Number of Industry Reports by Brokerage House

The table presents the distribution of industry reports across brokerage houses in the sample.

Brokerage House	<i>No. of Reports</i>	Brokerage House	<i>No. of Reports</i>
Bank of America Global Research	177,292	Canaccord Genuity	8,759
JP Morgan	154,018	Scotiabank Gmb	8,727
UBS Equities	151,538	Kepler Cheuvreux	8,340
Morgan Stanley	121,237	Santander	7,391
Deutsche Bank	78,571	Societe Generale	7,233
Jefferies	66,508	Wedbush Securities	7,199
Credit Suisse	64,776	CIBC Capital Markets	6,058
Barclays	57,201	BNP Paribas	5,475
RBC Capital Markets	40,603	National Bank	5,261
Wells Fargo	33,882	Guggenheim	5,233
Macquarie Research	28,709	Raymond James Ltd.	4,785
BMO Capital Markets	23,605	Cantor Fitzgerald	4,611
Evercore ISI	22,144	SVB Leerink	4,257
Cowen and Company	21,912	Berenberg	3,287
HSBC	20,674	NatWest Markets	1,660
Wolfe Research	16,487	Nordea Markets	1,405
SunTrust Robinson / Truist Securities	14,480	Auerbach Grayson	1,165
Piper Sandler	11,498	CA Cheuvreux	343
Oppenheimer	9,775	Equita SIM	172
Nomura	8,964	Warburg Research Gmb	114

Table IA3: Sample Construction and Representativeness

This table presents the sample construction and t-test results comparing key characteristics of brokerages and analysts within and outside our sample. Panel A summarizes the sample construction steps. The initial sample consists of one-year-ahead EPS forecasts for US common stocks from I/B/E/S, after applying standard data screening criteria. The screening criteria require valid analyst and brokerage identifiers, exclude team forecasts, and require available actual EPS for calculating forecast errors. Panel B compares brokerage and forecast characteristics between brokerages selected for our sample and those excluded. Panel C compares analyst characteristics within the selected brokerages, between analysts who have authored at least one industry report and those who have not. *Mean Diff* is the difference between the two groups. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Sample Construction				
			No. of Forecasts	No. of Analysts
One-year-ahead EPS forecasts for US common stocks from I/B/E/S			1,867,116	7,669
Restrict to 40 brokerages in our sample			968,316	3,794
Restrict to analysts who authored at least one industry report			900,317	3,112
Panel B: Brokerages in Our Sample vs. Excluded Brokerages				
Variable	Selected	Excluded	Mean Diff	t-stat
Forecast error	1.660	1.964	-0.304***	-49.42
Brokerage size	40.097	4.526	35.571***	18.56
No. of firms	415.181	43.467	371.713***	15.79
Average firm size	25.361	34.480	-9.118	-0.82
Panel C: Analysts with vs. without Industry Reports				
Variable	With Reports	No Reports	Mean Diff	t-stat
Forecast error	1.670	1.700	-0.030	-0.30
No. of firms	9.118	5.421	3.697***	11.28
Average firm size	23.323	14.873	8.450***	5.76

Table IA4: Distribution of Industry Recommendations by Industry

This table presents the distribution of industry recommendations extracted from industry reports using LLMs, covering 73 industries defined by GICS 6-digit codes over the period January 2012 to October 2025 (44,687 records in total). The table reports the top 20 industries ranked by total recommendation count, which account for 77% of all industry recommendations.

GICS Code	Industry	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	Total
401010	Banks	354	363	318	172	175	223	245	231	183	237	420	487	463	303	4,174
101020	Oil, Gas & Consumable Fuels	296	452	389	325	256	305	279	193	229	167	153	130	108	124	3,406
351020	Health Care Providers & Services	415	460	438	374	321	241	198	195	191	104	102	110	127	127	3,403
601010	Diversified REITs	221	223	255	225	97	129	134	129	175	200	213	172	132	161	2,466
352010	Biotechnology	70	72	89	51	110	126	159	151	402	422	225	155	147	256	2,435
253010	Hotels, Restaurants & Leisure	103	196	159	114	113	131	201	203	299	145	73	96	107	161	2,101
151010	Chemicals	93	129	104	94	68	80	118	107	144	140	154	131	168	79	1,609
402030	Capital Markets	139	146	184	129	107	87	63	54	70	82	116	149	115	128	1,569
451030	Software	78	82	65	64	88	78	57	69	103	118	165	163	201	169	1,500
403010	Insurance	115	127	113	129	106	108	81	88	118	99	98	49	80	96	1,407
502010	Media	215	158	131	126	106	115	84	81	84	86	27	32	35	24	1,304
201010	Aerospace & Defense	85	50	57	39	50	58	45	61	54	138	165	132	158	119	1,211
501010	Diversified Telecommunication Services	80	98	149	125	107	119	113	76	71	56	71	65	57	24	1,211
352020	Pharmaceuticals	152	179	110	53	58	64	80	53	83	43	73	50	41	20	1,059
203020	Passenger Airlines	93	144	105	35	65	87	73	49	13	32	58	63	83	95	995
351010	Health Care Equipment & Supplies	145	161	121	65	42	58	52	55	117	32	24	22	33	60	987
453010	Semiconductors & Semiconductor Equipment	109	102	66	42	25	46	84	57	45	52	54	62	94	105	943
201060	Machinery	72	66	145	107	58	35	53	60	122	47	32	14	41	43	895
452020	Technology Hardware, Storage & Peripherals	87	89	91	75	56	68	51	43	65	57	59	59	40	30	870
255040	Specialty Retail	136	118	79	49	19	26	40	26	42	66	69	97	34	51	852

