Mental Models and Financial Forecasts*

Francesca Bastianello Paul H. Décaire Marius Guenzel†

July 2, 2025

Abstract

We uncover financial professionals' mental models—the reasoning they use to explain their quantitative forecasts. We organize our analysis around a framework of top-down and bottom-up attention, where analysts endogenously choose both a valuation method and how to allocate attention across variables. Using the near-universe of 2.1 million equity analyst reports, we collect the valuation methods analysts adopt to compute their price targets. To measure attention, we then prompt large language models (LLMs) on a subset of over 300,000 reports to extract 11.8 million lines of reasoning—each combining a topic, valuation channel, time horizon, and sentiment. To validate the reliability of our output, we introduce a multi-step LLM prompting strategy and new diagnostic tools. We document five main findings. (1) Analysts exhibit sparse and rigid mental representations. (2) The choice of valuation methods and topic focus is closely linked. (3) Attention allocation across variables plays a bigger role than valuation methods in explaining both changes in valuations over time and disagreement across analysts. (4) Biases in analysts' forecasts are driven by over-reaction to firm-specific features and under-reaction to macro-related ones. (5) These biases translate into asset price patterns: topics analysts overreact (underreact) to predict lower (higher) realized returns.

Keywords: Mental Models, Professional Forecasters, Large Language Models, Belief Formation, Attention

*First Version: October 30, 2024. This Version: July 2, 2025. All errors and omissions are our own.

†Bastianello is at the University of Chicago, Booth School of Business. Décaire is at the W. P. Carey School of Business at Arizona State University. Guenzel is at the Wharton School of the University of Pennsylvania. We thank Federico Bastianello, Josefina Cenzon, Ricardo De La O, Olivier Dessaint, Diego Garcia, Stefano Giglio, Sam Hanson, Alex Imas, Chad Kendall, Spencer Kwon, Will Mullins, Nick Roussanov, Josh Schwartzstein, Andrei Shleifer, Adi Sunderam, Sean Wu, and seminar and conference participants at the W. P. Carey School of Business at Arizona State University, Harvard Business School, McGill University Desautels Faculty of Management, the 2024 ASU-UA Junior Conference, the 1st Workshop on LLMs and Generative AI for Finance, the 14th ITAM Finance Conference, the 2025 Adam Smith Workshop, the MIT Sloan Junior Finance Conference, the NBER Asset Pricing Workshop, the 2025 FSU Truist Beach Conference, the 2025 University of Kentucky Finance Conference, the 2025 Western Finance Association Meetings, the SFS Cavalcade North America 2025, the 7th Future of Financial Information Conference, and the First Annual Isenberg School of Management Finance Conference for useful comments. We thank Rasoul Foroughfard, Daniil Garbuzov, Aria Liu, Biwon Lee, and Shuman Zhang for their exceptional research assistance. Bastianello gratefully acknowledges support from the Liew Fama Junior Faculty Fellowship, the Fama Faculty Fellowship, as well as the Fama-Miller Center and the Center For Applied AI Research at the University of Chicago Booth School of Business. Décaire gratefully acknowledges financial support from the Center for Responsible Investing at Arizona State University. Guenzel gratefully acknowledges financial support from the Jacobs Levy Equity Management Center for Quantitative Financial Research. Contact information: francesca.bastianello@chicagobooth.edu, paul.decaire@asu.edu, and mguenzel@wharton.upenn.edu.

1 Introduction

How do people form expectations in financial markets? According to the rational expectations benchmark, individuals use all available information to produce statistically optimal forecasts. Initially, this view was challenged with largely indirect evidence, and with the documentation of anomalies in financial markets that were difficult to reconcile with fully rational expectations.[1] Then, over the last few decades, progress accelerated with the growing availability of survey data, which allowed researchers to directly measure individuals' quantitative forecasts. However, the process through which people arrive to such forecasts still remains an open question, at least in part due to limited measurement. To better understand how market participants mentally represent firms, and how these representations might ultimately drive investor behavior, this paper goes a step further by using sell-side analysts' written reports to measure the *reasoning* behind their forecasts.

Specifically, we study how analysts reason when valuing stocks. Firm valuation is an inherently hard problem for several reasons. First, analysts operate in a high-dimensional environment, and, given limited attention, they must engage in dimensionality-reduction to choose which variables to attend to and which to neglect. Second, valuation is computationally complex. Applying the present value relationship exactly—that prices equal the present value of all expected future dividends—requires forecasting an infinite stream of future cash-flows and discount rates. In practice, however, this is not how analysts apply this formula. Instead, they rely on simplifying assumptions that give rise to different valuation methods.

To formalize these processes, we define a mental model to be a combination of attention weights and valuation methods, and study how analysts tackle the dual challenge of dimensionality reduction and computational tractability by exploring three sets of questions. First, how do analysts allocate attention and choose valuation methods? Second, how do these choices shape their quantitative forecasts? Third, how do analysts' mental models lead to biases and observed asset pricing anomalies?

[1] See De Bondt and Thaler (1990), Vissing-Jorgensen (2003), Bacchetta et al. (2009), Malmendier and Nagel (2011), Case et al. (2012), Amromin and Sharpe (2014), Greenwood and Shleifer (2014), Bordalo et al. (2020b), Bordalo et al. (2020a), Giglio et al. (2021), Nagel and Xu (2022) among others, and Adam and Nagel (2023) for a review.

To answer these questions, we begin by introducing a conceptual framework of top-down and bottom-up attention, where analysts endogenously choose both attention allocation and valuation methods based on the subjective relevance of variables and cost of acquiring information about them. To measure attention and valuation methods, we then analyze the written text from the near-universe of 2.1 million equity reports from analysts at 43 brokerage houses over a 26-year period. We then prompt large language models (LLMs) to extract valuation methods and price targets from all reports, and an additional 11.8 million separate lines of reasoning from the text portions of a subset of over 300,000 reports. Our resulting dataset links analysts' mental models to their quantitative forecasts for a sample with firm coverage and characteristics that closely aligns with commercial databases, such as the Institutional Brokers' Estimate System (IBES).

Building on our model and the novel dataset we construct, we document five sets of facts. First, analysts' mental models are sparse and rigid, and shaped both by top-down and bottom-up factors. Second, analysts' attention allocation is tightly linked to their choice of valuation method. Third, analysts' mental models translate into their quantitative forecasts and affect both changes in valuation over time and disagreement across analysts. While variation in valuation methods and in attention both contribute to valuation changes and disagreement, differences in attention account for the largest share of the variation. Fourth, analysts' price targets are biased and reflect overreaction to firm-related topics and underreaction to macro-related ones. Finally, these biases correlate with key asset pricing patterns. Topics which analysts overreact (underreact) to are associated with lower (higher) realized returns.

Conceptual framework. To lay the foundations for our analysis, we start by introducing a conceptual framework with a model of top-down and bottom-up attention. In the model, the true value of a stock is determined by many variables (e.g., accounting metrics or intangible factors) that differ in how relevant they are for valuation and in how costly they are to process. The model has two key components. First, analysts do not directly observe these variables, but instead receive noisy signals about them. The more attention analysts pay to a given variable, the more precise the signal they obtain. Second, analysts have access to a set of valuation methods, each placing different weights on these variables. For example, we may

think of discounted cash flow (DCF) models as assigning greater weight to long-run outcomes, while more reduced-form multiples approaches may emphasize variables associated with relevant comparables. Analysts then subjectively choose valuation methods and attention weights to minimize mean squared forecast error, subject to a budget constraint on attention.

This framework yields a series of testable predictions. First, analysts should allocate more attention to variables that they perceive to be more relevant, more volatile, and less costly to acquire information about. Second, analysts face a bias-variance trade-off. The optimal valuation method minimizes processing costs as well as discrepancies along variables that are perceived to be important. Third, analysts' attention allocation and their choice of valuation methods are tightly interlinked. Specifically, analysts should tilt their attention towards variables that are weighed more in their chosen valuation method, while simultaneously selecting valuation methods they perceive to be best at processing these variables. Finally, changes in asset valuations and disagreement in analysts' forecasts may arise from differences in valuation methods or differences in attention allocation. Our empirical investigation tests these predictions and assesses the empirical relevance of these channels, while also exploring their resulting asset pricing implications.

Data collection and methodological contribution. The challenge with testing these models is that we typically do not observe what information market participants attend to, nor do we observe how they simplify their valuation task. To overcome these challenges, we leverage the fact that sell-side equity analysts not only publish quantitative forecasts—which have been widely studied in empirical asset pricing—but also produce accompanying reports in which they explain the key arguments underlying those forecasts.

To test the predictions of our model, we use large language models (LLMs) to collect information from the near-universe of 2.1 million equity analyst reports. To structure our data-collection effort, we then define a mental model as a combination of a valuation method (the formula analysts use to compute their price target, e.g., DCF, multiples), and a series of lines of reasoning, each with a topic (e.g., customer demand, brand, inflation), valuation channel (e.g., whether it impacts sales, costs, margins, discount rates, etc), time outlook (past, present, near-future, distant-future), and sentiment (positive, negative, neutral).

In constructing our dataset, we make two methodological contributions that ensure the

output is comprehensive, stable, and reliable.

First, we show that when applied to lengthier, multi-page documents—such as equity reports—naïve prompting strategies that ask LLMs to extract information in a single step often yield incomplete outputs. This limitation arises in part from LLMs' tendency to focus disproportionately on the beginning of the input. To address this, we introduce a refined LLM approach in which we *iteratively* provide sub-segments of each report. After processing all segments using this "chunking" method, we then present the LLM with both the full document and the extracted content in a final step that validates the selected topics. This multi-step prompting strategy substantially outperforms the single-step approach in both the number and comprehensiveness of individual lines of reasoning extracted.

Second, and relatedly, we introduce diagnostic tools akin to bootstrapping to show that the output of our multi-step prompt is highly stable. To do so, we repeatedly feed a subset of reports to the LLM using both the naïve and the multi-step prompting strategies and calculate the likelihood that a given argument is extracted across multiple runs. Unlike naïve one-step approaches, the output of our multi-step approach is highly stable.

Sparse and rigid mental representations. Leveraging our dataset, and guided by our theoretical framework, we turn to the key results of our paper. The first fact that we establish is that analysts' mental representations are sparse and rigid. In the average report, analysts attend to 16 out of 139 topics, illustrating how their attention allocation in indeed selective.[2] At an aggregate level, analysts predominantly focus on firm-specific topics (75–80%) and top-line items (40–50% of arguments), and their arguments are mostly forward-looking (55-60%). Moreover, sentiment and attention to top-line items is highly pro-cyclical, while attention to discount rates and macro-related topics is counter-cyclical.

At a finer level, we are able to confirm that, consistent with the first prediction of our model, analysts pay more attention to variables that are more relevant, volatile, and easily available. To illustrate the role of relevance and volatility, we identify three types of topics. First, there are topics that are relevant and volatile, and therefore receive persistent attention over time. These include fundamental drivers of cash flow generation, such as customer

[2] We present a series of tests to show that topics omitted from textual discussions are outside forecasters' mental models, rather than missed by the AI or excluded due to analyst incentives or space constraints.

demand, market share, and cost structure. Second, there are topics that are relevant but mostly stable, and therefore only receive heightened attention when they significantly deviate from their steady state levels. These include topics such as public health, taxation, inflation, or tariffs. Third, there are emerging topics, whose true or perceived relevance rises over time. These include environmental policies and ESG and corporate social responsabilities (CSR), as well as new categories like AI and cognitive applications, automation and robotics, or cybersecurity and data protection. While these three types of topics illustrate the importance of variable relevance and volatility in shaping attention allocation, we next exploit cross-sectional differences across analysts to provide evidence on the role of a variable's availability. We show that greater proximity to a company's headquarters leads analysts to allocate more attention to firm-related topics at the expense of macro-related topics, and that greater exposure to high and volatile inflation leads analysts to allocate more attention to inflation, even controlling for firm-year fixed effects.

Next, we explore the choice of valuation methods, and find that 43% of reports use discounted cash flow (DCF) methods, 32% use price-to-earnings (P/E) multiples, and 62% use at least one multiples-based approach.[3] This heterogeneity reflects both analysts' attempts to tailor their model to a given firm, and analyst-specific factors, such as their training. To illustrate how feature relevance impacts the choice of valuation methods, we exploit differences in adopted methods across industries and in the cross-section. For example, DCF methods (structured to emphasize future cash flows) are, relatively speaking, more commonly used for small, young, and growth firms, even with analyst-firm fixed effects. In contrast, multiples-based approaches are more frequently applied to large, mature, and value firms.[4] To show how analyst-specific features also impact the choice of valuation method, we exploit differences in training across continents. For example, we show that analysts located in Europe are much more likely to use DCF methods than analysts located in North America, even when covering the same firm at the same point in time. This suggests that analysts approach valuing a firm with a default method, and only partially adjust that choice across

[3] Our valuation method prompt extracts all methods used to derive the price target forecast within a report. Since analysts often use a blend of methods, the shares described above sum to more than 100%.

[4] Reflecting these differences, we also show that valuation method usage varies intuitively across industries—for example, DCFs are relatively more common in information and professional services, whereas multiples approaches are more prevalent in retail.

firms and over time (Bastianello and Imas, 2025).

Consistent with this intuition, we show that analysts' mental representations are rigid both over time and in the cross-section of stocks they cover, meaning that they do not adapt their mental models enough to reflect changes in fundamentals. Despite such rigidity, analysts are more likely to update their models following large forecast errors.

Valuation methods and attention are tightly interlinked. The next prediction of our model is that the choice of valuation method and attention allocation are tightly interlinked. Consistent with this, we find that using the same valuation method is associated with significantly more similar topic focus, as well as greater joint congruence in topic selection, valuation channel, time outlook, and sentiment.

Moreover, analysts allocate attention to topics that are differentially emphasized by their chosen valuation method. For example, even when covering the same firm at the same point in time, analysts using discounted cash flow (DCF) methods tend to adopt a long-run perspective, and emphasize topics such as discount rates, innovation, and corporate investment; consistent with the logic of projecting future cash flows and discounting them back to present value. In contrast, analysts using multiples-based approaches adopt a shorter-term focus, and place greater weight on product- and customer-related topics that are more salient for comparables-based valuation.

Changes in valuation over time and disagreement. Having established some of the key properties of analysts' mental models, we study how our measured reasoning translates into their quantitative forecasts. To do so, we explore how analysts' mental models affect changes in valuations over time and disagreement. In line with the model's prediction, both differences in attention and differences in valuation methods contribute to valuation changes and disagreement. Empirically, however, differences in attention play a quantitatively larger role.

Starting from changes in valuation over time, analysts exhibit a 40% overlap in topics covered for a given firm from one year to the next, and use the same valuation method 82% of the time. Consistent with this, changes in attention and in how analysts mentally represent firms over time contribute more to time-varying valuations than do changes in how analysts price different features (64% vs. 36%). The greater importance of attention weights than

valuation weigths holds both in the aggregate and at the feature level, with two notable exceptions: innovation/R&D and environment. For these two topics, differences in how analysts have priced them over time contribute more to valuation changes than differences in their fundamentals or analysts' attention allocation. This is consistent with both variables having been associated with potential mispricing.

Turning to disagreement, when considering two analysts covering the same firm at the same point in time, they share only 33% of the combined topics used, while they have some overlap in valuation methods used 81% of the time. Consistent with these patterns, we find that differences in attention weights contribute substantially more to total disagreement than do differences in valuation methods (83% vs. 17%). Two topics for which differences in valuation weights remain statistically significant are innovation/R&D and housing markets, both of which have previously been associated with displacements, bubbles and crashes (Kindleberger, 1978).

Overall, this analysis allows us to make the drivers of valuation changes and disagreement interpretable, and it also suggests that analysts converge more on how they handle computational complexity (which algorithm/valuation method to use), than in how they resolve representational complexity, such as deciding which aspects of the information environment to attend to most (Ba et al., 2022, Bordalo et al., 2024a).

Biased beliefs and return predictability. Having established that our measured reasoning translates into analysts' quantitative forecasts, in the final part of the paper we explore what features analysts over and underreact to, and the resulting asset pricing implications of such biases. At the aggregate level, our measure of sentiment correlates with Shiller's CAPE ratio in the time-series, and correlates with patterns of realized returns for the main asset pricing factors in the cross-section. However, our data allow us to move beyond sentiment by identifying which specific features analysts over- or underreact to, and how that influences return predictability. We find that analysts overreact to firm-related topics and underreact to macro-related ones. Moreover, the features that analysts overreact to are associated with lower realized returns, while features that analysts udnerreact to are associated with higher realized returns. This highlights the potential of the granularity of our measured mental models to deepen our understanding of market expectations, return predictability,

and mispricing in an interpretable way.

Literature. Our paper contributes to three strands of literature. First, we contribute to the literature in empirical asset pricing that has studied the properties and implications of investors' potentially biased beliefs. This includes papers that have analyzed market participants' quantitative forecasts (Greenwood and Shleifer, 2014, Giglio and Kelly, 2018, Bouchaud et al., 2019, Ma et al., 2020, Giglio et al., 2021, Nagel and Xu, 2022, De La O and Myers, 2021, Bordalo et al., 2019, 2024b, Afrouzi et al., 2023, Ben-David and Chinco, 2024, Decaire and Graham, 2024, Bastianello, 2025, de Silva and Thesmar, 2024, de Silva et al., 2025, and Decaire and Guenzel, 2025), an emerging literature at the intersection of machine learning and finance which often uses text data to study beliefs (Asquith et al., 2005, Giglio et al., 2022, Bybee, 2023, Gormsen and Huber, 2023, Van Binsbergen et al., 2023, Gabaix et al., 2023, Bianchi et al., 2024, Bybee et al., 2024, Charles and Sui, 2024, Chen, 2024, Lopez-Lira and Tang, 2024, Decaire and Graham, 2024, Decaire et al., 2024, Decaire and Guenzel, 2025, Sarkar, 2025), as well as a literature in accounting that has documented variation in the usage of valuation methods and their accuracy (Bradshaw, 2002, Demirakos et al., 2004, Imam et al., 2013, Gleason et al., 2013, Erkilet et al., 2022). While these earlier papers study market participants' beliefs either through their quantitative forecasts or through text data alone, we add to this line of research by exploiting a setting that allows us to create a new dataset that *links* the text data in sell-side analysts' equity reports to their subjective quantitative beliefs. This allows us to study the reasoning behind their forecasts. In contemporaneous work, Chen et al. (2025) leverage equity reports to study buy-side analyst recommendations, and De Rosa (2024) and Ke (2025) link sell-side equity reports to IBES forecast data to study memory and attention, respectively.

Second, we contribute to the literature on mental models. A growing body of research uses open-ended and structured survey questions to capture the thought process in people's reasoning in a variety of different fields, including macroeconomics and finance (Bailey et al., 2019, Andre et al., 2022, Chinco et al., 2022, Chopra and Haaland, 2023, Andre et al., 2024a, Andre et al., 2024b, Bauer et al., 2024, Binetti et al., 2024, Stantcheva, 2024, Laudenbach et al., 2024).[5] We contribute to this work by considering a high-stakes field setting, which

[5] See Haaland et al. (2024) for a recent survey of this literature.

allows us to exploit two key advantages.[6] First, we are able to link the reasoning and quantitative forecasts of financial professionals in their own natural habitat using output that is part of their day-to-day job. Second, since we do not have to elicit information on people's reasoning via additional surveys, we have access to a whole time-series and cross-section, rather than having to rely on the snapshot in time when the survey was run.

Third, we contribute to an active body of work in behavioral economics that studies how people represent problems in potentially high-dimensional environments, and how they allocate their limited attention (Woodford, 2001, Sims, 2003, Maćkowiak and Wiederholt, 2009, Van Nieuwerburgh and Veldkamp, 2009, Van Nieuwerburgh and Veldkamp, 2010, Veldkamp, 2011, Gabaix, 2014, Hanna et al., 2014, Schwartzstein, 2014, Kacperczyk et al., 2016, Gagnon-Bartsch et al., 2018, Gabaix, 2019, Kohlhas and Walther, 2021, Fan et al., 2021, Schwartzstein and Sunderam, 2021, Ba et al., 2022, Bordalo et al., 2023, Link et al., 2023, Bordalo et al., 2024a, Charles and Kendall, 2024, Flynn and Sastry, 2024, Bastianello and Imas, 2025). One of the challenges of testing these theories is a lack of data on what information people are in fact attending to. Our paper contributes to this literature by exploiting a field setting that enables large-scale measurement of the components of attention and information processing.

Organization of the paper. The rest of the paper proceeds as follows. Section 2 introduces a conceptual framework to guide the empirical analysis. Section 3 introduces the data, presents key institutional features, and discusses our LLM-based mental model extraction method. Section 4 presents aggregate facts about analysts' mental models. Section 5 shows how analysts' mental models are sparse and rigid. Section 6 studies how analysts' mental models translate to their quantitative forecasts by analyzing by linking analysts' reasoning to changes in valuation, and forecast disagreement. Section 7 considers the sources of analysts' biased beliefs and resulting asset pricing implications. Section 8 discusses how our methodology can be applied to study other components of analysts' mental models and concludes.

[6] An early mostly theoretical exploration in which analysts use simplified reasoning includes Hong et al. (2007).

2 Conceptual Framework

The first challenge in analyzing the text of equity analysts' reports is to organize and structure this data in a way that enables a systematic examination. In deciding what data to collect about analysts' reasoning, we start from the canonical present value relation that prices reflect the present value of expected future dividends, and focus on how analysts apply this formula:

$$P_t = \mathbb{E}_t \left[\sum_{j=0}^{\infty} \frac{D_{t+1+j}}{1 + R_{t+1+j}} \right] \tag{1}$$

Applying this formula requires making two high level choices. First, since analysts typically do not forecast an infinite stream of future dividends and discount rates, they must instead choose an appropriate approximation—in other words, they must select a valuation method. A common example is a discounted cash flow (DCF) model, which projects and discounts expected cash flows over a finite horizon, before assuming a constant terminal growth rate. Alternatively, analysts may rely on a more reduced-form, multiples-based approach, inferring value by comparing the firm's metrics (e.g., earnings or sales) to those of comparable peers or its own historical trends. Second, once a particular valuation method is chosen, analysts must decide which features to pay attention to when estimating its relevant inputs. Analyzing the present value relationship in equation (1) reveals three key dimensions that analysts may reason through when determining the impact of a given piece of news on valuation: the valuation channel through which it operates (cash flows or discount rates), the time horizon it affects, and its directional impact (sentiment).

These aspects of analysts' reasoning then pin down how analysts draw relationships across variables. Motivated by this thinking, we define a mental model as a combination of i) a valuation method and ii) a set of attention weights on different inputs, where each input is characterized by a topic, along with its associated valuation channel, time outlook, and sentiment. We discuss our data collection efforts of these features in detail in Section 3, and we now turn to introducing a conceptual framework that illustrates how analysts might choose valuation methods and allocate attention across features when faced with limited cognitive resources (Veldkamp, 2011, Gabaix, 2019).

2.1 Setup

To model the core choices analysts make with a minimal working example, we recast the present-value relationship from equation (1) in reduced-form, and model prices as a linear function of different variables:

$$p \;=\; \sum_{k=1}^{K} v_k \, x_k \tag{2}$$

where v_k capture valuation weights, and x_k relate to relevant features that impact valuation. These may include both standard financial metrics (e.g., cash-flow forecasts or discount-rate proxies) as well as harder-to-measure components (e.g., brand strength or regulation).

Given this environment, we introduce the two key elements that analyst ought to choose: valuation methods, and attention allocation across variables. Starting from the first component, we assume that when valuing a firm analysts can choose from a set of valuation methods $m \in M$. Each method m assigns a distinct set of weights $\{m_k\}_{k=1}^{K}$ to the relevant variables $\{x_k\}_{k=1}^{K}$. Formally, the price target under method m is given by:

$$p^m = \sum_{k=1}^{K} m_k \, x_k, \tag{3}$$

where m_k indicates the weight that method m places on variable x_k. Importantly, we think of the weight that method m places on variable x_k as a reduced form way of also capturing how easy a variable is to process under the chosen valuation method. In other words, a higher m_k can implicitly reflect the ability of valuation method m to create a simpler mapping between variable x_k and the resulting price target. For example, by explicitly modeling long term growth rates, DCF methods generate a simpler mapping between price targets and long term growth than multiples-based approaches.

Definition 1 (Valuation Method)**.** *Let $\{x_k\}_{k=1}^{K}$ be the relevant input variables for valuing a firm. A* valuation method *m is determined by a vector of weights $\{m_k\}_{k=1}^{K}$, where each $m_k \in \mathbb{R}$ determines the weight of the corresponding variable x_k in valuation. Under valuation method m the price target is then given by: $p^m = \sum_{k=1}^{K} m_k x_k$.*

To allow for the role of attention in this framework, we assume that analysts do not observe each feature x_k (e.g., they do not observe future cash-flows or discount-rates precisely).

Instead each feature has a prior distribution $x_k \sim N(\mu_k, \tau_{0k}^{-1})$, where μ_k is the prior mean and τ_{0k}^{-1} is the prior variance. Analysts then receive noisy signals:

$$s_k = x_k + u_k, \quad \text{where} \quad u_k \sim N(0, \tau_{sk}^{-1}) \tag{4}$$

where τ_{sk} is the precision of signal s_k. Upon receiving each signal, analysts update their beliefs via Bayesian updating, such that their posterior beliefs are given by: $\mathbb{E}[x_k|s_k] = \mu_k + \frac{\tau_{sk}}{\tau_{sk}+\tau_{0k}}(s_k - \mu_k)$ and $\mathbb{V}[x_k|s_k] = (\tau_{sk} + \tau_{0k})^{-1}$, and the resulting price target is:[7]

$$\mathbb{E}[p^m|\{s_k\}_{k=1}^K] = \sum_{k=1}^K m_k \mathbb{E}[x_k|s_k] = \sum_{k=1}^K m_k(s - \mu_k)a_k + m_k\mu_k \tag{5}$$

where $a \equiv {}^{\tau_{sk}}/_{(\tau_{sk}+\tau_{0k})}$. Finally, we allow analysts to choose the precision of each signal (τ_{sk}) by allocating attention across variables, subject to a linear budget constraint: $\sum_{k=1}^K c_k\tau_{sk} \leq C$, where c_k denotes the marginal cost of devoting one more unit of increasing attention to feature x_k, and C is the total attention budget. More generally, we think of c_k as also capturing bottom-up factors and prior experiences which make a variable more prominent or salient, more easily available, or easier to process.

Definition 2 (Attention)**.** *Analysts* attention *is captured by the precision parameters* $\{\tau_{sk}\}_{k=1}^K$ *of the signals they acquire about each feature* $\{x_k\}_{k=1}^K$*. Allocating more attention to a particular features* x_k *increases* τ_{sk}*, resulting in a more precise signal for that feature.*

Having specified how both valuation methods and attention weights enter our model, we now define a mental model as combination of these two components.

Definition 3 (Mental Model)**.** *A mental model* $\mathcal{M}(m, \tau_s)$ *is defined by the combination of (i) a valuation method* $m = \{m_k\}_{k=1}^K$*, and (ii) a vector of attention weights* $\tau_s = \{\tau_{sk}\}_{k=1}^K$*.*

Given this framework, we are interested in understanding how analysts should optimally choose a valuation method and allocate attention across features. To do so, we first establish a benchmark in which analysts choose attention weights under the true valuation function

[7] We assume that all features x_k are uncorrelated with one another, so that signal s_k is only relevant for forecasting feature x_k.

(i.e., by setting $\{m_k\}_{k=1}^K = \{v_k\}_{k=1}^K$). We then consider how the adoption of alternative valuation methods leads to deviations from this benchmark.

2.2 Attention Allocation Across Variables Under the "True" Valuation Model

To gain intuition in the simplest possible setting, we focus on the case with just two features. Moreover, in this section we assume that analysts use the true valuation function as their valuation method, i.e., $\{m_k\}_{k=1}^K = \{v_k\}_{k=1}^K$. Analysts then choose how much attention to allocate to each feature by minimizing mean squared forecast errors, subject to their linear budget constraint on attention:[8]

$$\min_{\tau_{s1},\tau_{s2}} \quad \mathbb{E}\left[(p - \mathbb{E}[p|s_1, s_2])^2\right] \qquad s.t. \qquad c_1\tau_{s1} + c_2\tau_{s2} \leq C, \quad \tau_{s1} \geq 0, \quad \tau_{s2} \geq 0 \tag{6}$$

Since analysts in our model perform rational Bayesian updating and use the true valuation function for valuation, their forecasts are unbiased on average. Therefore, the problem is equivalent to minimizing their posterior variance:

$$\min_{\tau_{s1},\tau_{s2}} \quad \mathbb{V}[p|s_1, s_2] = v_1^2\left(\tau_{01} + \tau_{s1}\right)^{-1} + v_2^2\left(\tau_{02} + \tau_{s2}\right)^{-1} \tag{7}$$

$$s.t. \qquad c_1\tau_{s1} + c_2\tau_{s2} \leq C, \quad \tau_{s1} \geq 0, \quad \tau_{s2} \geq 0 \tag{8}$$

Setting up the Lagrangian and taking first order conditions:

$$v_1^2(\tau_{01} + \tau_{s1})^{-2} = \lambda c_1 \tag{9}$$

$$v_2^2(\tau_{02} + \tau_{s2})^{-2} = \lambda c_2 \tag{10}$$

which equates the marginal benefit of acquiring an additional unit of information with its marginal cost. From this simple expression, it is clear that the marginal benefit of acquiring information about x_k is increasing in v_k (its relevance for valuation), and decreasing in τ_{0k} (its volatility).

[8] Analysts have incentives to focus their efforts on producing informative valuation theses and credible price targets and recommendation (see Section 3.1).

Proposition 1 (Allocation of Attention Across Features)**.** *Ceteris paribus, an analyst will allocate more attention to feature x_k when it has a greater weight in the price function (higher v_k), is more volatile (lower τ_{0k}), and is less costly to acquire information about (lower c_k).*

Proof. All proofs are in Appendix A, unless otherwise stated. □

These comparative statics help us understand why analysts may allocate differential attention to features both in the cross-section and over time. Proposition 1 implies that if two industries or stocks differ in how relevant, volatile or costly different variables are (v_k, τ_{0k}, c_k), then analysts will allocate attention differently when valuing them. For example, brand perception may be highly relevant in consumer-goods or tech-platform firms, whereas regulated utilities may demand more focus on policy changes or regulatory filings. Similarly, pharmaceutical companies often exhibit greater volatility in research outcomes, prompting analysts to devote more resources to assessing clinical data than they would for stable consumer-staples companies.[9] Combining the two first order conditions and using the budget constraint we then obtain:

$$\tau_{s1}^* = \frac{C + c_2\tau_{02} - \frac{v_2}{v_1}\sqrt{c_1}\sqrt{c_2}\tau_{01}}{c_1 + \frac{v_2}{v_1}\sqrt{c_1}\sqrt{c_2}}. \tag{11}$$

and $\tau_{s2}^* = {}^{(C\,-\,c_1\tau_{s1}^*)}/_{c_2}$. If either of the interior solutions is negative, analysts optimally set the corresponding precision to zero, i.e., they pay no attention to that feature, and allocate the full attention budget to the remaining variable.

Proposition 2 (Sparsity)**.** *If the marginal benefit from acquiring the first unit of information is lower than its marginal cost ($v_k^2(\tau_{0k})^{-2} < \lambda c_k$), then analysts will pay no attention to that variable, setting $\tau_{sk} = 0$ and $\tau_{sj} = {}^{C}/_{c_j}$. A variable is more likely to be neglected if it is less relevant for valuation (low v_k), more stable over time (high τ_{0k}) or has a higher marginal cost (high c_k).*

Thus, analysts allocate attention only to variables with a sufficiently high marginal value of information. Variables that are less important for valuation, highly stable, or costly to learn about are more likely to be ignored, leading to sparse mental representations.

[9] Moreover, notice how the expression in (11) shows how attention across all features may change even when the volatility of a single variable changes.

2.3 Endogenous Choice of Valuation Method

Thus far, we have assumed that analysts use the true valuation function when valuing firms (i.e., we imposed that $\{m_k\}_{k=1}^K = \{v_k\}_{k=1}^K$). In practice, however, analysts often have access to a range of valuation methods that make simplifying assumptions and systematically assign more weight to some variables over others. For example, a DCF model tends to emphasize projected cash-flows and discount rates, whereas a multiples-based approach may lean more on peer performance or recent trends. In this section, we analyze how analysts should choose among (potentially misspecified) valuation methods, and how these choices affect their allocation of attention across variables.

Specifically, we assume that analysts know the true valuation model in (2), and that they have access to a range of different valuation methods $m = \{m_k\}_{k=1}^K$.[10] Analysts then choose $\mathcal{M}(m, \tau_s)$ to minimize the mean squared forecast error:

$$\mathbb{E}\left[(p - \mathbb{E}[p^m|s_1, s_2])^2\right] = \mathbb{V}\left[p - \mathbb{E}[p^m|s_1, s_2]\right] + \mathbb{E}\left[p - \mathbb{E}[p^m|s_1, s_2]\right]^2 \tag{12}$$

$$= \sum_{k=1,2} \left(\frac{v_k^2 - (v_k - m_k)^2}{\tau_{0k} + \tau_{sk}} + \frac{(v_j - m_j)^2}{\tau_{0j}} \right) + \left(\sum_{j=1,2} (v_j - m_j)\mu_j \right)^2 \tag{13}$$

When analysts use the true model ($m_k = v_k$), this expression reduces to the objective function in (7). Instead, adopting an alternative mental model ($m_k \neq v_k$) leads to changes in both the variance and bias components of the total mean squared error.

To solve the model, we proceed in two steps. First, we compute the optimal attention allocation τ_s for a given valuation method $m = \{m_k\}_{k=1}^K$. Second, we solve for the valuation method that minimizes mean squared errors under these optimal attention weights. Starting from the first step, we notice that attention weights τ_{sk} appear only in the first term of the objective function in (13). Therefore, analysts choose their attention allocation to minimize:

$$\min_{\tau_{s1}, \tau_{s2}} \quad \frac{v_1^2 - (v_1 - m_1)^2}{\tau_{01} + \tau_{s1}} + \frac{v_2^2 - (v_2 - m_2)^2}{\tau_{02} + \tau_{s2}} \qquad s.t. \qquad c_1\tau_{s1} + c_2\tau_{s2} = C \tag{14}$$

[10] The fact that analysts know the true valuation model simply reflects the fact that analysts often have a sense of the relevant importance of different variables in valuing a firm. This assumption can easily be relaxed. If analysts have misspecified views about the relevance of each variable, such that their perceived weight is $\tilde{v}_k \neq v_k$, then the analysis goes through by replacing v_k with $\tilde{v}_k$ as we take expectations under analysts' subjective expectations. This also applies to misperceptions about other primitives of the model.

Relative to the objective function in equation (7), this expression shows that analysts are now more cautious about allocating attention to variables for which the true and misspecified models diverge more. This becomes even clearer when solving for the resulting optimal attention allocation, and comparing it to the true valuation model benchmark in (11):

$$\tau_{s1}^{m} = \frac{C + c_2\tau_{02} - \sqrt{\frac{v_2^2-(v_2-m_2)^2}{v_1^2-(v_1-m_1)^2}}\sqrt{c_1}\sqrt{c_2}\tau_{01}}{c_1 + \sqrt{\frac{v_2^2-(v_2-m_2)^2}{v_1^2-(v_1-m_1)^2}}\sqrt{c_1}\sqrt{c_2}}. \tag{15}$$

Proposition 3 (Attention Allocation when Using a Misspecified Valuation Method)**.** *When analysts use a valuation method m that deviates from the true valuation function v (i.e., $m_k \neq v_k$ for some k), they allocate more attention to variable x_k and less attention to variable x_j—relative to the benchmark with no deviations—if and only if $\left(\frac{v_k-m_k}{v_k}\right)^2 < \left(\frac{v_j-m_j}{v_j}\right)^2$. That is, analysts allocate relatively more attention to variables where the true and misspecified models are more closely aligned, and relatively less attention to variables for which the discrepancy is larger.*

For instance, consider a growth stock whose true model places substantial weight on long-term growth prospects. A DCF method—which explicitly projects future cash flows—aligns more closely with this emphasis than a multiples-based approach that focuses on peer comparisons or historical trends. According to the proposition, because the DCF method's discrepancy from the true model is smaller in the growth dimension, analysts using DCF will allocate more attention to signals that inform these prospects, whereas those using a multiples-based approach will allocate less attention to them.

Moving to the second step, analysts choose the valuation model (and corresponding attention weights) that minimizes the expected mean squared error. Substituting the optimal attention weights in (15) into the mean squared errors expression in (13), we find that the

resulting MSE from using valuation method m can be written as:

$$MSE(m) = \underbrace{\frac{\left(\sum_{k=1,2}\sqrt{c_k\left(v_k^2-(v_k-m_k)^2\right)}\right)^2}{C+c_1\tau_{01}+c_2\tau_{02}} + \sum_{k=1,2}\frac{(v_k-m_k)^2}{\tau_{0k}}}_{\text{variance}} + \underbrace{\left(\sum_{k=1,2}(v_k-m_k)\mu_k\right)^2}_{\text{bias}} \tag{16}$$

This equation shows that using a misspecified model ($v_k \neq m_k$) unambiguously increases the bias term, but it may decrease the variance term. Consequently, analysts face a bias/variance trade-off, meaning that it is not necessarily optimal to choose the correct valuation model. Instead, accepting some degree of model misspecification (and thus some bias) can reduce mean squared error if it allows analysts to reallocate attention more effectively toward variables that are more important, more volatile, and are less costly to acquire information about.

Proposition 4 (Optimal Choice of Valuation Method)**.** *When analysts can choose among multiple valuation methods, they may optimally deviate from the true valuation model, as they face a bias-variance trade-off. Relative to a baseline with no wedges, it can be optimal to introduce a wedge $w_k^2 = (v_k^2 - m_k^2) > 0$ on variable x_k if doing so reduces overall MSE. Ceteris paribus, this becomes more likely when x_k is less relevant (lower v_k, μ_k), when the value of learning about the other variable is higher (lower τ_{0j}, c_j, or higher v_j), or when the attention budget C is lower. Conversely, analysts are more likely to align their model with the true valuation function (i.e, set $w_k^2 = 0$) when x_k is more relevant, the value of learning about the other variable is lower, or the total budget on attention is higher.*

Returning to our earlier example, growth stocks place a high weight on long-run outcomes. Therefore, when valuing growth stocks, analysts are more likely to use a DCF model that minimizes bias on that crucial dimension. In contrast, with value stocks, long-run growth is less central, making analysts more comfortable using simpler multiples–even if doing so might introduce some bias along that dimension (which is now less important).

This analysis illustrates how the choice of valuation methods both shapes and is shaped by what analysts attend to. These considerations can explain why analysts may adopt different

valuation methods and attention weights across industries or firm types or over time, as we examine empirically below.

2.4 Heterogeneity in Information Processing: Top-down and Bottom-up Factors

So far, we have worked under the assumption that all analysts know and agree on the true primitives of the models (v_k, τ_{0k}, c_k). In this section we explore instances where analysts have heterogeneous (and potentially misspecified) beliefs about those primitives ($\tilde{v}_k^i, \tilde{\tau}_{0k}^i, c_k^i$). Since analysts in Section 2.3 solve the model under their own subjective beliefs, we can study the implications of misspecified beliefs by simply replacing true primitives with analysts' perceptions of them in the solutions in (14) and (16).

Once we allow for these differences, the choice of valuation method is driven not only by the relevant firm's fundamentals, but also by an analyst's type and their individual level ability to process information, which can be affected by both top-down and bottom-up factors. Starting from top-down factors, if analysts have a different education and training in stock valuation (e.g., if located in different countries), they will find it easier to process information using one method over another. This can lead to differences across analysts in how they weight a given feature even when using the same valuation method ($m_{DCF,k}^i \neq m_{DCF,k}^j$). Turning to bottom-up factors, differences in analysts' experiences and exposures to certain variables (e.g., inflation) may make a given feature more prominent and available to some analysts than others ($c_k^i \neq c_k^j$).

In light of this heterogeneity, differences in mental models stem from both firm-specific factors and individual-specific experiences. Specifically, individual-specific experiences determine the default frame analysts are more likely to approach valuation with, while attention to firm-specific features leads analysts to adjust their frame to account for changes in the underlying environment (Bastianello and Imas, 2025). In our analysis, we verify the empirical relevance of both channels.

2.5 Empirical Predictions

The challenge in testing these models generally comes from the fact that we do not observe what people pay attention to, nor do we observe what simplifying rules they adopt to tackle

the problem at hand. In what follows, we overcome both these challenges by using the text in equity analyst reports to obtain proxies for $a(s-\mu)$ (analysts talk more about a feature when they are paying more attention to it, and when there is more news) and m (analysts' valuation methods).

With that at hand, there are four sets of predictions we test empirically. First, analysts attend to variables that are more relevant, more volatile, and are less costly to acquire information about. This holds both in the time-series and in the cross-section. Second, analysts choose valuation methods based both on how easy it is for them to implement different valuation methods (which is e.g., impacted by their training and experiences), and on how relevant different variables are. Conditional on analyst type, analysts choose valuation methods to emphasize dimensions that matter most. Third, the choice of valuation methods and attention are tightly interlinked. When analysts agree on valuation methods, they attend to more similar information, leading to more similar mental representations. Moreover, analysts should pay more attention to variables that receive greater emphasis under the chosen valuation method. Finally, the expression derived in (5) allows us to empirically study the key drivers of both changes in valuation over time and disagreement. Both of these quantities can either be driven by variation in how analysts mentally represent the firm (attention weights), or in how they price different features (valuation weights).

3 Data Collection

Our empirical analysis is centered on the information contained in equity analyst reports. In this section we document our data-collection efforts. We start by introducing the institutional details for our setting, and describing what features of analysts' mental models we collect from their reports. We then explain how we construct our sample of reports and the exact methodology used to extract information from them.

3.1 Institutional Details

Sell-side analysts write reports to offer valuation forecasts and stock recommendations for the stocks they cover. In these reports they provide both quantitative forecasts (which have been

used extensively in empirical asset pricing), as well as information regarding the underlying models and reasoning they use. In this section we describe details on analysts' incentives and the regulatory background that guide our data collection, measurement strategy, and empirical design.

The most comprehensive and central outputs of analyst reports are their price target and investment recommendation. As shown in Figure H.1, this information typically appears at the top of the first page, reflecting how this is part of analysts' incentives. Specifically, the two primary institutions that rank analysts, the Wall Street Journal (WSJ) and Institutional Investor (II), base their evaluations on the quality of these outputs. WSJ ranks analysts based on the performance of their stock recommendations relative to realized returns.[11] II further conducts surveys of key market participants (e.g., portfolio managers and buy-side analysts), asking which analysts provide the most valuable insights and advice. Holistically, this tends to reward analysts who construct high-quality valuation frameworks and identify firm-specific factors most relevant to price dynamics.