Table IA5: Heterogeneous Treatment Effects on Market Quality: Appendix Results

This table reports results for firms with more than 10 analysts on average covering during the pre-acquisition period. Panel A presents liquidity outcomes, and Panel B presents price informativeness outcomes. Treatment intensity is defined using analyst Knowledge and Quality scores. All specifications include the full set of control variables, firm fixed effects, and week fixed effects. *t*-statistics are reported in parentheses. Standard errors are clustered at the firm and week levels. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Market Liquidity				
Dependent Variable	<i>Knowledge</i>		<i>Quality</i>	
	Effective Spread (1)	Realized Spread (2)	Effective Spread (3)	Realized Spread (4)
<i>Post</i> × <i>Treatment</i> × <i>High</i>	-0.001 (-0.19)	-0.005* (-1.81)	-0.002 (-0.46)	-0.005** (-2.04)
<i>Post</i> × <i>Treatment</i> × <i>Low</i>	-0.001 (-0.28)	-0.004 (-1.55)	-0.000 (-0.03)	-0.004 (-1.38)
<i>Controls</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
N	70,085	70,085	70,085	70,085
<i>R</i> ²	0.851	0.566	0.851	0.566
Adj. <i>R</i> ²	0.848	0.560	0.848	0.560
Panel B. Price Informativeness				
Dependent Variable	<i>Knowledge</i>		<i>Quality</i>	
	Variance Ratio (1)	Lambda (2)	Variance Ratio (3)	Lambda (4)
<i>Post</i> × <i>Treatment</i> × <i>High</i>	0.006 (1.21)	0.175 (1.33)	0.003 (0.54)	0.145 (1.13)
<i>Post</i> × <i>Treatment</i> × <i>Low</i>	0.009 (1.63)	0.114 (0.85)	0.012** (2.28)	0.146 (1.06)
<i>Controls</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
N	70,085	70,085	70,085	70,085
<i>R</i> ²	0.849	0.765	0.849	0.765
Adj. <i>R</i> ²	0.847	0.762	0.847	0.762

A.1 Evaluating Industry Reports

A.1.1 Prompts

- **System Prompt:**

You are an expert evaluator of sell-side analyst research reports. Use the following rubric to score reports:

{rubric}

Score each dimension from 1-5 based on the indicators in the rubric.

- **User Prompt:**

Evaluate this analyst report:

Filename: {filename}

Report Content: {full_text}

Output a markdown table with columns: Dimension, Score (1-5), Why (brief evidence from report). Then provide the composite score and quality tier. Nothing else.

- **System Prompt:**

You are an expert evaluator of sell-side analyst research reports. Use the following rubric to score reports:

{rubric}

IMPORTANT: You are evaluating CHUNK {chunk_num} of {total_chunks} of a larger report. Your job is to identify and document evidence for each dimension found in THIS CHUNK. Focus on finding specific evidence: stock picks, price targets, risk sections, data sources, models, forecasts, etc.

- **User Prompt:**

Review this CHUNK of an analyst report and document evidence for each dimension:

Filename: {filename}

This is chunk {chunk_num} of {total_chunks}

(pages {chunk['start_page']}-{chunk['end_page']} of {chunk['total_pages']})

Chunk Content:

{chunk['text']}

For each dimension, list specific evidence found in this chunk. If no evidence, write "None in this chunk". Output format:

```
| Dimension | Evidence from this chunk |
|-----|-----|
```

- **System Prompt:**

You are an expert evaluator of sell-side analyst research reports. Use the following rubric to score reports:

{rubric}

You will receive evidence collected from multiple chunks of a large report. Your job is to synthesize ALL evidence across chunks and assign final scores.

- **User Prompt:**

Based on the combined evidence from all chunks of this report, assign final scores:

Filename: {filename}

{evidence_combined}

Now synthesize the evidence across ALL chunks and score the COMPLETE report. Consider cumulative evidence - e.g., technical expertise shown across multiple sectors in different chunks should score higher than expertise in just one area.