Analysts' incentives to provide both accurate forecasts and high quality investment theses is further corroborated by regulatory requirements, which ensure that the information we seek is systematically included in equity reports. Starting in July 2002, equity analysts working for brokerage houses intending to distribute financial material to U.S. clients have been subject to disclosure rules mandating that any valuation model directly used to derive a published price target be explicitly stated in the report (NASD Rule 2711, now superseded by FINRA Rule 2241).[12] These rules create a reliable link between stated valuation methodologies and the mental models we aim to measure. Not only does this mitigate endogeneity concerns about selective or strategic disclosure, but it also reflects regulators' attempts to ensure that analyst reports go beyond narrative framing.[13]

While analysts do have incentives to produce accurate forecasts, career concerns, repu-

[11] Since 2008, the Wall Street Journal has excluded earnings forecast accuracy from its methodology, focusing instead on the performance of recommendations relative to price movements.

[12] This includes non-U.S. analysts employed by a U.S.-registered broker-dealer or FINRA members, as well as any analyst whose reports are intended for distribution to U.S. clients. Nearly all leading broker-dealers, who constitute the focus of our sample, are U.S.-registered and routinely disseminate research globally, including to U.S. investors. As a result, equity reports produced by any major investment bank or brokerage with U.S. operations are almost certainly subject to these regulations.

[13] Noncompliance can result in significant penalties for both the analyst and their affiliated brokerage firm, including fines and suspension.

tational dynamics, and external pressures also play a role. Ultimately, however, prior work has shown that these quantitative forecasts are useful in explaining asset pricing dynamics, regardless of the incentives that produce them. Moreover, Bastianello (2025) shows that analysts' return expectations are less biased, more informative, and more relevant than commonly studied surveys, including the Graham-Harvey survey of CFOs, the Livingston survey of economists, and the Capital Market Assumptions (CMAs) of large asset managers. Using this as a starting point, we are interested in studying the arguments behind those quantitative forecasts, and Section 5.1 provides evidence on our ability to comprehensively capture the reasoning behind such forecasts, above and beyond these documents being written purely for the purpose of persuasion.[14]

The fact that price targets are the i) final output of the reports, that ii) there are rules to ensure that analysts provide us with the methods they used to obtain them, makes them an ideal starting point for our analysis. In particular, the fact that price targets are the end result of the report's analysis means that all sections of the report are relevant for understanding how analysts reason about this output.[15] This allows us to offer a comprehensive analysis that makes full use of the reports.[16] The fact that there are also regulations in place for analysts to disclose their valuation methods then further ensures our ability to capture models as well as reasoning of our chosen output.

3.2 Defining the Components of Analysts' Mental Models We Extract

The first choice we had to make to measure analysts' reasoning regarded the type of information to extract from the text portion of our reports. Prior work has either employed human workers (common in surveys) to organize text data in the form of interpretable directed-acyclic-graphs, which capture causal reasoning, or it has employed machine learning methods such as embeddings or LDAs, which capture semantic structure and organizes large bodies of text into high-dimensional vector spaces or topic distributions. To extract information

[14] Sections 5.1 and 5.2 show that the reasoning analysts provide in their reports well explains their quantitative forecasts, while topics they omit (and which others refer to) do not explain their quantitative forecasts but help in explaining their forecast errors.

[15] Our choice to focus on price targets rather than stock recommendations stems from the fact that stock recommendations are categorical, while the continuous nature of price targets allows for greater nuance.

[16] If instead we had decided to focus on a single input, such as earnings forecasts, we would have been constrained to only using sections of the reports relevant for that.

that is both directly interpretable and scalable, we instead employ Large Language Models (LLMs), which combine the key advantages of both human interpretability and machine learning scalability.[17]

Given our choice to adopt LLMs, this section outlines the information we collect about analysts' mental models from their equity reports. The term *mental model* generally refers to how individuals reason and form relationships across variables to support their arguments.[18] In our setting, we operationalize mental models by capturing these relationships as the combination of (i) the valuation methods analysts use and (ii) the structured reasoning they provide to support their forecasts. In what follows, we start by providing a high-level overview of the data we prompt LLMs to collect for both components of analysts' mental models. We then provide an in-depth description of how we extract information from reports in the next subsections.

Starting from the valuation methods analysts use to compute their price targets, we collect the following information:

- `Valuation Method Tied to Price Target:` the valuation method (formula)—or methods if more than one—directly tied to the report's price target. These are selected from a list that includes discount cash-flow (DCF) model, price-based multiples, earnings-based multiples, asset-based valuation, and so on.

[17] Traditional machine learning methods, such as LDA, are cheap, scalable, and efficiently identify topics in large documents, but the resulting topics are often difficult to interpret. In the context of equity reports, these methods tend to reliably produce broad thematic categories, such as grouping content by industry, institution, or region (Decaire and Guenzel, 2025). We are, however, interested in capturing a more granular picture of analysts' chains of thoughts as well. In contrast, human workers can interpret text, which allows for the extraction of topics that are interpretable, and for the organization of arguments via their causal relationships (Andre et al., 2024a). However, human annotation limits scalability, and makes it challenging to design methods to ensure the robustness of the outputs. We view LLMs as a way to make progress on both of these strategies limitations, while keeping their benefits. It enable us to generate clear and interpretable topics and chains of thought in a scalable way.

[18] Existing research, often using surveys, has typically focused on more aggregate narratives to distinguish between different mental models that people may entertain in a given situation. Examples include whether individuals adopt a demand- or supply-side view of inflation (Andre et al., 2024a), or whether they believe in market efficiency or mispricing (Andre et al., 2024b). Regardless of the adopted setting, mental models are used to capture how people think about relationships across variables. Our dataset allows us to identify such relationships at a more granular level. We chose this approach in order to offer a comprehensive window in all aspects of analysts' reasoning. Appendix D and Figure H.2 offer a complementary example of how our methodology can also be used to elicit people's mental models by imposing more structure ex-ante. Specifically, we instruct the LLM to capture how analysts reason about inflation during the most recent surge, using the same narrative classification as in Andre et al. (2024a). We view these type of exercises as interesting avenues for future research.

- `All Valuation Methods:` all other valuation methods mentioned in the report (e.g., as robustness checks or reference points).
- `Multiples Reference Point`: when multiples are referenced in the report, we additionally gather information on whether they are computed through forward-looking considerations, historical comparisons, or peer or industry comparisons.

Next, to capture the reasoning analysts use to justify their forecasts in a nuanced yet structured way, we classify analysts' arguments along the key dimensions that have informed much of the debate in empirical asset pricing: valuation channel (cash flows or discount rates), time horizon (past, present, near-future, distant-future), and sentiment (positive, neutral, negative). Each line of reasoning is then structured as follows:

`{Topic}//{Entities}//{Valuation Channel}//{Time Outlook}//{Sentiment}`

- `Entities`: firm, industry, macroeconomy.
- `Valuation Channel`: cash flows, earnings, sales, costs, profitability margins, discount rates, relative valuation.
- `Time Outlook`: past, present, near-future (1-3 years), distant-future (3+ years).
- `Sentiment`: positive, neutral, negative. This captures whether the argument in question contributes positively, negatively, or in a neutral way to the price target.

Appendix C and F provide examples of such output. In what follows we describe the sample and methodology used to extract this information from equity analyst reports.

3.3 Data Collection and Sample

This subsection details how we construct the samples to extract the two elements of analyst mental models: (i) the valuation methods, and (ii) the lines of reasoning. In brief, we download the universe of equity analyst reports available on Refinitiv which forms the basis for our analysis. For the more cost-intensive parts of the analysis, we then construct a smaller subsample, as described below.

Universe of equity reports in Refinitiv. In downloading the universe of equity reports from Refinitiv Eikon (LSEG Workspace), we focus on the period from 2000 to 2025 and restrict

our attention to the 43 most common brokerage houses in the data, following Decaire and Graham (2024). We apply a set of light data filters during the extraction process, including a maximum page count of 30, a restriction to English-language and company-specific reports (i.e., we exclude industry and macro reports), as well as the exclusion of reports related to M&A activity. This results in a total dataset of 2.1 million reports.

From the universe of equity reports, we instruct Gemini Flash 2.0 to extract the number of firms covered in the report for which a price target is generated, the market price at the time of publication, the price target and their currencies, analyst name, location, and phone number, as well as the word count of the valuation thesis (excluding regulatory boilerplate, figures, and tables) and the valuation methods employed (see Appendix C for prompt details). We exclude reports producing price target for more than one firm, those for which any of this information is missing, those for which the currencies do not match between price target and market price, and reports that contain less than 200 words or more than 5,000 words.[19]

Last, we convert the reports from PDFs to text format, by extracting all available text content. This includes the valuation thesis, figure titles and captions, tables, and regulatory boilerplate language. The final sample comprises 1.38 million reports in text format and serves as the basis for our analysis to extract the valuation methods.

Subsample used to measure reasoning. Since the process we employ to extract the lines of reasoning is computationally intensive and costly, we use a smaller set of reports to extract the second component of analysts' mental models. To do so, we select a subsample designed to be representative while maximizing both time-series and cross-sectional coverage at the analyst level.

Specifically, we begin by identifying the lead analysts from a random sample of 30,000 reports and all reports used in Decaire and Graham (2024). Then, we expand our selection to include all reports authored by those lead analysts in the first quarter of each calendar year (January 1 to April 1) over the entire range of our sample, i.e., 2000 to 2025. Ultimately,

[19] To validate the accuracy of the extracted data from the filtered set of reports, we randomly sampled 1,000 reports and instructed two research assistants to manually verify the analyst names, price targets, market prices, and phone numbers. The extracted information matched the manually verified data in all cases. Finally, to ensure that analyst names extracted from the reports can be reliably linked to the correct individual (e.g., distinguishing between "John H. Smith" and "John Smith" both working at Morgan Stanley, or both analyzing Apple Inc.), we perform a deduplication procedure combining fuzzy matching, large language model (LLM) assistance, and manual verification.

this procedure yields 301,364 reports covering 18,284 firms worldwide and authored by 8,578 lead analysts, forming the basis for our analyses of valuation methods and analyst reasoning.

Finally, we isolate the valuation thesis from the other parts of the reports by removing the text associated with regulatory boilerplate language, figures, and tables. This yields a cleaner version of text that focuses exclusively on the narrative content used to justify analysts' price targets.

3.4 LLM-Based Extraction of Mental Models

This section details how we apply LLMs to extract the two components of analysts' mental models: (i) the valuation methods, and (ii) the lines of reasoning. We discuss each in turn.

3.4.1 Extracting Valuation Methods

We extract valuation methods from the full universe of reports using Gemini 2.0 flash. To validate the accuracy of the prompt, we conducted manual checks on a random sample of 200 reports, with two research assistants and the authors independently reviewing the model outputs with their subjective assessment. This process not only provided a benchmark for evaluating the AI's performance but also revealed cases in which trained financial professionals' opinion diverged, allowing us to fine tune our approach. With this in mind, when extracting the valuation method tied to the price target, we adopted a conservative strategy: we required the valuation method to be *explicitly* linked to the generation of the price target. This rule minimizes both disagreement among human coders and between human coders and the AI, effectively permitting clear and consistent output. We then also collected all other valuation methods mentioned in the report, as well as information on relevant reference points when multiples were used. Appendix C provides the full prompt along with representative examples of the extracted valuation methods.

3.4.2 Extracting Lines of Reasoning: a three-step-approach

We extract the lines of reasoning from the subsample of 301,364 reports discussed in Section 3.3. Extracting data on lines of reasoning is a more involved task, as it requires the LLM to understand the subtleties of analysts' causal reasoning, and to organize these arguments

as directed acyclic graphs containing all the components we aim to collect.[20] To do so, we employed Claude 3.5 Sonnet from Anthropic and introduced a new three-step extraction method, as well as a new diagnostic tool that allows us to assess the completeness and stability of our output. Throughout our analysis, we set the temperature and top_p parameters to zero, which limits variation and prevents the model from introducing content beyond what is explicitly present in the reports.[21] Appendix B displays the full prompt, and Appendix F provides an example excerpt from an equity report along with the full corresponding model output.

As discussed in Section 3.2, we structure analysts' reasoning as a combination of a topic an analyst may be discussing, together with its associated entity, valuation channel, time outlook, and sentiment. To support the classification of the time outlook node, the report's publication date is provided. Moreover, to facilitate output verification, we also extract two snippets of words directly from the report. The first one identifies one or two keywords that best characterize the sentiment associated with each topic. The second one extracts the first five words linked to the relevant section of the text, enabling us to locate the topic within the document and assess whether the extracted topic is contextually plausible.

To extract this information we then proceed in three steps. The first two steps differ in the information we provide as input to the LLM. The third takes the output from the first two step and adds additional structure to the topic node.

Specifically, in Step 1, we divide each report into 200-word segments and prompt the AI iteratively on each segment. The 200-word threshold is motivated by the fact that the shortest documents in our sample contain at least 200 words, and we provide further discussion on this in Appendix B. Crucially, this *chunking method*—i.e., segmenting reports into smaller

[20] While smaller open-source LLMs, such as Gemini are considerable cheaper and perform well for simple tasks, in our internal tests as of April 2025, we found significant lower output quality and depth when using those smaller models to extract the mental models, in contrast to using Claude Sonnet.

[21] LLMs can be used for a range of tasks involving unstructured text which broadly fall into two categories: (i) descriptive extraction, (ii) predictive generation. Our approach is descriptive—we ask the LLM to summarize and structure information present in the report, rather than to predict outcome based on that information. Concerns about potential training and data leakage are therefore less of a concern (Sarkar and Vafa, 2024). Moreover, recent evidence suggest that leakage-induced biases may be modest in certain financial contexts (Engelberg et al., 2025, He et al., 2025). To confirm that the topics identify by the AI are effectively referenced in reports, the lack of data-leakage in our data collection by showing we are able to locate 100% of the arguments we extract in the reports themselves.

"sub-reports"—improves the AI's ability to detect the full set of topics and enhances output consistency, as we detail in Section 3.6.

Step 2 then addresses a potential concern with the chunking method—namely, that segmenting the text introduces the risk of missing topics that span multiple segments and are therefore not fully captured. In Step 2, we provide the AI with both the full equity report and all all topics identified in the first step. The AI then re-evaluates the report holistically to identify any potentially missing topics, and adds these topics to the final output if necessary.

Importantly, in the first two steps, we deliberately impose *no* ex-ante structure on the type of topics the AI is allowed to extract. As long as the AI determines a relation between a topic and the valuation exercise, we do not initially constraint topics to a predefined set of labels. In Step 3, we instead group the unstructured topics extracted in the first two steps into 139 standardized labels. The construction of this label set is guided by two inputs: a word2vec-based clustering of 200 commonly extracted topics, and financial intuition to ensure that the resulting labels are intuitive and economically meaningful in the context of valuation (see Appendix B for further details).[22] For the components of our analysis that rely on interpretability we further aggregate these topics into 32 broader categories (shown in Appendix Table E). We perform a series of checks to ensure that our results are not sensitive to specific aspects of the topic classification process, as detailed in robustness tests below.

3.5 Basic Report Anatomy: location of topics, sentiment, and time outlook

Before comparing our multi-step prompt to the performance of a naïve, single-step prompt, we first provide detail on the basic "report anatomy" of mental models. Consistent with the intuition that introductions and conclusions of reports are more content-dense, we find that the LLM extracts slightly more information and topics from the very beginning and end of reports (Panel A of Figure I; see also Panel C of Figure II, discussed further below).

[22] We use a two-tier structure for the topic labeling. We first instruct the AI to assign each raw topic to one of 128 refined standardized labels (Appendix Table E1). If no clear match is found, the AI selects from a second, coarser set of 11 categories (Appendix Table E2). This approach reflects the way analysts structure discussion in equity reports. For example, an analyst might explicitly state, "Inflation pressure will increase the cost of goods sold," linking inflation to costs. In other cases, they may simply write, "The cost of goods sold will increase," omitting the causal factor. To capture both types of expressions systematically, our final prompt is designed to favor more granular labeling when possible, while remaining flexible enough to include broader statements when necessary.

Moreover, the nature of arguments varies systematically across a report's progression (Panels B to D of Figure I). Early sections ("executive summary") tend to be more backward-looking and review past developments and earnings. The middle section ("investment thesis") is more forward-looking, with sales-related topics dominating the discussion.[23] In the final section ("key risks"), analysts tend to focus on future downside scenarios, with sentiment declining sharply and attention shifting toward industry- and macro-level topics.

One important takeaway from this report anatomy evidence is that different types of information appear in different parts of analyst reports. As a result, less sophisticated prompting strategies that disproportionately focus on, for example, the beginning of a report risk producing mental models that are not only incomplete but also systematically distorted.

3.6 Diagnostic Tools: the importance of a multi-step LLM approach

The first step we take to validate our output is to check that the snippets associated with each line of reasoning are contained in the report. We find that 100% of the snippets used to validate topic extraction are indeed drawn from the original document.[24] We next provide evidence on the importance of using a sophisticated LLM prompting strategy to extract mental models comprehensively—at least given current capabilities of LLMs. To this end, we compare the output of our multi-step LLM prompt to that produced by a single-step naïve prompt, i.e. a prompt that collapses the three-step procedure from the previous section into a single query.

Comprehensiveness of Extracted Topics. Panel B of Figure II shows that the output generated by a naïve one-step prompting strategy exhibits both lower breadth (i.e., fewer topics collected) and lower depth (i.e., fewer individual arguments per topic) relative to our multi-step approach. When provided with the full text at once, the AI tends to extract only one argument per topic, a pattern that holds across reports of varying lengths. (The median report contains just under 1,000 words, but the distribution is right-skewed, with

[23] This is consistent with our findings below on the dominance of top-line items in analyst discussions (Section 4.2), as well as survey evidence in Graham (2022), which shows that managers primarily focus on forecasting sales, reflecting the widespread use of percent-of-sales techniques in pro forma planning.

[24] In rare cases (1.4% of snippets), the AI fixed typos or combined non-contiguous text to enhance interpretability. For instance, in a January 13, 2011 report for RAND.AS from RBC Capital Markets, the AI extracted "*expect further recovery from later-cycle*" to represent the text "*expect further recovery from **[the]** later-cycle*." We view these instances as acceptable in context.

a substantial number of longer reports, as shown in Panel A of Figure II.) By contrast, our multi-step prompt consistently captures multiple arguments per topic.[25] Moreover, the AI displays a form of "laziness" under the one-step strategy, concentrating data collection almost exclusively at the beginning of the report (Panel D), whereas our multi-step approach ensures coverage throughout the full document (Panel C; see also Panel A of Figure I). This pattern for the single-prompt approach is especially concerning given that, as discussed above, the nature of information varies systematically across different sections of the report, and that arguments throughout all sections of a report help in predicting analysts' quantitative beliefs.[26]

Stability of Extracted Topics. We also introduce a novel method for assessing the robustness of extracted mental models and quantifying the degree of AI-induced variability. To do so, we use an approach that is akin to evaluating coefficient stability across bootstrapped samples, or to asking multiple equally well-trained human workers to extract mental models from the same report. We note this approach can be applied to LLM-based text extraction across contexts. Specifically, we prompt the AI to extract the mental model from a given equity report *repeatedly* (ten times) and compute the Jaccard topic similarity score for all pairs of attempts, defined as:

$$\text{Jaccard similarity} = \frac{|A \cap B|}{|A \cup B|} \tag{17}$$

where A and B represent the sets of topics extracted in attempts A and B of a given report, $|A \cap B|$ is the number of overlapping topics across attempts, and $|A \cup B|$ is the total number of distinct topics identified in at least one of the attempts. The Jaccard similarity thus quantifies the degree overlap between the topic sets, ranging from 0 (no overlap) to 1 (complete overlap).

[25] Appendix Table H.1 confirms this pattern using within-report statistics. In the average report, 23% of topics are linked to more than one valuation channel (e.g., sales, costs), and in 24% of cases, a topic is associated with multiple sentiments (e.g., inflation is expected to *positively* impact valuation through sales but *negatively* through costs).

[26] Specifically, as one concrete example for this, Appendix Table H.2 shows that argument sentiment at all points in the report significantly predicts analysts' price target forecasts (and does so positively, as one would expect). This result holds both when sentiment from different report sections is included separately and when it is included jointly as predictors.

We implement repeated prompting using a stratified random subsample of 240 reports.[27] With our multi-step approach, the Jaccard score for topic overlap across repeated prompts is consistently very close to 1, across the full distribution of report lengths. This confirms that the LLM output is highly stable across repeated LLM extractions (Panel E of Figure II). By contrast, the single method exhibits greater variability in topic output, with Jaccard similarity scores consistently lower than those produced by the multi-step prompt. These discrepancies are even more pronounced when we evaluate Jaccard similarity as a function of the number of arguments in Panel F.[28]

4 Mental Model Summary Facts

This section starts by presenting aggregate patterns for both components of analysts' mental models. We then corroborate our theory's predictions that analysts pay more attention to topics that are more relevant, more volatile, and easily available, and that the choice of valuation methods and attention weights are tightly interlinked.

4.1 Valuation Methods are Tailored to Emphasize Important Dimensions (and to Minimize Processing Costs)

Across the near-universe of equity reports, Panel A of Figure III shows that the two most prominent valuation methods used to form price target forecasts are discounted cash flow methods (used in 43% of reports), and price-to-earnings multiples (32%), followed by enterprise value-to-EBITDA multiples (22%), sum-of-the-parts (16%), and net asset value (6%). Moreover, 62% of reports use at least one multiples-based approach.[29] Panel B then shows

[27] We stratify the random subsample based on report length, by increments of 100 words and randomly select 5 reports in each stratum. This allows us to study the robustness of the LLM extraction approach over the entire range of report lengths using a well-populated subsample.

[28] The "missing" dots for the single-step prompt beyond just above 20 arguments reflect the fact that it never extracts more than this number of arguments to begin with (cf. Panel B of Figure II), let alone consistently across repeated extraction attempts.

[29] Figure H.5 highlights additional variation we have collected to add nuance to multiples-based valuation methods. Benchmarking multiples to historical values is less common than benchmarking them to multiples implied by the current market price. Moreover, in a majority of cases multiples are justified using forward-looking considerations, and peer comparisons are also common. We caveat these statistics noting that the LLM-based extraction may somewhat overstate the use of future reference points when it classifies a trailing multiple applied to a future estimate as future-oriented. Appendix C and Appendix G provide the prompt we used, together with several examples of reference points from our sample.

how the usage of different methods fluctuates over time. Most notably, periods with a large share of expected negative earnings-per-share estimates see an increase in the usage of DCF away from P/E multiples.[30] This pattern is especially pronounced during the financial crisis.

According to our model, this high prevalence of multiples-based approaches may reflect i) the ability of multiples to emphasize firm-features that are actually more important, ii) the ability of multiples to emphasize features that analysts perceive to be more important, iii) the greater ease of applying multiples relative to DCF, which allows for lower processing costs. While the first component implies that the choice of valuation method should be based on firm fundamentals, the latter two components allow for analyst-specific characteristics, such as their background and training, to also play a role. In what follows, we provide evidence that both firm- and analyst-specific traits contribute to the choice of valuation methods.

Starting from the relevance of firm-specific features in shaping the choice of valuation method, Panel A of Figure IV reveals substantial variation in valuation methods across industries and Figure V focuses on the cross-section of stocks. For example, Figure V shows that, even after controlling for the share of negative EPS estimates, DCF methods are relatively more common for small, young, and growth firms (for which future cash-flows play a larger role), while multiples-based approaches become more prevalent for more mature firms (where observable comparables are more available and relevant).[31] These patterns still hold once we include analyst-firm fixed effects (Figure H.3). This is consistent with our model's prediction that, conditional on an analyst's type (e.g., holding fixed their background, general expertise, and location/culture), analysts do indeed tailor valuation methods to emphasize more important dimensions as firms change over time.

Turning to the role of analyst-specific traits in driving their choice of valuation methods, we consider how the choice of valuation methods varies with analyst's expertise and familiarity with different valuation methods. Different levels of training may introduce differences in

[30] Applying a P/E multiple to a negative EPS estimate would result in a negative valuation, meaning that analysts need to resort to other valuation methods. Panel B also shows how the share of reports using DCF was rising in the first part of our sample. This reflects changes in regulation that required analysts to disclose their valuation methods.

[31] Patterns across industries in Panel D of Figure III are also consistent with our theory's prediction. For example, the relatively higher usage of DCF methods in information and professional services is consistent with the forward-looking nature of these industries, which often focus on innovation and long-run investment horizons. Instead, the dominance of multiples-based approaches in industries such as retail is consistent with their emphasis on observable comparables.

analysts' ease and ability to process information using a given valuation method, leading analysts to sort into the usage of different methods for the same firm. To illustrate the relevance of this channel, Panel B of Figure IV shows that European analysts have a higher propensity to employ DCF models than their American counterparts. This holds even once we include firm-year fixed effects (Figure H.4), meaning that we are not only capturing analysts in different continents sorting into covering different firms.

Overall, these results suggest that conditional on an analyst's type (and ability to process information using different valuation methods), analysts tailor their valuation methods to highlight the most important dimensions of the firm they are covering. However, training and experience with processing information through different methods also plays a meaningful role.

4.2 Attention Allocation to More Relevant, Volatile, and Available Topics

Table I shows basic summary statistics of the lines of reasoning we extract using the main sample of 301,364 reports described in Section 3.3. The average report contains 39 distinct lines of reasoning, comprising 16 out of 139 distinct topics per report, with each topic discussed through 2.3 individual arguments.[32]

At an aggregate level, Figure VI shows that analysts predominantly focus on top-line items (with 40-50% of arguments relating to sales), firm-related topics (75-80% of arguments), and forward-looking arguments (50-60% of arguments relate to the near-future). Moreover, our measure of sentiment (coded as $+1$ = positive, 0 neutral, and -1 = negative), attention to top-line items, and attention to firm-related topics are pro-cyclical, while focus on discount rates and macro-related topics is counter-cyclical and doubles during the financial crisis.[33] Table II additionally shows how differences in attention to the two most prominent valuation channels (sales and costs) translate into analysts' quantitative forecasts. When news have a positive (negative) sentiment, greater attention to a given variable leads to an increase

[32] For example, analysts may talk about inflation (a topic) both in terms of costs and in terms of sales (two different valuation channels), resulting into two different lines of reasoning associated with the same topic.

[33] Appendix Figure H.6 further shows that sentiment is most volatile for arguments that relate to the near-future, while distant-future sentiment is both more stable and more positive than sentiment at other horizons.

(decrease) in analysts' price targets.[34] Moreover, Table H.3 shows that the average time outlook and sentiment of a report are also both individually and jointly predictive of analysts' quantitative forecasts.[35]

At a finer level, Figures VII and VIII present heatmaps showing the average share of attention that the typical report in our sample allocates to each topic in a given year. The first figure shows attention to the 50 most frequently discussed topics, whereas the second figure shows attention to the remaining 75 topics.[36] Our theory predicts that analysts should allocate more attention to topics that are more relevant, volatile, and easy to process. In what follows, we start by exploiting time-series and cross-sectional coverage of stocks to illustrate the role of relevance and volatility. We then exploit differential experiences across analysts to provide evidence that the availability of a variable also affects attention allocation.

Consistent with the predictions that attention allocation is based on a variable's relevance and volatility, topics vary both in the amount of attention they receive, and in how stable their attention allocation is over time. With this in mind, we identify three types of topics. First, there are topics that are *relevant and volatile*, and which therefore receive a large and persistent share of attention. This includes topics related to fundamental drivers of cash flow generation, such as customer demand, market expansion and market share, capital expenditures, cost structure and efficiency, and mergers and acquisitions. Second, there are topics that are *relevant but stable*, and which only receive heightened and transitory attention when they deviate significantly from their steady state values. Most notably, taxation becomes more prominent during the first term of the Trump administration, attention to tariffs spikes during both terms of the Trump administration, and topics related to public health, supply chain issues and inflation receive heightened attention starting in 2020. The third type of topics reflects *emerging* ones, and which receive attention as either their true or perceived relevance rises over time. Examples of topics whose perceived relevance changed over time

[34] For example, a one standard deviation increase in positive-sentiment (negative-sentiment) sales attention is associated with an upward (downward) shift in price target forecasts of 6% (12%) relative to the mean, including when estimated using only within-analyst-firm and within-firm-year variation. The corresponding effects for attention to cost channels are slightly smaller but remain economically meaningful.

[35] Section 6 below presents additional evidence that variation in attention, not just to valuation channels, but to specific topics and categories, significantly predicts changes in analyst valuations, disagreement, and both overreaction and underreaction.

[36] We note that the color intensities are only comparable within each figure, not across the two figures. The figures do not include topics belonging to the valuation or undetermined categories (cf. Appendix E).

include environmental policies and ESG as well as corporate social responsibilities (CSR), which are much more prominent in the latter part of the sample. This also illustrates how perceived relevance can be endogenous. Examples of topics whose true relevance changed over time include AI and cognitive applications, automation and robotics, or cybersecurity and data protection, which emerged as a new topics only recently. This last category of topics highlights our ability to flexibly capture new categories that analysts deem relevant for valuation, and allows us to effectively study changes in representations over time, as further discussed in Section 6.[37] Finally, Appendix Figures H.7 through H.14 show how attention to topics varies based on relevance and volatility even across industries.[38]

Having established that analysts pay more attention to more prominent and more relevant variables, we turn to the last component of our model that influences attention allocation: the cost of acquiring information. To do so, we exploit variation in experiences across analysts, which may make certain variables more readily available to some analysts than others. Specifically, our data includes the identities of individual forecasters as well as their geographic locations, inferred from regional codes in phone numbers. We use this data to assess whether differences in analysts' exposure to certain variables affects how much attention they allocate to them. Table III shows that analysts that are closer to a company's headquarters have a greater focus on firm-related topics than analysts located further away, who instead focus more on macro-related topics.[39] Moreover, Table IV shows that inference from local inflation experiences also influences analysts' attention allocation. Analysts located in countries with higher inflation pay more attention to inflation, even for firms headquartered in *other* countries (and controlling for the level of inflation in the firm's country). Specifically, this positive association emerges when local inflation is sufficiently salient—specifically, when

[37] In the appendix, we also show intuitive relationships between topics (and topic categories) and both time outlook and sentiment (Figures H.16 to H.19).

[38] For example, in resource-dependent sectors such as oil & gas and mining, forecasters tend to focus on commodity and raw material prices, oil and gas prices, and production capacity—firms' ability to secure natural resource reserves (Appendix Figure H.11). By contrast, dynamic and innovation-driven industries, such as the information sector and professional and scientific services, place greater emphasis on capital investment, customer acquisition, market share, and product and service mix (Appendix Figures H.9 and H.10). Appendix Figure H.15 shows how topic attention varies by whether analysts' arguments refer to the firm, industry, or macroeconomy.

[39] Table H.7 shows a strong mental model alignment effect associated with geographic proximity. Analyst pairs located in the same country exhibit significantly greater similarity in both narrative topics and arguments, with economic magnitudes comparable to those observed for shared valuation techniques in Table H.4.

it exceeds the commonly referenced 2% threshold. This analysis suggests that bottom-up factors also play a meaningful role in determining analysts' attention allocation.

4.3 Valuation Methods and Attention Allocation are Closely Linked

In this section we validate our theory's prediction that the choice of valuation method and attention allocation are tightly interlinked. There are two dimensions to this argument, which we discuss in turn.

First, conditional on analyst type, our theory predicts that analysts should choose both valuation methods and attention weights to emphasize the more relevant variables. An illustrative example of this pattern stems from looking at how changes in attention and valuation methods jointly vary across value and growth stocks. Consistent with the fact that long-term outcomes are more relevant for growth stocks, Figures V and H.23 show that, when covering growth relative to value stocks, analysts are not only more likely to use DCF methods (which put more emphasis on the long-run) but they are also more likely to focus on a more forward-looking future-outlook. These results hold even when we include analyst-firm fixed effects. This allows us to control for analyst-type (e.g., their background and training) and capture how analysts' choice of valuation method and attention allocation vary as the underlying firm changes over time. The significance of these results therefore confirms that analysts' focus on more future-oriented arguments when using DCF instead of multiples reflects their take on fundamentals and is not purely driven by their training or firm-specific knowledge.

Second, our theory also predicts that if analysts disagree in their choice of valuation method when covering the same firm at the same point in time, their attention allocation should also differ in a way that reflects those choices. Specifically, topic attention should vary across valuation methods depending on which variables are most relevant under the chosen model. Consistent with the fact that DCF and multiples put differential emphasis on different variables, Figure IX and Table V show that time outlook and topic coverage differ systematically across DCF and multiples-based methods. Starting from differences in time outlook, Table V shows that DCF usage is associated with arguments that are more forward-looking even once we control for firm-year fixed effects, or additionally analyst-firm

fixed effects.[40]

Turning to topic coverage, Figure IX shows that analysts using DCF focus more on topics related to systemic risk, interest rates, corporate investment, and GDP growth. These patterns are consistent with analysts reasoning in terms of projecting cash flows into the future and discounting them back to present value. Instead, analysts covering the same firms using multiples-based approaches have a greater focus on customers, products, and pricing—topics that reflect analysts' reasoning in terms of comparables.

These differences in reasoning then translate into systematically different forecasts, as shown in the last column of Table V: evaluating a firm using DCF produces systematically higher price targets than evaluating the same firm at the same time using multiples.[41]

5 Properties of Analysts' Mental Models

So far we have shown that analysts' attention allocation is determined by both top-down and bottom-up factors. On average, Table I shows that this results in them only considering 16 out of 139 possible topics in their reports. Moreover, there is only a 33% overlap in topics coverage across analysts evaluating the same firm in a given year (Panel D). These results persist even when we examine the longest decile of reports (exceeding 2,400 words), and when we vary the granularity of our 139-topic classification scheme by creating coarser categories

[40] The specification with analyst-firm fixed effects in Panel B of Table V relies on within-analyst-firm variation (i.e., changes in how an analyst evaluates a given firm over time). Changes in valuation methods over time may then be driven by: i) analysts' beliefs of changes in firm fundamentals over time, ii) changes in an analyst's familiarity with different valuation methods over time. To show that results are not only driven by the second channel, Panel C shows that DCF usage is still associated with more forward-looking arguments even when including analyst fixed effects and focusing on cases where analysts did not change their valuation methods for the specific firm. This suggests that at least part of the effect that we are picking up does in fact reflect that attention allocation and valuation methods are interlinked. In the analysis below, when examining valuation methods, we include similar specifications with analyst-by-firm fixed effects and/or condition on consistent valuation methods over time. For the sake of brevity, we do not repeat the interpretation of these specifications.

[41] Relatedly, Table H.4 shows that, when comparing pairs of analysts forecasting the same firm at the same time, those who align in their valuation method also align more closely in their topic attention and arguments. Topic and argument similarity across analyst pairs increase by 6% and 4% relative to the mean, respectively, when analysts use the same valuation method. Additionally, Table H.5 provides further evidence of mental model alignment across analysts, extending beyond valuation methods and attention to include alignment across argument components such as time outlook, valuation channels, and sentiment.

with much lower topic similarity (Panel E).[42]

These patterns raise three questions. First, whether the arguments we extract reflect a complete representation of analysts' mental models, or whether there may be other components of their reasoning that we systematically do not capture. Second, whether analysts' seemingly parsimonious models are justified by fundamentals, or lead to predictable mistakes in their quantitative forecasts. Third, if mental models are indeed sparse, we analyze whether such sparsity reflects rigidity in analysts' mental models, and if so what leads analysts to update their models. We address each question in turn.

5.1 Mental Model Completeness

To assess whether the mental representations that we are capturing are complete, or whether we are missing important arguments of analysts' reasoning that they omit from their reports, we analyze whether there exist topics that are absent from an analyst report but still relevant for explaining analysts' quantitative forecasts. The most natural candidates for such relevant yet omitted topics are those covered by other analysts for the same firm-year. We find that topics excluded from a given report but included by other analysts do not meaningfully increase the explanatory power of analysts' quantitative forecasts. Specifically, Panel A of Figure X shows that adding excluded topics discussed by peers increases the in-sample R^2 only slightly, with almost 90% of the explainable variation in price target forecasts stemming from topics analysts include in their own reports. These results hold consistently across different report lengths. Thus, while topics included contribute to explaining analysts' quantitative forecasts, topics excluded but discussed by others play a smaller role, suggesting that the mental models we capture are in fact complete.

[42] To assess the extent to which varying the granularity of topic definitions affects the extent of topical overlap, we calculate the semantic similarity between topics using cosine similarity scores based on OpenAI-generated text embeddings of each topic labels. Appendix Table H.6 illustrates this procedure for the topic *antitrust*. We find that scores of 0.85 and above typically identify logically similar topic pairs. In contrast, scores around 0.80 often group conceptually distinct topics. We use these thresholds as benchmarks for evaluating the robustness of our findings. We also conduct a "kill test" to determine the level of topic coarsening required for at least 90% of analyst pairs to achieve a Jaccard score of 1. The resulting threshold is implausibly low, requiring the aggregation of topics with similarity scores as low as 0.76. This is akin to grouping *antitrust* with *commodity prices* and *housing demand*.

5.2 Mental Models are Sparse

Next, we explore whether the sparsity of analysts' mental models is warranted by fundamentals. To do so, we analyze whether omitted topics can help explain analysts' forecast accuracy. Panel B of Figure X compares the R^2 obtained from regressing the size of analysts' forecast errors on included topics alone versus including both included and excluded topics. Excluded topics account for almost 40% of the explainable variation in analysts' forecast accuracy, again consistently across all report lengths. The fact that omitted topics do not meaningfully help in explaining analysts' quantitative forecasts, but instead contribute substantially to explaining their forecast errors, suggests that the sparse nature of analysts' mental model is not justified by fundamentals and may instead be driven by limited cognitive resources.

5.3 Mental Models are Rigid (and Adjust with Large Forecast Errors)

Having provided evidence on mental model sparsity, this section examines how rigid analysts' mental models are, both in the time series and in the cross-section. Along the time-series dimension, we assess the persistence of analysts' mental models and ask what, if anything, leads them to revise them. Similarly, along the cross-sectional dimension, we are interested in understanding whether analysts adopt a similar mental model across all stocks they cover, or whether they tailor their models to each firm.

Mental Model Persistence Over Time. Panel A of Table VI shows that, when covering the same stock over time, there is more than an 80% year-on-year overlap in valuation methods (DCF vs. multiples-based approach), and a 40% overlap in the topics analysts discuss. To assess whether this level of persistence reflects objective changes in fundamentals, or analyst-specific rigidity in mental models, we use other analysts' topic coverage as a benchmark.[43] Panel A of Table VI then shows that about 37% of the year-on-year valuation method overlap, and 20% of the year-on-year topic overlap (i.e., Jaccard similarity), stem from idiosyncratic analyst rigidities, with the remainder reflecting changes in the information

[43] Specifically, we assess the similarity in mental models between an analyst's report on a given firm in a particular year and reports on the same firm in the following year, regardless of whether the subsequent reports are written by the same analyst or by others. Table VI then shows that when pairing an analyst with themselves rather raises the overlap in valuation methods by 0.33, relative to a baseline overlap of 0.49, and the overlap in topics covered by 0.08, relative to a baseline overlap of 0.32.

environment that are common to all analysts.[44]

Despite this rigidity, Panel B of Table VI shows that analysts shift their topic focus by 2.5% more following large forecast errors, including when we estimate this using within-firm-year differences. Instead, large switches in analysts' choice of valuation methods are more likely to occur in unison. The first two columns in Panel B of Table VI show that, while analysts are more likely to change valuation methods following periods with large forecast errors, these patterns are not robust to including firm-year fixed effects. This likely reflects the fact that a large fraction of analysts tend to adjust their valuation methods following large events, such as during the financial crisis, as least in part due to practical constraints in applying certain valuation methods at these points in time. Consistent with this, Panel B of Figure III shows a strong correlation between the use of specific valuation methods over time and the presence of positive earnings per share estimates.

Taken together, this evidence indicates that analysts' mental models exhibit substantial inertia: while a meaningful share of the year-on-year stickiness is common across analysts, a substantial portion reflects individual specific styles and model rigidity. Moreover, larger forecast errors can contribute to analysts making bigger changes to their mental models.

Mental Model Persistence in the Cross-section. Table VII shows that there is more than a 70% congruence in valuation methods, and a 30% overlap in topics, across stocks covered by a given analyst in the same year. Table H.11 shows how these overlaps increase further for firms located in the same country or operating in the same industry (Panel A), and decrease with greater industry distance (Panel B).

To understand whether this level of rigidity reflects analyst-specific styles or whether it is warranted by the stocks an analyst covers having similar fundamentals, we perform a similar exercise to that we used in the time series and we once again use comparisons with other analysts as a benchmark. The estimates in Table VII show that 33% of the across-stock valuation-method overlap, and 13% of the across-stock topic overlap come from analyst

[44] The specifications with fixed effects in the even-numbered columns provide the tightest possible estimation, holding firm-pair and year constant and comparing residual differences between cases where the analyst is matched to themselves in the following year versus cases where a different analyst covers the same firm. Table H.8 presents additional evidence indicating analyst-specific valuation method styles, showing that analyst fixed effects account for more variation in valuation method choices than year, brokerage, or firm fixed effects.

rigidities.[45]

Taken together, this evidence indicates that analysts' mental models are rigid in the cross-section as well. In light of our theoretical framework, this may reflect synergies in paying attention to the same topic across firms: for example, if an analyst is paying attention to inflation when covering one stock, then the marginal cost of paying attention to inflation when covering another stock is lower.

6 Changes in Valuations Over Time and Disagreement

Having established the key properties of analysts' mental models, we next analyze how their different components shape analysts' quantitative forecasts, and provide evidence on which topics contribute most to i) changes in valuation over time, ii) disagreement, and ii) over or underreaction to information.

To do so, we start from the theoretical framework in Section 2, according to which price targets can be written as:

$$\mathbb{E}_t^i[p_{ft}] = \sum_{k=1}^{K} m_{kft}^i \left(s_{kft} - \mu_{kft}\right) a_{kft}^i + \sum_{k=1}^{K} m_{kft}^i \, \mu_{kft} \tag{18}$$

By re-arranging this expression, we can re-write it in terms of expected price appreciation (or expected returns):

$$\mathbb{E}_t[\hat{p}_{f,t+1}^i] \equiv \mathbb{E}_t^i[p_{f,t+1}] - \sum_{k=1}^{K} m_{kft}^i \, \mu_{kft} = \sum_{k=1}^{K} m_{kft}^i \left(s_{kft} - \mu_{kft}\right) a_{kft}^i \tag{19}$$

By interacting our measure of attention with sentiment, our data allows us to obtain a proxy for $\alpha_{kft}^i \equiv a_{kft}^i(s_{kft} - \mu_{kft})$. Specifically, our measure of attention proxies for $|\alpha_{kft}^i| \equiv a_{kft}^i|(s_{kft} - \mu_{kft})|$, meaning that an analyst talks about a topic more either because they pay more attention to it, or because there is a greater news component associated with it. Our measure of sentiment then allows us to further sign this quantity, so that the interaction

[45] Furthermore, as shown in Table H.12, cross-sectional rigidity in analysts' reports persists for technical, less narrative-oriented topics (e.g., covenants and inventory management, as opposed to market share and expansion), which may align more naturally with internal belief formation than solely with persuasive messaging.

of these two components provides us with a proxy for $\alpha^i_{kft} \equiv a^i_{kft}(s_{kft} - \mu_{kft})$.[46] To highlight the components of (19) that are observable, we re-write it as:

$$\mathbb{E}^i_t[\hat{p}_{f,t+1}] = \sum_{k=1}^{K} m^i_{kft}\, \alpha^i_{kft} \tag{20}$$

where $\alpha^i_{kft} \equiv (s_{kft} - \mu_{kft})a^i_{kft}$. Building on this expression, we can then analyze the key drivers of changes in valuation, disagreement, and forecast errors. In order to limit the dimensionality of the variables we include on the RHS of the regressions that follow, and to improve interpretability, we group our 139 topics into the 32 categories described in Table E1.