CALIBRATION GUIDANCE:

- Score 5 (Exceptional) should be RARE - only 1-2 dimensions per report at most
- Score 4 (Strong) is appropriate for solid professional work that meets high standards
- Score 3 (Adequate) is the baseline for competent professional work
- Be critical: having evidence doesn't automatically mean a high score - evaluate QUALITY not just presence
- A typical high-quality report should average 3.5-4.0, not 4.5+
- For Originality: educational/overview reports score lower (3) even if they contain some contrarian views; score 5 requires the report's PRIMARY purpose to be a novel thesis
- For Timeliness: scheduled events (teach-ins, conferences) score 3; score 4-5 requires event-driven/breaking analysis
- For Technical Expertise: breadth across multiple sectors (MLPs + drilling + LNG + shipping + refining) warrants 5

Output a markdown table with columns: Dimension, Score (1-5), Why (cite specific evidence from chunks). Then provide the composite score and quality tier. Nothing else.

A.1.2 Rubric

Analyst Report Quality Rubric

18 Scoring Dimensions (1–5 Scale)

Part A. General Analyst Report Quality (10 Dimensions)

#	Dimension	What to Look For
1	Originality	Contrarian view; differentiated perspective relative to consensus
2	Data Richness	Source tier; use of proprietary data
3	Analytical Depth	Explanation of “why” and “so what”; second-order effects
4	Actionability	Specific investment ideas, catalysts, or entry points
5	Structure	Clear thesis and organization; layered presentation for different audiences
6	Forward-Looking	Multi-year perspective; discussion of inflection points
7	Technical Expertise	Depth of domain knowledge; operational and industry nuances
8	Quantitative Rigor	Explicit calculations; use of specific numerical support
9	Risk Awareness	Thesis-specific risks; scenario analysis
10	Timeliness	Responsiveness to events; insight ahead of consensus

Part B. Industry Knowledge (8 Dimensions)

#	Dimension	What to Look For
11	Industry Trends	Industry cycles; secular versus cyclical forces; implications of structural change
12	Industry Technology	Key technologies; technology shifts; effects on costs, productivity, or competition
13	Supply Chains	Upstream and downstream structure; key inputs; capacity constraints or bottlenecks
14	Distribution Models	Go-to-market structure; role of intermediaries; margin or competitive implications
15	Margins	Margin drivers; cost structure; pricing power or competitive pressure
16	Customers	Customer segments; demand drivers; usage patterns; customer economics
17	Labor	Labor availability; wage dynamics; productivity or skill constraints
18	Management Teams	Observable actions; execution outcomes; capital allocation and strategic decisions

Quick Scoring Guide

Five-Point Scoring Guide

Score	Interpretation
5	Exceptional: proprietary evidence or deep industry-mechanism insight; paradigm-shifting analysis
4	Strong: differentiated and well-supported
3	Solid: competent and provides some value added
2	Basic: limited interpretation and modest analytical content
1	Weak: largely descriptive or repetitive, with little original insight

Source Tier Hierarchy

1. Proprietary surveys, channel checks, and expert networks
2. Specialized industry bodies (e.g., IGU, IEA, OPEC, EIA)
3. General data providers (e.g., Bloomberg, Reuters)
4. Company filings and management guidance
5. News articles and press releases

Quality Tiers

Composite Score Quality Tiers

Composite Score	Tier	Example
4.5–5.0	Exceptional	—
3.5–4.4	High Quality	BNP Clean Pulse (4.05), Kepler GTT (3.5)
2.5–3.4	Average	Kepler OFS (3.3), Truist Darlings (2.8), UBS Refiners (2.6)
1.5–2.4	Below Average	Kepler Political (2.3), Truist Data Comp (1.8)
1.0–1.4	Low Quality	—

Key Patterns Identified from 14 Reports

1. Report type predicts the ceiling: Weekly monitor-style reports tend to cap out around 3.0, whereas thematic reports can exceed 4.0.
2. Originality is the main differentiator: Variation in originality explains a substantial share of differences across quality tiers.
3. Quantitative rigor is not the same as insight: Reports with extensive tables but limited interpretation receive relatively low scores.
4. Actionability is distinct from rigor: A report may contain recommendations without strong analysis, or detailed analysis without explicit recommendations.
5. Flash notes can still be high quality: Short reports can score highly when they incorporate proprietary work.
6. Contrarian and quantitative analysis tends to score highest: Testable claims supported by evidence are a common feature of top-scoring reports.
7. Data compilation alone is not research: Reports with little narrative interpretation score poorly regardless of length.
8. Templates improve consistency: Standardized frameworks increase the likelihood that all relevant dimensions are addressed.

Red Flags

- No narrative discussion (tables only)
- Generic or boilerplate risk language
- Recycled standing views presented as event-driven research
- High page count but low informational density
- No source citations
- No specific numerical support or calculations

Green Flags

- Explicit contrarian estimate relative to consensus
- Proprietary data or proprietary modeling
- Bottom-up calculations shown explicitly
- Multi-dimensional analytical framework
- Testable or falsifiable thesis
- Part of a systematic research series