6.1 Changes in Valuations Over Time: how analysts represent features plays a bigger role than how they price those features

Given the theoretical framework and equation (20) above, we can write changes in valuations over time as:

$$\mathbb{E}^i_t[\hat{p}_{ft+1}] - \mathbb{E}^i_{t-1}[\hat{p}_{ft}] = \sum_{k=1}^{K} \left(m^i_{kft}\Delta\alpha^i_{kft} + \Delta m^i_{kft}\alpha^i_{kft-1} \right) \tag{21}$$

where $\alpha^i_{kft} \equiv (s_{kft} - \mu_{kft})\, a^i_{kft}$ and $\Delta\alpha^i_{kft} = (s_{kft} - \mu_{kft})\Delta a^i_{kft} + \Delta s_{kft}\, a^i_{kft}$. Thus, changes in valuation over time may reflect either shifts in how analysts represent the firm (changes in attention and perceived changes in fundamentals) or in how they price different attributes (changes in valuation weights). Once we estimate this as a regression, Figure XI shows how much each component contributes to changes in price targets over time.[47] Specifically, Panel A of Figure XI plots how much each category contributes to changes in valuation over time due to changes in how analysts represent it ($m^i_{kft}\Delta\alpha^i_{kft}$), while Panel B shows each category's contribution due to changes in how analysts price it ($\Delta m^i_{kft}a^i_{kft-1}$).

Three patterns are worth highlighting. First, when we aggregate the two channels of interest across all categories, changes in representations account for a bigger source of valuation

[46] Strictly speaking, we think of our attention measure as capturing news in terms of percentage changes: $\left(\frac{s_{kft}-\mu_{kft}}{\mu_{kft}}\right) a^i_{kft}$. This does not change the interpretation of the results that follow, as we can multiply and divide the relevant coefficients by μ_{kft}.

[47] To estimate this regression, we exploit the fact that our data collection provides us with proxies for both $|(s_{kft} - \mu_{kft})a^i_{kft}|$ and the associated sentiment. We interact these two components to obtain proxies for $(s_{kft} - \mu_{kft})a^i_{kft}$. See the caption of Figure XI for additional details.

changes over time than do changes in valuation weights (64% vs. 36%).[48] These patterns are consistent with the results in Section 5.3 showing that analysts are more rigid in their choice of valuation method than in their coverage of different topics over time. Second, while the above pattern typically holds at the topic level as well, there are two exceptions: innovation/R&D and environment. For these two categories, changes in valuation weights contribute just as much as changes in representations. This is consistent with recent trends showing increased investor focus on intangible assets and ESG considerations, which may prompt analysts to adapt both how frequently they discuss these topics and how they factor them into valuation models.[49] Third, when looking at the overall contribution of each category, we notice that attention to firm-related topics contribute the most to analysts increasing their price targets, while macro related topics contribute the most to analysts revising their price targets down.

6.2 Disagreement: how analysts attend to features plays a bigger role than how they price those features

Next, we turn to studying the drivers of disagreement across analysts forecasting the same firm at the same point in time. Given our framework, we can express disagreement among analyst pairs as:

$$\mathbb{E}_t^i[\hat{p}_{ft+1}] - \mathbb{E}_t^j[\hat{p}_{ft+1}] = \sum_{k=1}^{K} \left(m_{kft}^i(\alpha_{kft}^i - \alpha_{kft}^j) + (m_{kft}^i - m_{kft}^j)\alpha_{kft}^j \right) \tag{22}$$

where $\alpha_{kft}^i \equiv (s_{kft} - \mu_{kft})\, a_{kft}^i$ and $\alpha_{kft}^i - \alpha_{kft}^j = (a_{kft}^i - a_{kft}^j)(s_{kft} - \mu_{kft})$. The expression in (22) shows that analysts can disagree either because they use different valuation weights $(m_{kft}^i - m_{kft}^j)$ and/or because they use different attention weights across variables $(a_{kft}^i - a_{kft}^j)$. We start by showing how each component contributes to disagreement, and then consider their relative roles.

Columns 1 to 4 of Table VIII show that differences in valuation methods and differences in topic and argument alignment both individually contribute to disagreement.[50] The latter

[48] This decomposition is based on each panel's share of the sum of coefficients (in absolute value) across Panels A and B. In each panel, we can reject that all coefficients are jointly zero ($p < 0.0001$).

[49] See Pástor et al. (2021) for a recent review on sustainable investing.

[50] Table H.9 shows robustness to alternative topic classifications, grouping topics that are semantically sufficiently close based on either a 85% or 80% similarity cutoff.

holds even with analyst-firm fixed effects and when, similarly to Table V and shown in Table H.10, estimated on the subsample of analysts who have consistently used the current valuation method for the firm. The results in columns 5 and 6 of Table VIII then further show that when jointly examining the roles of valuation method overlap and topic and argument overlap in explaining forecast disagreement, topic and argument similarity are associated with larger effect sizes. In other words, differences in attention weights seem to have a larger impact on disagreement than differences in valuation methods.

To corroborate this conclusion, and to shed further light on which topics contribute most to disagreement, we run the regression outlined in equation (22). We do so by considering analyst pairs (forecasting the same firm at the same point in time) where one analyst bases their price target on a multiples-based valuation (without any mention of DCF), and the other on a DCF-based valuation (without any mention of multiples). We systematically sort the pairs as (multiples, DCF), so that the difference in valuation weights captures the effect of multiples- relative to DCF-based methods for the same firm on the same year.[51]

Figure XII then shows how much each component contributes to disagreement. Specifically, Panel A plots how much each category contributes to disagreement due to differences in attention weights $\left(m^i_{kft}(\alpha^i_{kft} - \alpha^j_{kft})\right)$, while Panel B shows each category's contribution due to differences in how analysts price them $\left((m^i_{kft} - m^j_{kft})\alpha^j_{kft}\right)$. Comparing the two panels of Figure XII confirms that differences in how analysts attend to various variables contribute substantially more to overall disagreement than differences in how they price them (83% vs. 17%).[52] Panel A of Figure XII then shows that it is specifically attention to firm-related categories that generates the most disagreement, while macro-related categories play a smaller role. Turning to Panel B, we notice that there are two different types of categories that contribute to disagreement via differences in valuation weights. The first type is once again firm related, and includes product, production/utilization, and governance, indicating that analysts do disagree on how to price topics related to the firm's core operations. The second type includes categories associated with major trends or events which have generally

[51] For completeness, Figure H.20 reports results when analyst pairs are systematically sorted as (DCF, multiples), such that differences in valuation weights capture the effect of DCF- relative to multiples-based methods. The results remain virtually identical.

[52] Again, we can reject that all coefficients in each panel are jointly zero ($p < 0.0001$ and $p = 0.0016$, respectively).

been associated with displacements, bubbles and crashes (Kindleberger, 1978, Pástor and Veronesi, 2009, Bastianello and Fontanier, 2025):[53] the positive and significant coefficient on innovation/R&D (which includes topics like AI and cognitive applications, technology lifecycle, trade secrets and patents, clinical trials) suggests that there is disagreement in how analysts price new products and technologies, while the significant coefficient for housing market (which includes mortgage rates, home price, housing demand, and housing supply) is consistent with disagreement on these issues in the lead up to the financial crisis.

Overall, our analysis suggests that both differences in valuation weights and differences in how analysts represent firms contribute to disagreement. While the latter component plays a larger overall role, there are meaningful differences in how analysts price topics that have traditionally been associated with mispricing, such as innovation and housing markets.

7 Biased Beliefs and Asset Pricing

This section explores the extent to which analysts' mental models are associated with biased beliefs, and how such biases translate into asset prices.

7.1 Which Topics Drive Over- and Underreaction?

While it is well-known that analysts' price targets are on average over-optimistic (leading to negative forecast errors on average, as shown in in Panel C of Table I), in this section we study the extent to which different components of analysts' mental models contribute to over- or underreaction to information.

One of the standard approaches to study over- and underreaction comes from running Coibion-Gorodnichenko (CG) regressions, which involve regressing forecast errors on forecast revisions. A positive coefficient is usually interpreted as evidence of underreaction: although forecasters revised their beliefs upward (downward), they did not update up (down) sufficiently. Conversely, a negative coefficient is generally interpreted as evidence of overreaction, as it implies that forecasters updated their beliefs too strongly in the direction of the news.

[53] See Brunnermeier and Oehmke (2013), Xiong (2013), Barberis (2018), and Sufi and Taylor (2022) for exhaustive surveys on bubbles and crashes, and financial crises.

The benefit of running CG regressions is that they do not require knowledge of the underlying data generating process. However, they do require observing both forecast errors and forecast revisions for a fixed end date. Because analysts consistently provide 12-month ahead price targets, we do not observe the relevant forecast revisions. Therefore, we are not generally able to run such regressions for price targets.

Our data collection efforts allow us to make significant progress on this question by exploiting our measure of sentiment, and running the following CG-like regression:

$$\hat{p}_{ft+1} - \mathbb{E}_t^i[\hat{p}_{ft+1}] = \sum_{k=1}^{K} \left(\beta_k^+ \alpha_{kft}^{i+} + \beta_k^- \alpha_{kft}^{i-} \right) + \phi_f + \epsilon_{fkt} \tag{23}$$

where $\alpha_{kft}^{i+} \equiv \alpha_{kft}^i \times \mathbf{1}(s_{kft} - \mu_{kf} > 0)$ and $\alpha_{kft}^{i-} \equiv \alpha_{kft}^i \times \mathbf{1}(s_{kft} - \mu_{kf} < 0)$, while β_k^+ and β_k^- are our coefficients of interest, and ϕ_f are firm fixed effects. Importantly, our measure of sentiment indicates whether the specific argument an analyst puts forward contributes to an increase or decrease in the associated price target. In other words, it provides information about the direction of the analyst's valuation revision. Conditional on knowing that analysts revised their beliefs up, a positive forecast error is then indicative of underreaction (the analyst updated upward but not enough), while a negative forecast error is indicative of overreaction (the analyst updated upward but too much). Similarly, the same logic applies in the case of downward revisions.

Therefore, not only does our data allow us to run a variant of the traditional CG regression on price targets (which prior literature was not able to do), but we are also able to provide a much more granular picture as to the exact types of information to which analysts tend to over- or underreact to when valuing stocks.

Figure XIII shows the results from this exercise once we aggregate positive and negative coefficients. There are two sets of results worth highlighting. First, the figure shows that analysts tend to overreact to firm related news and underreact to macro-related news. Most notably, we observe that they underreact to monetary policy and rates, consistent with with evidence in prior work that has shown prevalence of underreaction in forecasts about interest rates (Bordalo et al., 2020b). Moreover, overreaction to firm-related topics and underreaction to macro-related topics both contribute to overly optimistic forecasts, on average. To see this,

recall that Figure XI showed that increased attention to firm-related variables contributes to higher price targets, while increased attention to macro-related variables decreases price targets. Therefore, amplifying the upward movement and dampening the downward ones both push forecasts towards being on average too optimistic.

Second, we wish to revisit the three topics discussed in prior sections: innovation/R&D, housing market, and environment. We observe overreaction to the first two topics (consistent with these having been associated with bubbles and crashes by prior work), while we observe no systematic bias in how analysts process information about the environment. One potential explanation for the lack of over or underreaction to this latter topic may be that analysts correctly predict investors' mispricing of it. This would also be consistent with the results in Figure XII, which shows relatively little disagreement among analysts on this dimension (both in terms of attention and pricing). Intuitively, if analysts are aware of mispricing related to environment-related news, they will correctly price this into their valuation. Instead, if analysts disagree on the pricing of a given topic (as with innovation/R&D and housing market), then they will not all incorporate it into their valuation in the same way, leading to biases in their forecasts.

7.2 Asset Pricing Implications: topics that are over-(under-)reacted to predict lower (higher) realized returns

We end our analysis by exploring how biases in analysts' beliefs contribute to return predictability. Intuitively, if analysts overreact (underreact) to a given topic, we should expect that topic to be associated with lower (higher) realized returns.

To investigate this hypothesis, we regress realized returns on our measure of attention to each category (interacted with sentiment):

$$\hat{p}_{ft+1} = \sum_{k=1}^{K} \beta_k^+ \alpha_{kft}^{i+} + \beta_k^- \alpha_{kft}^{i-} + \gamma X_{ft} + \phi_f + \epsilon_{ft+1} \tag{24}$$

where $\alpha_{kft}^{i+} \equiv \alpha_{kft}^{i} \times \mathbf{1}(s_{kft} - \mu_{kf} > 0)$ and $\alpha_{kft}^{i-} \equiv \alpha_{kft}^{i} \times \mathbf{1}(s_{kft} - \mu_{kf} < 0)$, X_{ft} includes a series of known predictors of returns, namely book-to-market ratio, earnings-to-price ratio, prior 12-month return (excluding the most recent month), asset growth, and operating profitability

(Jensen et al., 2023), and ϕ_f are firm fixed effects. Finally, β_k^+ and β_k^- capture our coefficients of interest.

Figure XIV shows the results from this regression (with and without the additional control variables described above), together with the associated over/underreaction coefficients from the CG-like regression we ran in Section 7.1, now split across positive and negative coefficients. The results are broadly consistent with our hypothesis that topics that analysts overreact to more tend to be associated with lower realized returns going forward, while topics that analysts underreact to more tend to be associated with higher realized returns going forward. Specifically, we estimate robust positive correlations between the estimates of β_k^+ and β_k^- in Figure XIV and the corresponding coefficients from the CG forecast error regression capturing over- and underreaction, both without the predictor variables ($corr^+ = .81,\ corr^- = 0.79$) and with the predictors included ($corr^+ = 0.74,\ p < 0.001;\ corr^- = 0.65,\ p < 0.001$).

Relatedly, the predictive power of our topic variables for subsequent realized returns is both statistically and economically meaningful. We reject the null that all estimated topic coefficients are jointly zero, with a p-value below 0.001.[54] Economically, the set of included predictor variables X_{ft} yields an R^2 of 0.014 for realized returns (excluding firm fixed effects), compared to an R^2 of 0.012 when we include just our topic variables. As another way to quantify this, a one–standard deviation change in all predictor variables (in the direction associated with positive returns) corresponds to a 14% increase in realized returns in the regression underlying Panel B of Figure XIV. In comparison, a one–standard deviation change in all topic category variables (in the direction aligned with the predicted over- and underreaction based on the CG regression) is associated with an *additional* 7.8% increase in returns.

Broadly, these associations between mental models and asset prices point to promising directions for linking the structure of mental models to the formation of market expectations

[54] The joint significance test is appropriate here (consistent with the approaches in Figures XI and XII above) as we are interested in whether the topic variables have predictive power in combination. Individual statistical tests may underestimate significance in the presence of other, relatively similar topic categories.

and to the study of return predictability and mispricing in an interpretable way.[55]

8 Discussion and Conclusion

This paper analyzes the text of the near-universe of equity reports to measure the mental models underlying equity analysts' quantitative forecasts. We define mental models as a combination of valuation methods and attention weights over different arguments—with each argument defined by a topic and its associated valuation channel, time outlook, and sentiment. We show that the choice of valuation methods and the allocation of attention across variables are tightly interlinked, shaping variation in analysts' quantitative forecasts and disagreement, while also correlating with key asset pricing patterns.

Having explored many dimensions of analysts' reasoning, this paper also illustrates the granularity of the data that can be collected from these reports. Future work may zoom into certain topics and explore how analysts reason about specific variables or events. For example, Figure H.2 shows results from analyzing analysts' reasoning about inflation during its most recent surge. To do so, we instruct an LLM to identify all lines of reasoning that relate to inflation, and to classify them according to the narratives identified in Andre et al. (2024a) (see Appendix D for the prompt instructions). We find that analysts overwhelmingly have a supply-side view of inflation, and that they are much more likely to reason about inflation through costs than sales. This type of analysis further illustrates the benefits of our data and methodology in complementing and extending prior survey work. By offering both a rich cross-section and an extended time series, our data and methodology open the door to a new wave of empirical work on how financial agents interpret, explain, and respond to

[55] In the time-series, we find evidence of significant co-movement between sentiment in analyst reasoning and Shiller's CAPE index (Panel F of Figure VI). In the cross-section, Figure H.21 shows how average sentiment across analysts varies across bins in various cross-sectional sorts: CAPM beta, market-to-book, size, momentum, subsequent-year return, CapEx investment, and profitability. Broadly, these patterns are consistent with those of realized returns. However, our dataset allows us to go beyond sentiment. Figure H.22 shows how average time outlook varies across the same set of cross-sectional characteristics. Interestingly, time outlook is more forward-looking for growth firms than for value firms. Additionally, time outlook is increasingly forward-looking in firm investment. Finally, Figure H.23 shows even more granular patterns by examining the difference in attention focus across categories for the highest and lowest sorts of each factor. Growth stocks attract more attention around product strategy, brand, and competition than value stocks. Similarly, large stocks attract more attention to payout policy and FX, reflecting larger cash positions, more maturity, and a more international focus than small stocks.

economic developments, and how their reasoning can help predict financial market outcomes.

References

Adam, K., and S. Nagel. 2023. Expectations data in asset pricing. In *Handbook of economic expectations*, pp. 477–506. Elsevier.

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar. 2023. Overreaction in expectations: Evidence and theory. *The Quarterly Journal of Economics* 138:1713–1764.

Amromin, G., and S. A. Sharpe. 2014. From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science* 60:845–866.

Andre, P., I. Haaland, C. Roth, M. Wiederholt, and J. Wohlfart. 2024a. Narratives about the Macroeconomy. Tech. rep., SAFE Working Paper.

Andre, P., C. Pizzinelli, C. Roth, and J. Wohlfart. 2022. Subjective models of the macroeconomy: Evidence from experts and representative samples. *The Review of Economic Studies* 89:2958–2991.

Andre, P., P. Schirmer, and J. Wohlfart. 2024b. Mental Models of the Stock Market. Tech. rep., Working Paper.

Asquith, P., M. B. Mikhail, and A. S. Au. 2005. Information content of equity analyst reports. *Journal of financial economics* 75:245–282.

Ba, C., J. A. Bohren, and A. Imas. 2022. Over-and underreaction to information. *Available at SSRN 4274617* .

Bacchetta, P., E. Mertens, and E. Van Wincoop. 2009. Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance* 28:406–426.

Bailey, M., E. Dávila, T. Kuchler, and J. Stroebel. 2019. House price beliefs and mortgage leverage choice. *The Review of Economic Studies* 86:2403–2452.

Barberis, N. 2018. Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 1, pp. 79–175. Elsevier.

Bastianello, F. 2025. Time-series and cross-section of risk premia expectations: Evidence from Financial Analysts. *Available at SSRN 4204968* .

Bastianello, F., and P. Fontanier. 2025. Partial equilibrium thinking, extrapolation, and bubbles. *Working Paper* .

Bastianello, F., and A. O. Imas. 2025. Underreaction and Overreaction in Inference and Forecasting. *Working Paper* .

Bauer, R., K. Gödker, P. Smeets, and F. Zimmermann. 2024. Mental Models in Financial Markets: How Do Experts Reason about the Pricing of Climate Risk? *European Corporate Governance Institute–Finance Working Paper* .

Ben-David, I., and A. Chinco. 2024. Expected EPS× Trailing P/E. Tech. rep., National Bureau of Economic Research.

Bianchi, F., S. C. Ludvigson, and S. Ma. 2024. What hundreds of economic news events say about belief overreaction in the stock market. Tech. rep., National Bureau of Economic Research.

Binetti, A., F. Nuzzi, and S. Stantcheva. 2024. People's Understanding of Inflation. Tech. rep., National Bureau of Economic Research.

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer. 2023. How people use statistics. Tech. rep., National Bureau of Economic Research.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer. 2020a. *Expectations of fundamentals and stock market puzzles*. National Bureau of Economic Research.

Bordalo, P., N. Gennaioli, G. Lanzani, and A. Shleifer. 2024a. A Cognitive Theory of Reasoning and Choice .

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer. 2019. Diagnostic expectations and stock returns. *Journal of Finance* 74:2839–2874.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer. 2020b. Overreaction in macroeconomic expectations. *American Economic Review* 110:2748–2782.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer. 2024b. Finance without exotic risk. *Working paper* .

Bouchaud, J.-P., P. Kruger, A. Landier, and D. Thesmar. 2019. Sticky Expectations and the Profitability Anomaly. *Journal of Finance* .

Bradshaw, M. T. 2002. The use of target prices to justify sell-side analysts' stock recommendations. *Accounting Horizons* 16:27–41.

Brunnermeier, M. K., and M. Oehmke. 2013. Bubbles, financial crises, and systemic risk. *Handbook of the Economics of Finance* 2:1221–1288.

Bybee, J. L. 2023. The ghost in the machine: Generating beliefs with large language models. *Working Paper* .

Bybee, L., B. Kelly, A. Manela, and D. Xiu. 2024. Business news and business cycles. *The Journal of Finance* 79:3105–3147.

Case, K. E., R. J. Shiller, and A. Thompson. 2012. What have they been thinking? Home buyer behavior in hot and cold markets. Tech. rep., National Bureau of Economic Research.

Charles, C., and C. Kendall. 2024. Causal narratives. *Available at SSRN 4669371* .

Charles, C., and P. Sui. 2024. Marketwide Memory. *Working paper* .

Chen, H., B.-H. Hwang, and Z. Peng. 2025. Why do investors like short-leg securities? Evidence from a textual analysis of buy recommendations. Tech. rep.

Chen, Z. 2024. Memory and Beliefs in Financial Markets: A Machine Learning Approach. *Working Paper* .

Chinco, A., S. M. Hartzmark, and A. B. Sussman. 2022. A new test of risk factor relevance. *The Journal of Finance* 77:2183–2238.

Chopra, F., and I. Haaland. 2023. Conducting qualitative interviews with AI .

De Bondt, W. F., and R. H. Thaler. 1990. Do security analysts overreact? *The American economic review* pp. 52–57.

De La O, R., and S. Myers. 2021. Subjective Cash Flow and Discount Rate. *Journal of Finance* .

De Rosa, L. 2024. (Deep) Learning Analyst Memory. Tech. rep., Working Paper.

de Silva, T., E. Larsen-Hallock, A. Rej, and D. Thesmar. 2025. Expectations Formation with Fat-Tailed Processes: Evidence and Theory. Working Paper NBER Working Paper, National Bureau of Economic Research.

de Silva, T., and D. Thesmar. 2024. Review of Financial Studies. *Journal of Finance* 37:1494–1537.

Decaire, P., and M. Guenzel. 2025. What Drives Very Long-Run Cash Flow Growth Expectations? *Working Paper* .

Decaire, P. H., and J. Graham. 2024. Valuation Fundamentals. *Working paper* .

Decaire, P. H., D. Sosyura, and M. Wittry. 2024. Resolving Estimation Ambiguity. *Working paper* .

Demirakos, E. G., N. C. Strong, and M. Walker. 2004. What valuation models do analysts use? *Accounting horizons* 18:221–240.

Engelberg, J., A. Manela, W. Mullins, and L. Vulicevic. 2025. Entity Neutering. Working paper.

Erkilet, G., G. Janke, and R. Kasperzak. 2022. How valuation approach choice affects financial analysts' target price accuracy. *Journal of Business Economics* 92:741–779.

Fan, T., Y. Liang, and C. Peng. 2021. The inference-forecast gap in belief updating. *Unpubished Manuscript* .

Flynn, J. P., and K. Sastry. 2024. Attention cycles. Tech. rep., National Bureau of Economic Research.

Gabaix, X. 2014. A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics* 129:1661–1710.

Gabaix, X. 2019. Behavioral inattention. In *Handbook of behavioral economics: Applications and foundations 1*, vol. 2, pp. 261–343. Elsevier.

Gabaix, X., R. S. Koijen, R. Richmond, and M. Yogo. 2023. Asset embeddings. *Available at SSRN 4507511* .

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein. 2018. *Channeled attention and stable errors*. Harvard Business School Boston.

Giglio, S., and B. Kelly. 2018. Excess volatility: Beyond discount rates. *The Quarterly Journal of Economics* 133:71–127.

Giglio, S., B. Kelly, and D. Xiu. 2022. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics* 14:337–368.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2021. Five facts about beliefs and portfolios. *American Economic Review* 111:1481–1522.

Gleason, C. A., W. Bruce Johnson, and H. Li. 2013. Valuation model use and the price target performance of sell-side equity analysts. *Contemporary Accounting Research* 30:80–115.

Gormsen, N. J., and K. Huber. 2023. Corporate discount rates. Tech. rep., National Bureau of Economic Research.

Graham, J. R. 2022. Presidential address: Corporate finance and reality. *The Journal of Finance* 77:1975–2049.

Greenwood, R., and A. Shleifer. 2014. Expectations of returns and expected returns. *The Review of Financial Studies* 27:714–746.

Haaland, I. K., C. Roth, S. Stantcheva, and J. Wohlfart. 2024. Measuring what is top of mind. Tech. rep., National Bureau of Economic Research.

Hanna, R., S. Mullainathan, and J. Schwartzstein. 2014. Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics* 129:1311–1353.

He, S., L. Lv, A. Manela, and J. Wu. 2025. Chronologically Consistent Large Language Models. Working paper.

Hong, H., J. C. Stein, and J. Yu. 2007. Simple forecasts and paradigm shifts. *The Journal of Finance* 62:1207–1242.

Imam, S., J. Chan, and S. Z. A. Shah. 2013. Equity valuation models and target price accuracy in Europe: Evidence from equity reports. *International Review of Financial Analysis* 28:9–19.

Jensen, T. I., B. Kelly, and L. H. Pedersen. 2023. Is there a replication crisis in finance? *The Journal of Finance* 78:2465–2518.

Kacperczyk, M. T., S. van Nieuwerburgh, and L. Veldkamp. 2016. A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica* 84:571–626.

Ke, B. 2025. Analysts' Belief Formation in Their Own Words. Tech. rep., Working Paper.

Kindleberger, C. P. 1978. *Manias, panics, and crashes; a history of financial crises*. Basic Books, New York.

Kohlhas, A. N., and A. Walther. 2021. Asymmetric attention. *American Economic Review* 111:2879–2925.

Laudenbach, C., A. Weber, J. Wohlfart, and R. Weber. 2024. Beliefs about the Stock Market and Investment Choices: Evidence from a Field Experiment. *Review of Financial Studies* .

Link, S., A. Peichl, C. Roth, and J. Wohlfart. 2023. Attention to the Macroeconomy .

Lopez-Lira, A., and Y. Tang. 2024. Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. *Working paper* .

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar. 2020. A quantitative analysis of distortions in managerial forecasts. Tech. rep., National Bureau of Economic Research.

Maćkowiak, B., and M. Wiederholt. 2009. Optimal sticky prices under rational inattention. *American Economic Review* 99:769–803.

Malmendier, U., and S. Nagel. 2011. Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics* 126:373–416.

Nagel, S., and Z. Xu. 2022. Asset pricing with fading memory. *The Review of Financial Studies* 35:2190–2245.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. Sustainable investing in equilibrium. *Journal of financial economics* 142:550–571.

Pástor, L., and P. Veronesi. 2009. Technological revolutions and stock prices. *American Economic Review* 99:1451–1483.

Sarkar, S. K. 2025. Economic Representations. *Working Paper* .

Sarkar, S. K., and K. Vafa. 2024. Lookahead Bias in Pretrained Language Models. Tech. rep., Harvard Business School, Working Paper.

Schwartzstein, J. 2014. Selective attention and learning. *Journal of the European Economic Association* 12:1423–1452.

Schwartzstein, J., and A. Sunderam. 2021. Using models to persuade. *American Economic Review* 111:276–323.

Sims, C. A. 2003. Implications of rational inattention. *Journal of monetary Economics* 50:665–690.

Stantcheva, S. 2024. Why do we dislike inflation? Tech. rep., National Bureau of Economic Research.

Sufi, A., and A. M. Taylor. 2022. Financial crises: A survey. *Handbook of international economics* 6:291–340.

Van Binsbergen, J. H., X. Han, and A. Lopez-Lira. 2023. Man versus machine learning: The term structure of earnings expectations and conditional biases. *The Review of Financial Studies* 36:2361–2396.

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

Van Nieuwerburgh, S., and L. Veldkamp. 2010. Information acquisition and under-diversification. *The Review of Economic Studies* 77:779–805.

Veldkamp, L. L. 2011. *Information choice in macroeconomics and finance*. Princeton University Press.

Vissing-Jorgensen, A. 2003. Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions. *NBER Macroeconomics Annual* 18:139–194.

Woodford, M. 2001. Imperfect common knowledge and the effects of monetary policy.

Xiong, W. 2013. Bubbles, crises, and heterogeneous beliefs. Working paper, National Bureau of Economic Research.

Figures and Tables

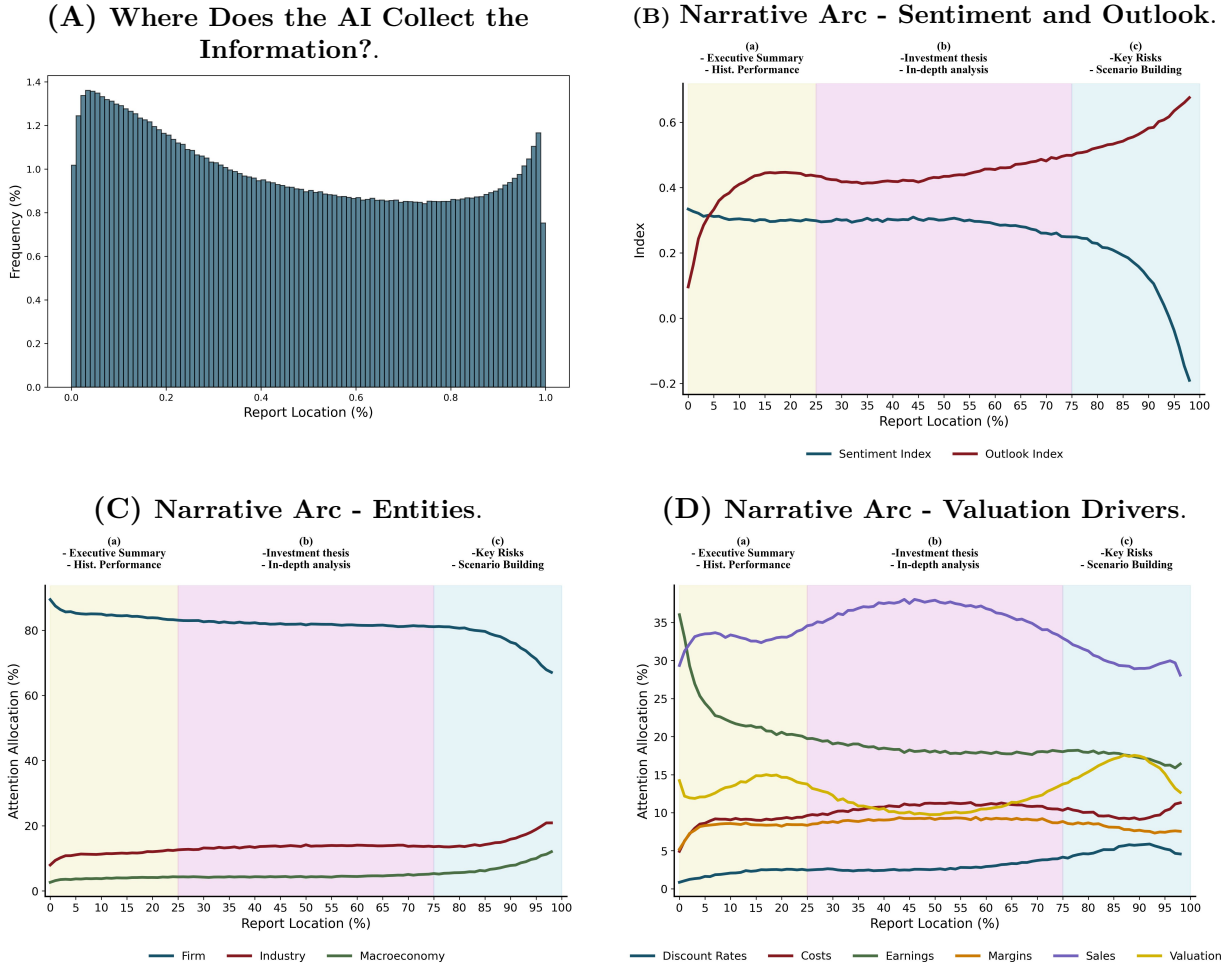


Figure I:
Report Narrative Arc

This figure shows where information is collected within the average report and the type of information collected at each stage. The x-axis for all four panels denotes the position of the report, in percentiles from start to finish. The sample period is 2000–2025 and includes all firms in our main sample. Panel A plots the share of topics extracted at each point in the report. Panel B plots average sentiment and time outlook. Sentiment is calculated using values of −1, 0, and +1 for negative, neutral, and positive sentiment, respectively. Time outlook Index values of −1, 0, +1, and +2 for past, present, near-future, and distant-future outlook, respectively. Panel C plots the allocation of attention to entities (i.e., firm, industry, and the macroeconomy), while Panel D plots attention to valuation drivers (i.e., sales, costs, earnings/cash flows, margins, discount rates, and relative valuation statements). Shaded areas in the figures indicate the three commonly observed sections of report structure, providing intuitive reference points along the x-axis.

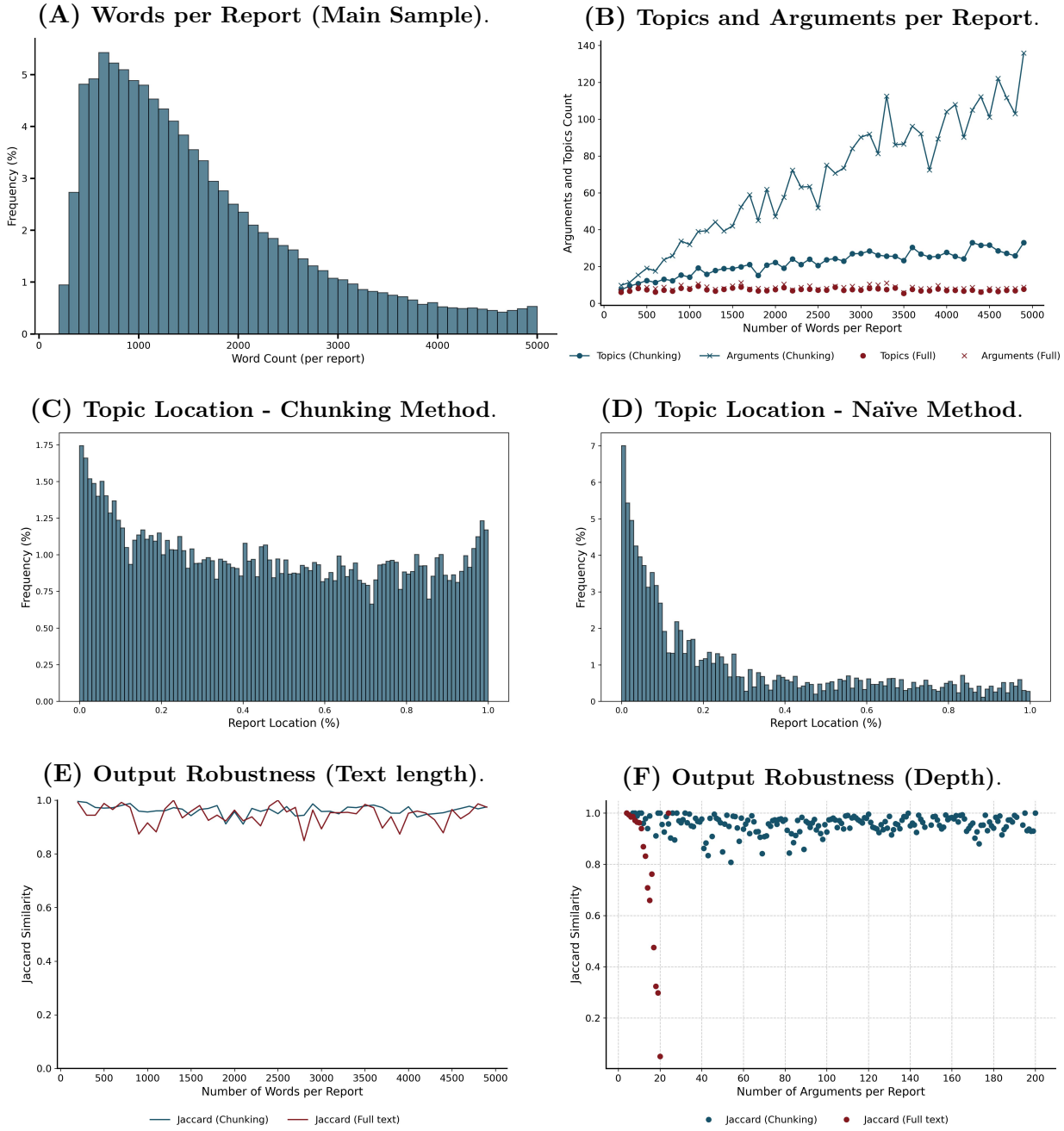
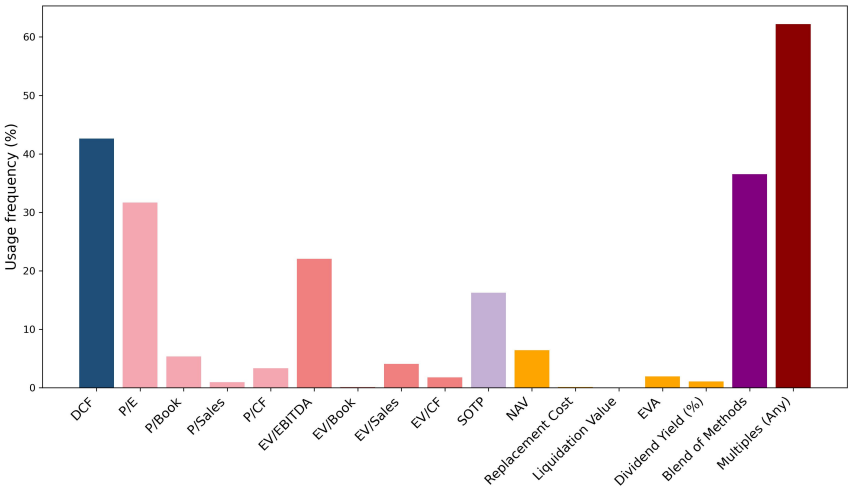


Figure II:
Diagnostic Tools

This figure presents results from diagnostic tests validating the reliability of our LLM-based extraction method. Panel A uses the main sample and plots the distribution of report word counts. Panels B through F use a subsample of 240 randomly selected reports (stratified by report length), each processed 10 times. Panel B shows the number of topics and arguments (combinations of topic, valuation channel, time outlook, and sentiment) as a function of report length, comparing our multi-step extraction method (blue) with a naïve single-step method (red). Panels C and D plot the average position within the report at which topics are collected, for the multi-step and single-step approaches, respectively. Panels E and F report the average topic similarity across the 10 extraction attempts, by report length and number of arguments extracted, respectively.

(A) Valuation Method Frequency.



(B) By Year.

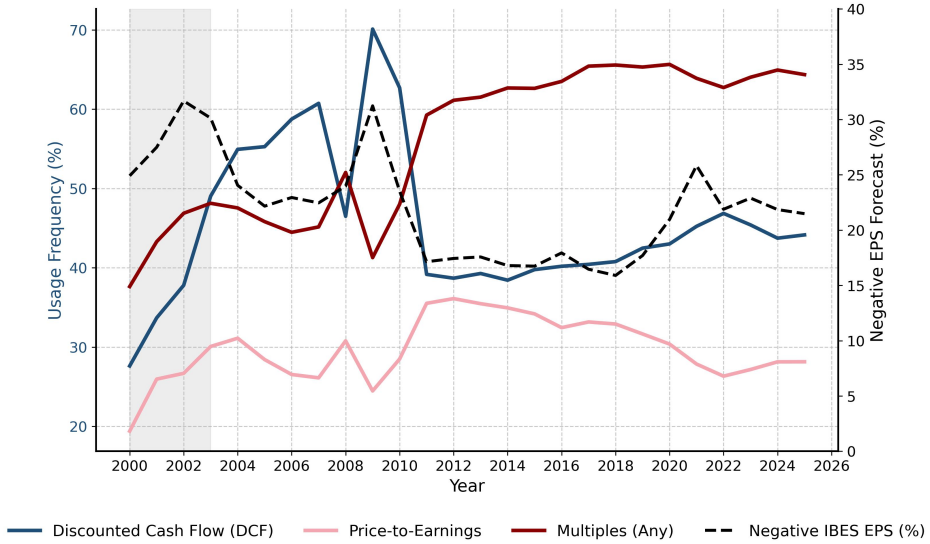
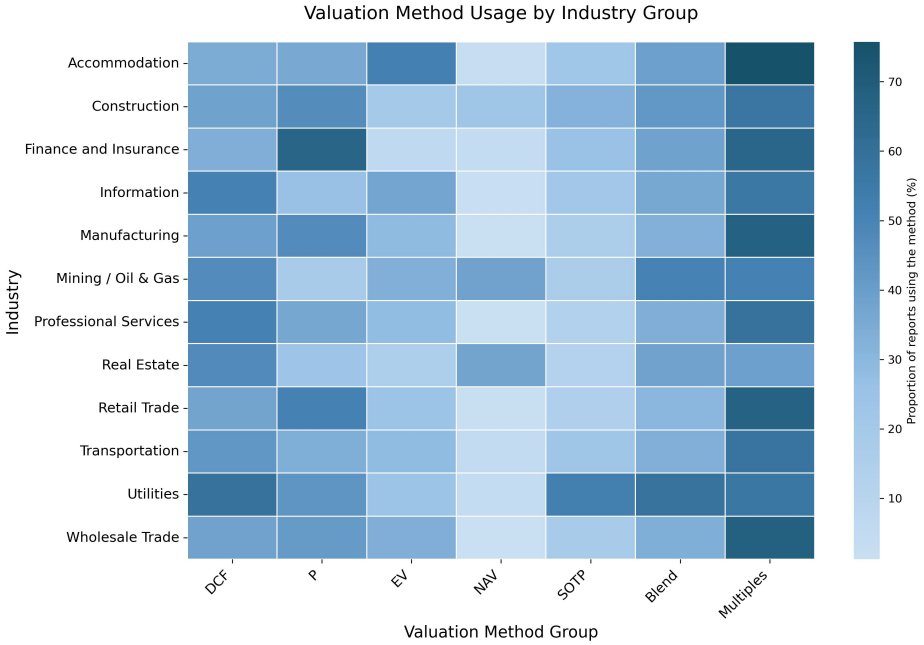


Figure III:
Valuation Methods

This figure shows the usage of valuation methods across the near-universe of equity reports in Refinitiv from 2000 to 2025 for which the LLM extraction identified the price-target valuation method(s) used ($N = 1,379,912$ reports). Panel A reports the frequency of various valuation methods, including DCF, multiples, and blended approaches. Panel B shows variation in the usage of valuation methods over time.

(A) By Industry (10 largest in sample).



(B) By Continent.

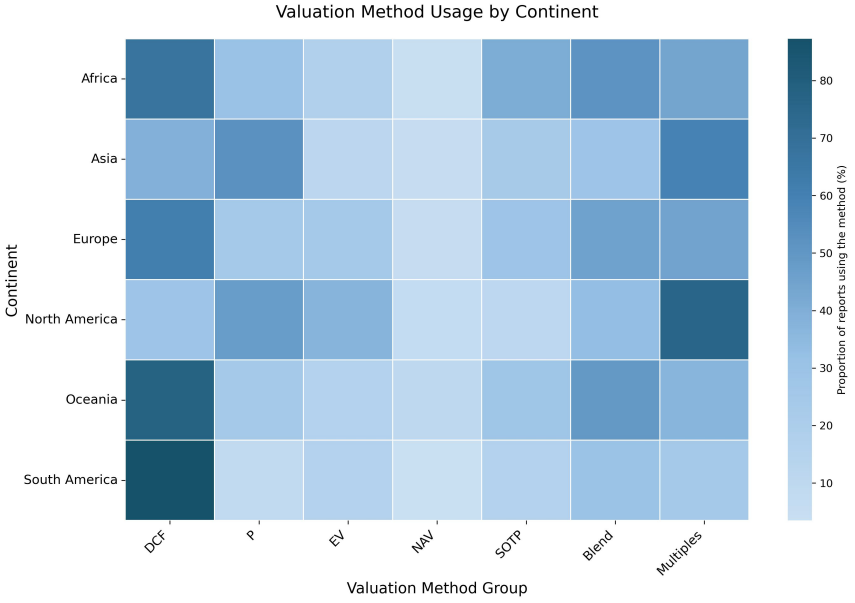
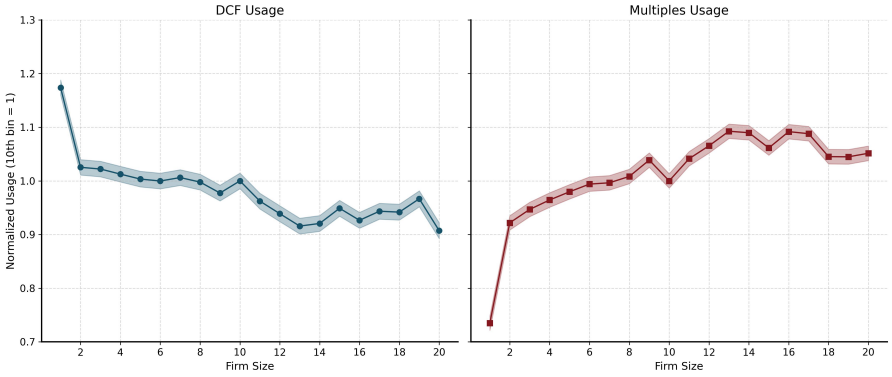


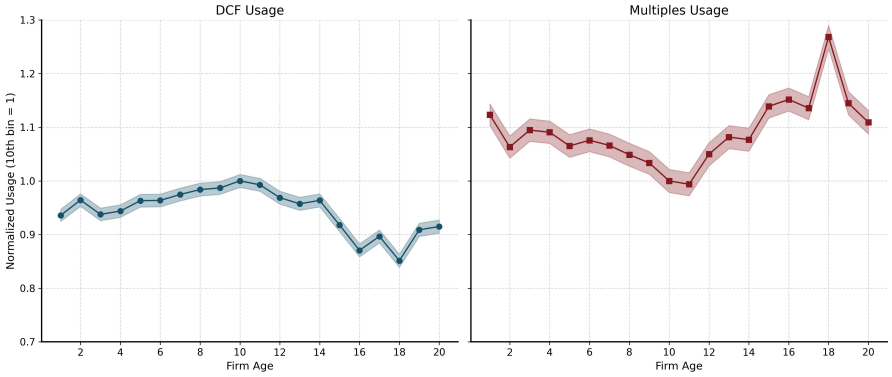
Figure IV:
Valuation Methods – Heterogeneity by Industry and Geography

This figure shows heterogeneities in the usage of valuation methods across the near-universe of equity reports in Refinitiv from 2000 to 2025 for which the LLM extraction identified the price-target valuation method(s) used ($N = 1,379,912$ reports). Panel A shows variation across industries, whereas Panel B shows variation across continents.

(A) Size.



(B) Firm Age.



(C) Market-to-Book.

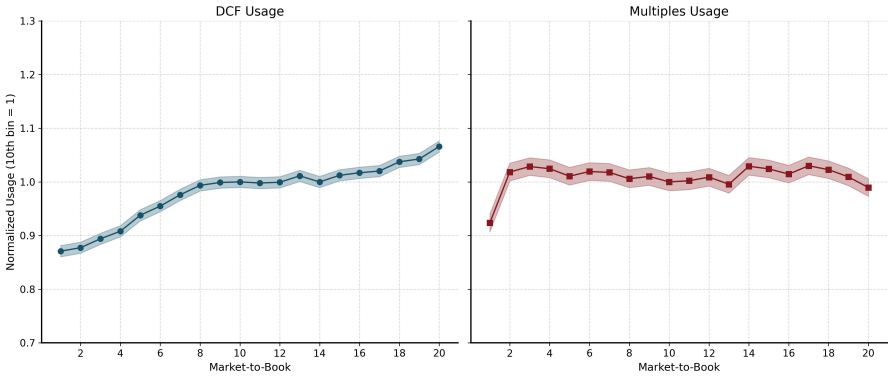


Figure V:
Valuation Method Choice and the Cross-Section of Firms

This figure plots the propensity to use either a DCF method (blue) or a multiples approach (red) as a function of firm characteristics. Panels A, B, and C show results by firm size (total assets), firm age (measured from the year of incorporation), and market-to-book ratio, respectively. For each characteristic, firms are sorted into 20 equal-sized bins (5%) within each year of the sample. The propensity to use valuation methods based on firm characteristics is estimated using the multivariate regression $1_{\text{Valuation Method}} = 1_{\text{Size Bin}} + 1_{\text{Age Bin}} + 1_{\text{Market-to-Book Bin}} + 1_{\text{Negative Expected EPS}}$, where $1_{\text{Valuation Method}}$ is a binary variable equal to 1 if the focal valuation method is used in the report and zero otherwise; $1_{\text{Size Bin}}$, $1_{\text{Age Bin}}$, and $1_{\text{Market-to-Book Bin}}$ are each sets of 20 fixed effects equal to 1 if the firm is included in the respective bin and zero otherwise; and $1_{\text{Positive Expected EPS}}$ takes the value 1 if all of the annual EPS forecasts in IBES are positive for the firm in that year, and 0 otherwise. All series are normalized to equal 1 at ventile 10 to allow for comparison of relative trends. Confidence intervals are at the 90th percentile, and standard errors are heteroskedasticity-robust.

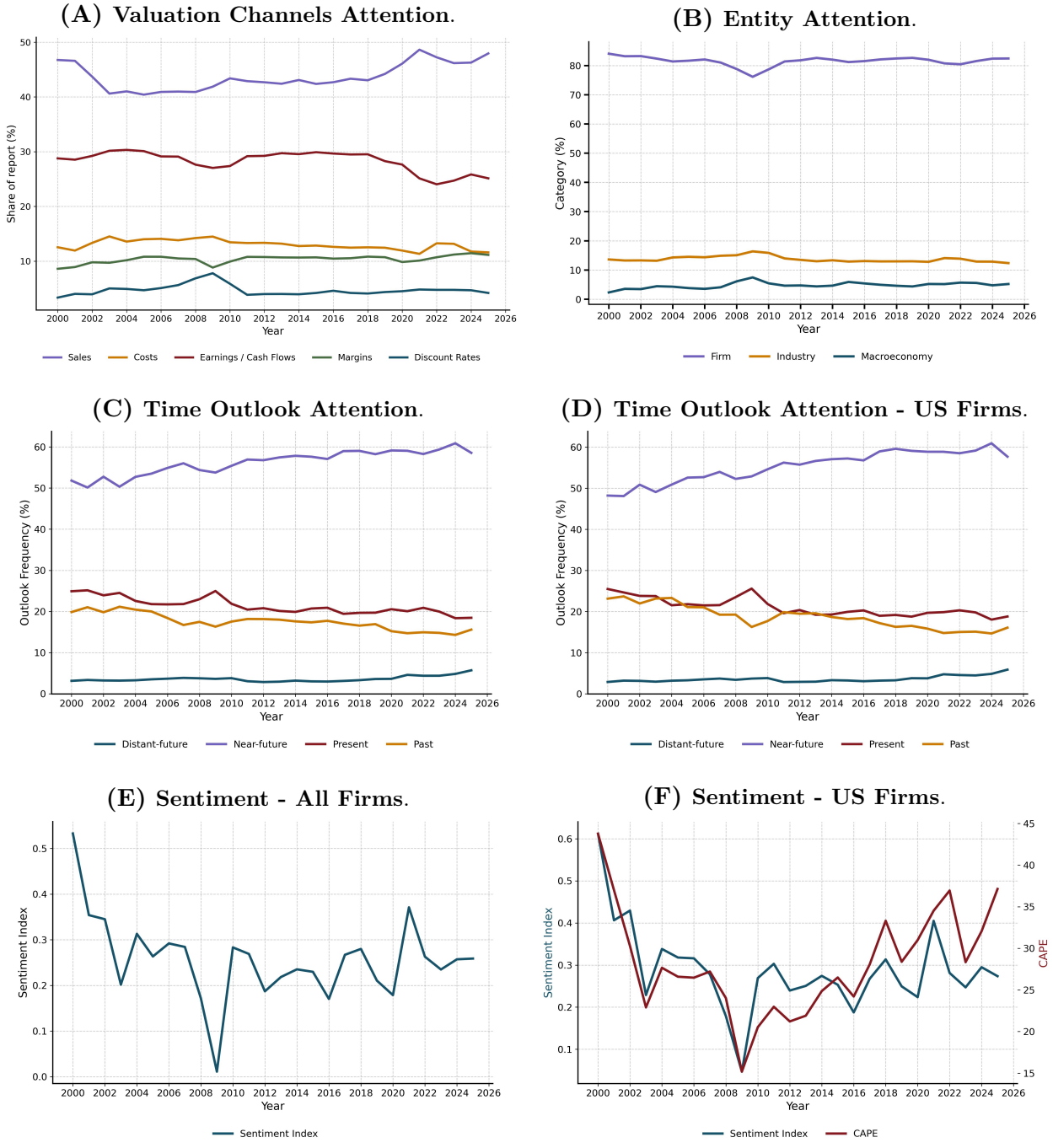


Figure VI:
Valuation Channels, Entities, Outlook, and Sentiment

This figure plots time trends in mental model features across our sample, spanning 2000–2023. The x-axis denotes years in all panels. Panel A shows the share of statements in the average report that reference different valuation drivers—sales, costs, earnings, margins, and discount rates. Panel B reports the share of statements referencing the three entities: firm, industry, and macroeconomy. In Panels C and D, the y-axis indicates the percentage of statements in the average report that discuss topics related to the past, present, near future (< 3 years), and distant future (> 3 years), for all firms and U.S. firms only, respectively. Panels E and F display average sentiment across reports written in a given year, measured on a scale from -1 (negative) to $+1$ (positive), with 0 denoting neutral sentiment. In Panel F, the right-hand y-axis represents the scale for Shiller's CAPE index.

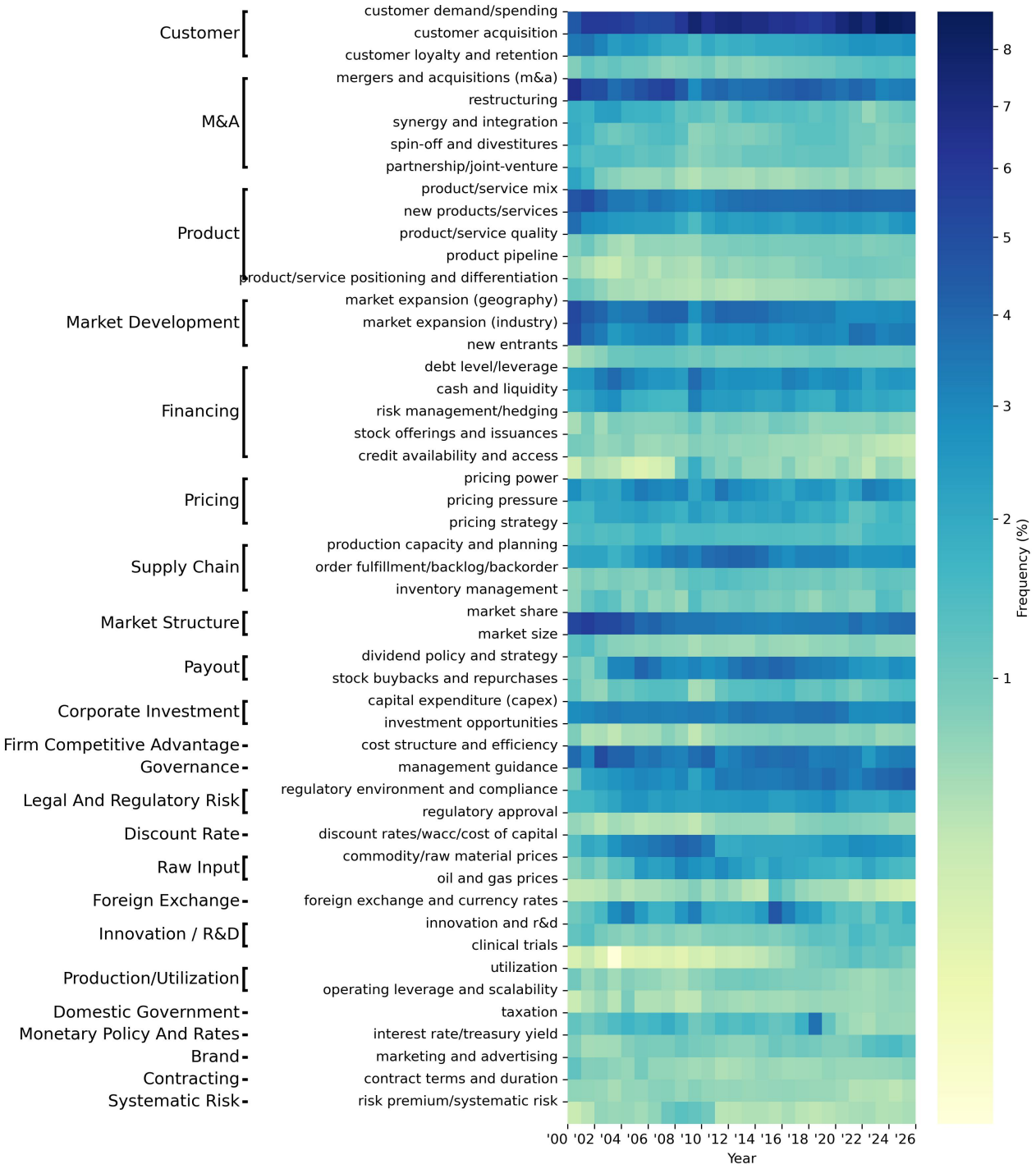


Figure VII:
Topic Attention Allocation Over Time (Top 50 Topics)

This figure plots attention allocation to the 50 most frequently discussed across all firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic. Note that color intensities are not directly comparable between this figure and Figure VIII.

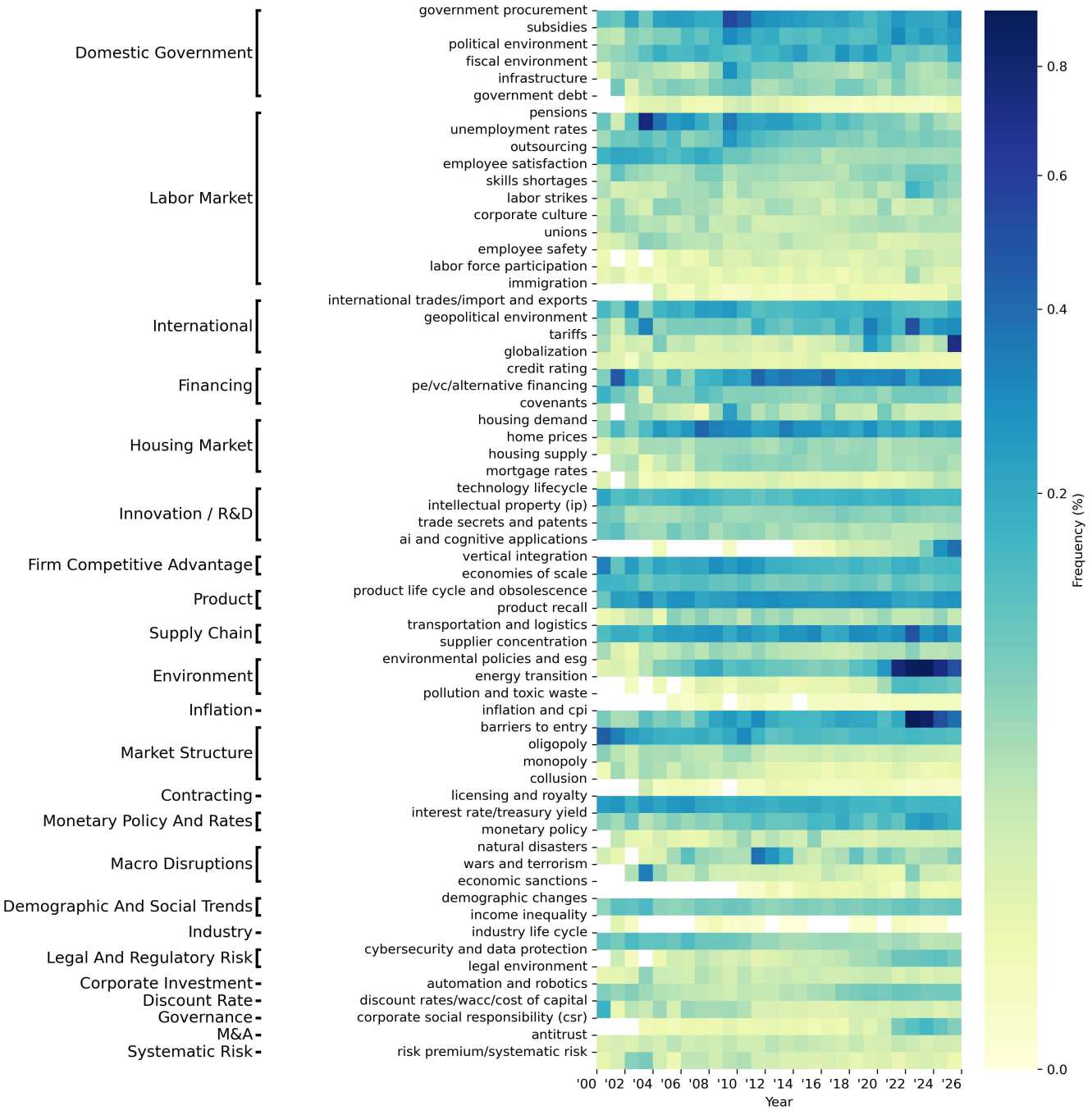


Figure VIII:
Topic Attention Allocation Over Time (Bottom 75 Topics)

This figure plots attention allocation to the 75 least frequently discussed topics across all firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the bottom 75 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic. Note that color intensities are not directly comparable between this figure and Figure VII.

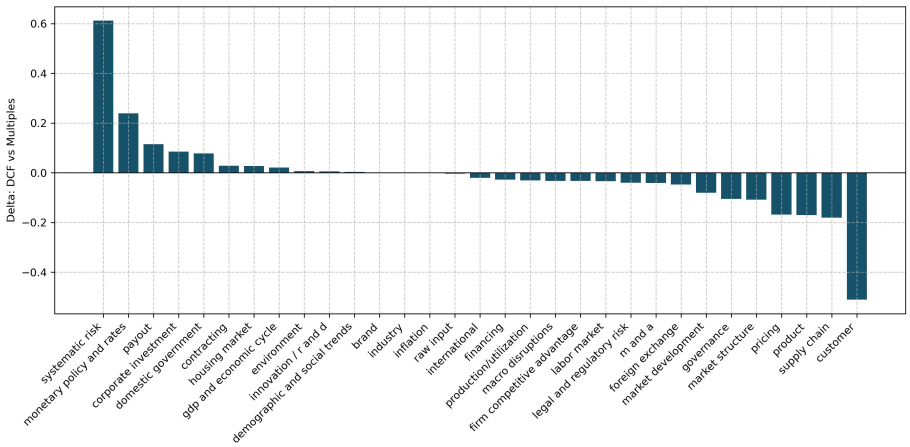
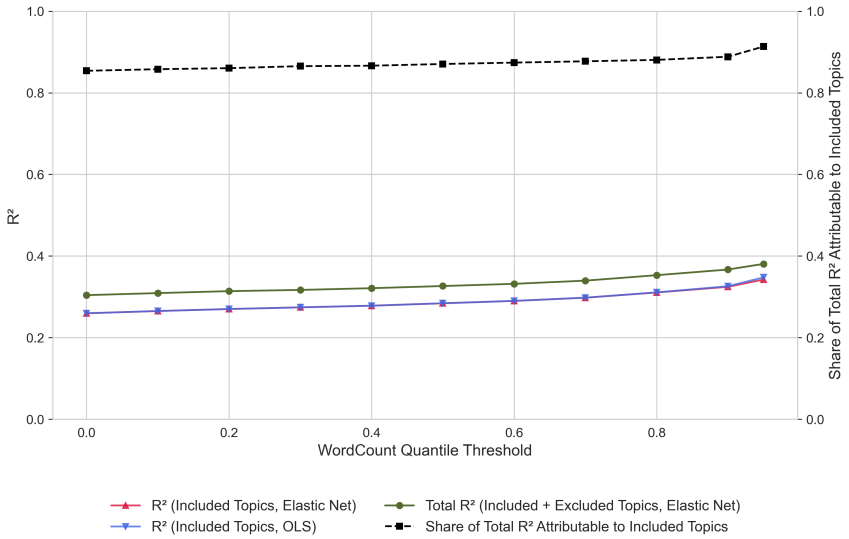


Figure IX:
Attention Allocation Delta: DCF Versus Multiples

This figure plots the average difference in attention weights allocated to topic categories between reports which base the price target on a multiples-based valuation (without any mention of DCF) versus on a DCF-based valuation (without any mention of multiples). Coefficients are estimated from a regression of topic category attention on an indicator for DCF-based valuations, including firm-by-year fixed effects. A positive value indicates that DCF-driven valuations allocate more attention to a given topic category, whereas a negative value indicates less attention than a multiples-driven approach.

(A) Completeness.



(B) Sparsity.

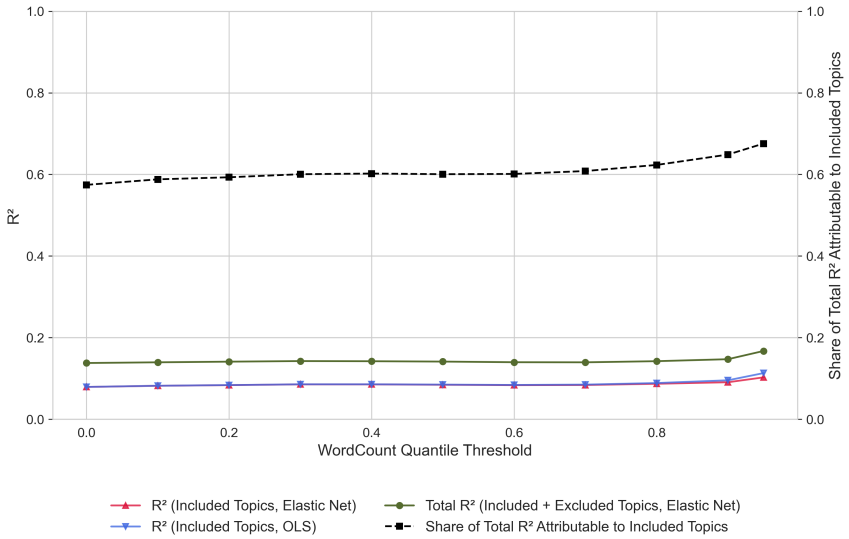
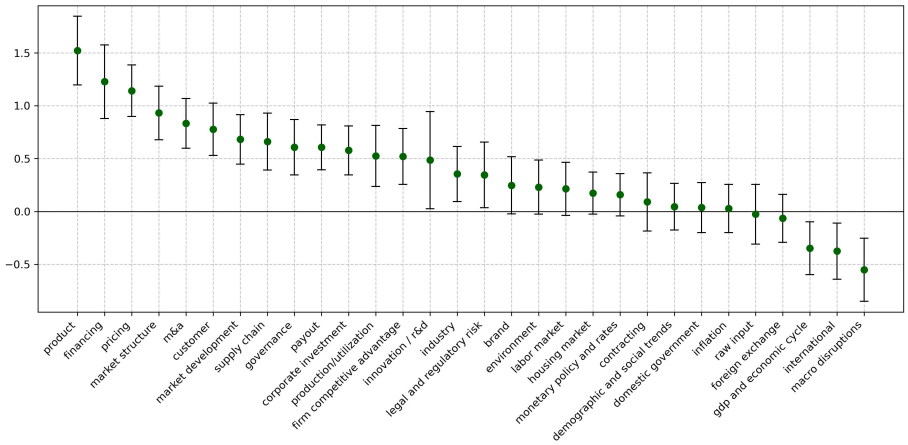


Figure X:
Mental Model Completeness and Sparsity

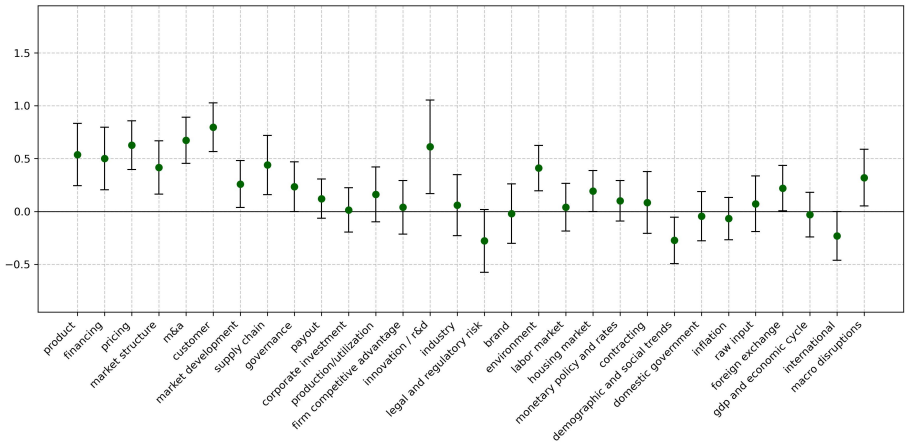
This figure presents results on mental model completeness (Panel A) and sparsity (Panel B). Each panel reports the R^2 from three specifications: (i) in blue, an OLS regression of either the expected capital gain (Panel A) or the magnitude of the analyst's normalized forecast error as a measure of forecast accuracy (Panel B) on indicator variables for whether each topic is discussed in the report, with separate indicators for positive and negative sentiment; (ii) in red, the same regression as in (i), estimated using an elastic net approach that shrinks coefficients (possibly to zero); and (iii) the sum of the R^2 from (i) and the R^2 from a regression of the residuals from (i) on indicators for whether excluded topics are included by other analysts covering the same firm in the same year. We also plot the share of the total R^2 in (iii) that is attributable to (i). Together, the panels show the share of explainable variation in price targets and forecast error magnitudes accounted for by included versus excluded topics, thereby capturing mental model completeness and sparsity, respectively.

(A) Contribution from Changes in Representation.



p-value from joint test of all coefficients equal to zero: < 0.0001

(B) Contribution from Changes in Pricing.



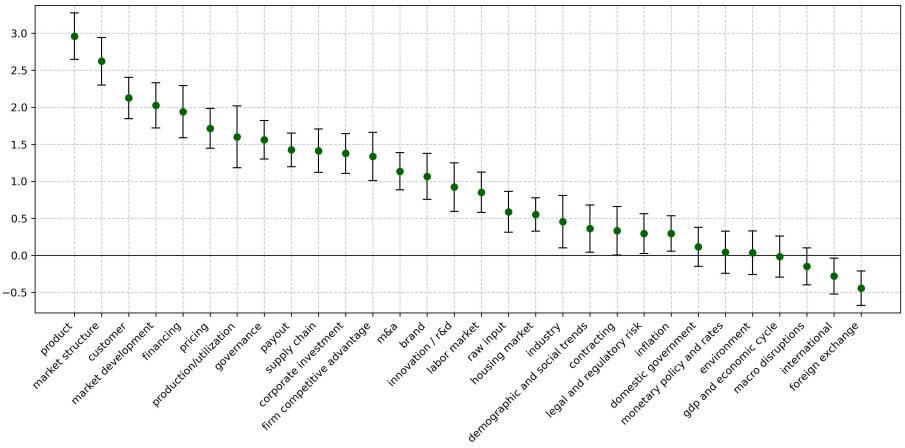
p-value from joint test of all coefficients equal to zero: < 0.0001

Figure XI: Changes in Valuations. This figure shows the decomposition of changes in price targets over time into the components specified in equation (21), reflecting shifts in analyst attention to topic categories (Panel A) or in how they price different attributes (Panel B). Specifically, we estimate:

$$\mathbb{E}_t^i[\hat{p}_{f,t+1}] - \mathbb{E}_{t-1}^i[\hat{p}_{f,t}] = \sum_{k=1}^{K} \left(m_k^+ \Delta\alpha_{kft}^{i,+} + m_k^- \Delta\alpha_{kft}^{i,-} + \Delta m_k^+ \alpha_{kft-1}^{i,+} + \Delta m_k^- \alpha_{kft-1}^{i,-} \right) + \epsilon_{f,t}$$

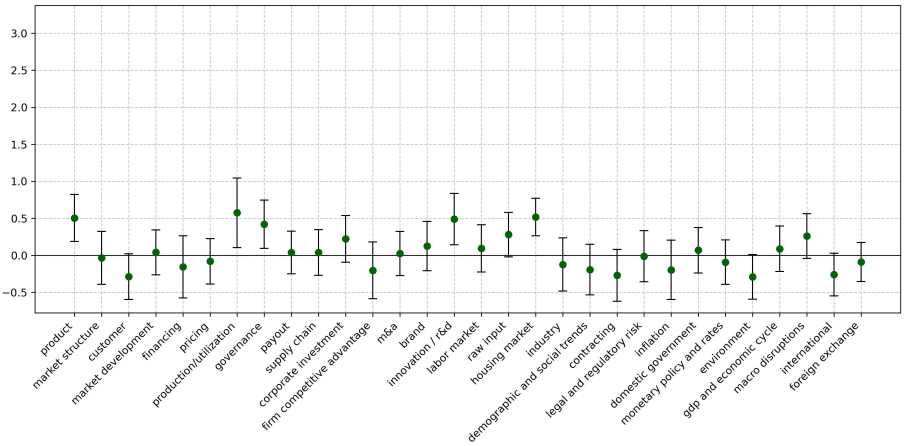
with $\alpha_{kft}^{i,+}$ ($\alpha_{kft}^{i,-}$) denoting the fraction of analyst i's report on firm f in year t allocated to topic k when it is discussed with positive (negative) sentiment, and with $\Delta\alpha_{kft}^{i,+}$ ($\Delta\alpha_{kft}^{i,-}$) capturing the change in the fraction of analyst i's report on firm f allocated to topic k with positive (negative) sentiment between year $t-1$ and year t. $\alpha_{kft}^{i,+}$, $\alpha_{kft}^{i,-}$, $\alpha_{kft-1}^{i,+}$, and $\alpha_{kft-1}^{i,-}$ are scaled by their in-sample standard deviation to enable direct comparisons across topics and sentiment directions. In Panel A, dots represent the average of estimates m_k^+ and m_k^- for each category. In Panel B, dots represent the average of estimates Δm_k^+ and Δm_k^- for each category. The "undetermined" and "valuation" categories are excluded from the exercise.

(A) Contribution from Differences in Representation.



p-value from joint test of all coefficients equal to zero: < 0.0001

(B) Contribution from Differences in Pricing.



p-value from joint test of all coefficients equal to zero: 0.0016

Figure XII: Sources of Disagreement. This figure shows the decomposition of differences in price targets across analysts evaluating the same firm at the same time into the components specified in equation (22), reflecting differences in analyst attention to topic categories (Panel A) or in how they price different attributes (Panel B). Specifically, we estimate:

$$\mathbb{E}_t^i[\hat{p}_{f,t+1}] - \mathbb{E}_t^j[\hat{p}_{f,t+1}] = \sum_{k=1}^{K} \left(m_k^+ \Delta\alpha_{kft}^{i,+} + m_k^- \Delta\alpha_{kft}^{i,-} + \Delta m_k^+ \alpha_{kft}^{j,+} + \Delta m_k^- \alpha_k^{j,-} \right) + \epsilon_{f,t}$$

with $\alpha_{kft}^{i,+}$ ($\alpha_{kft}^{i,-}$) denoting the fraction of analyst i's report on firm f in year t allocated to topic k when it is discussed with positive (negative) sentiment, and with $\Delta a_{kft}^{i,+}$ ($\Delta a_{kft}^{i,-}$) capturing the difference in the fraction of analyst i's and analyst j's report on firm f in year t allocated to topic k with positive (negative) sentiment. We run this regression on analyst pairs where one analyst bases their price target on a multiples-based valuation (without any mention of DCF), and the other on a DCF-based valuation (without any mention of multiples). We systematically sort the pairs as (i =multiples, j =DCF), so that the difference in valuation weights captures the effect of multiples- relative to DCF-based methods. $\alpha_{kft}^{i,+}$, $\alpha_{kft}^{i,-}$, $\alpha_{kft}^{j,+}$, and $\alpha_{kft}^{j,-}$ are scaled by their in-sample standard deviation to enable direct comparisons across topics and sentiment directions. In Panel A, dots represent the average of estimates m_k^+ and m_k^- for each category. In Panel B, dots represent the average of estimates Δm_k^+ and Δm_k^- for each category. The "undetermined" and "valuation" categories are excluded from the exercise.

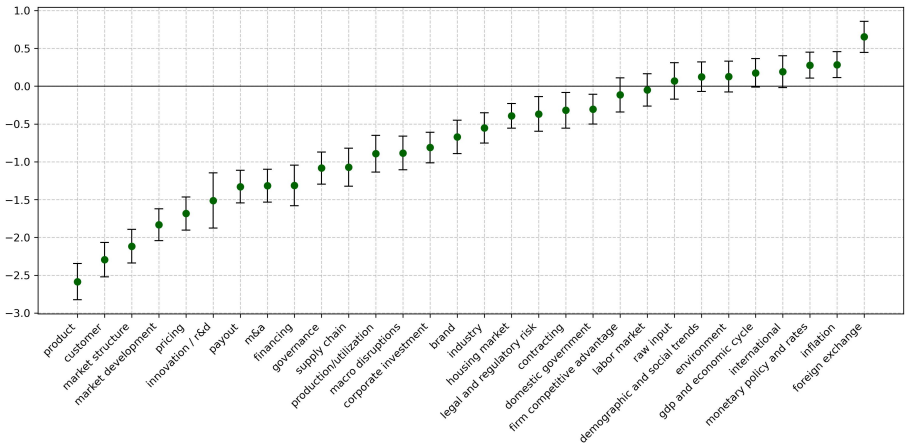
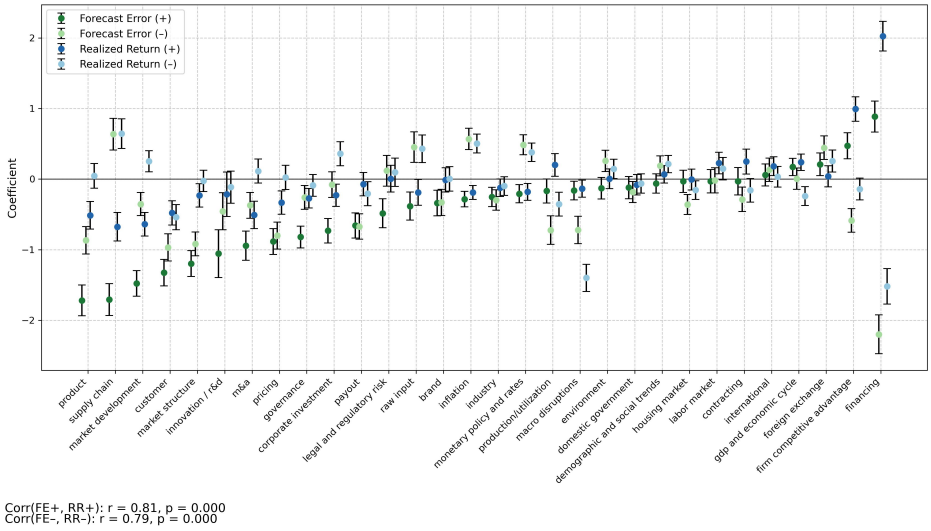


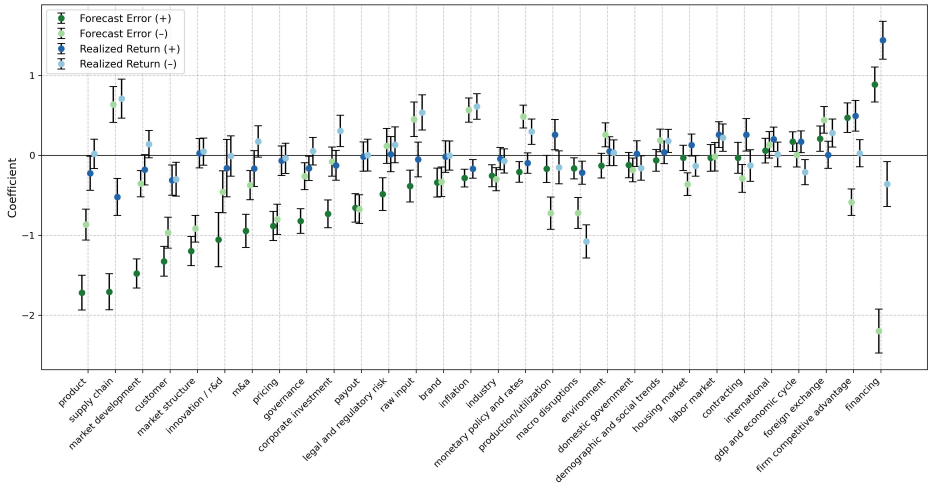
Figure XIII: Sources of Over- and Undervaluation. This figure shows estimates of the degree of over- or underreaction associated with attention to different topic categories, estimated using the Coibion-Gorodnichenko-like regression in equation (23) and including firm fixed effects. Positive coefficients indicate underreaction, whereas negative coefficients indicate overreaction. The "undetermined" and "valuation" categories are excluded from the exercise.

(A) Baseline Estimates.



Corr(FE+, RR+): r = 0.81, p = 0.000
Corr(FE–, RR–): r = 0.79, p = 0.000

(B) Controlling for Known Return Predictors.



Corr(FE+, RR+): r = 0.74, p = 0.000
Corr(FE–, RR–): r = 0.65, p = 0.000

Figure XIV:
Over- and Underreaction and Realized Returns

This figure shows the predictive power of topic categories associated with over- and underreaction for subsequent realized returns, estimating equation (24). Panel A includes topic variables and firm fixed effects. Panel B includes a series of known predictors of returns, namely book-to-market ratio, earnings-to-price ratio, prior 12-month return (excluding the most recent month), asset growth, and operating profitability, as additional control variables.

Table I:
Descriptive Statistics on Equity Reports and Analysts' Mental Models

This table presents summary statistics based on our main sample of equity reports spanning 2000–2025. Panel A reports basic report characteristics. Panel B summarizes various features of extracted information, including topics, valuation channels, entities, time outlook, and sentiment. Panel C presents statistics on quantitative forecasts, forecast errors, and forecast disagreement. Panel D provides statistics on topic and sentiment similarity measures, while Panel E focuses on semantic similarity. *Long Reports* in Panels D and E refer to those in the top decile of word count.

Panel A: Dataset Structure	Avg. Obs. per Group-Year	Avg. Obs. per Group	Nb. Obs.
Group:			
Equity Report			301,364
Year		11,591	26
Brokerage House	270	7,008	43
Firm	1	16	18,284
HQ Country	116	3,014	100
Industry	580	15,068	20

Panel B: Mental Model Features	Mean	25^{th} pct.	Median	75^{th} pct.	S. D.	Nb. Obs.
Equity Report Text:						
Nb. Topics per Report	16.32	12.00	16.00	20.00	5.63	301,364
Nb. Arguments per Report	39.13	23.00	33.00	49.00	22.56	301,364
Nb. Arguments per Topic	2.32	1.73	2.15	2.74	0.82	301,364
Nb. Words per Report	1,198.02	573.00	926.00	1,533.00	878.06	301,364
Valuation Channels:						
Sales (%)	44.05	30.00	44.44	58.00	20.16	301,364
Costs (%)	12.76	4.55	10.71	18.75	11.06	301,364
Margins (%)	10.60	0.00	8.89	16.67	9.92	301,364
Earnings (%)	28.05	15.38	25.00	37.50	17.50	301,364
Discount Rate (%)	4.47	0.00	0.00	6.25	7.72	301,364
Entities:						
Firm (%)	82.19	75.00	84.62	92.31	13.35	301,364
Industry (%)	12.92	4.44	10.34	18.75	11.38	301,364
Macroeconomy (%)	4.90	0.00	3.17	7.50	6.08	301,364
Time Outlook:						
Time Outlook	0.45	0.27	0.48	0.67	0.29	301,364
Past (%)	14.74	4.00	11.11	22.22	13.74	301,364
Present (%)	28.05	19.05	26.98	35.90	12.75	301,364
Near-future (%)	52.39	41.53	52.63	63.49	15.97	301,364
Distant-future (%)	3.49	0.00	1.56	5.26	5.08	301,364
Sentiment:						
Sentiment	0.24	-0.03	0.28	0.54	0.40	301,364
Sentiment^{Past}	0.29	-0.14	0.42	1.00	0.66	253,845
$\text{Sentiment}^{Present}$	0.19	-0.18	0.25	0.60	0.52	298,766
$\text{Sentiment}^{Near-future}$	0.25	-0.06	0.30	0.60	0.46	300,992
$\text{Sentiment}^{Distant-future}$	0.46	0.00	0.67	1.00	0.64	157,196

Panel C: Forecast Variables	Mean	25^{th} pct.	Median	75^{th} pct.	S. D.	Nb. Obs.
Expected Capital Gain ($\ln(\frac{E_{i,j,t}[P_{j,t+1}]}{P_{j,t}})$)	0.13	0.03	0.12	0.22	0.19	301,364
Forecast Error	-0.13	-0.33	-0.08	0.14	0.46	257,376
Forecast Disagreement	0.14	0.05	0.10	0.19	0.13	514,784

Panel D: Similarity Measures	Mean	25^{th} pct.	Median	75^{th} pct.	S. D.	Nb. Obs.
Topic Similarity (Analyst-Pairs):						
Topic Overlap	0.33	0.26	0.33	0.40	0.10	514,784
$\text{Topic Overlap}^{LongReports}$	0.40	0.34	0.40	0.46	0.09	6,623
Argument Similarity (Analyst-Pairs):						
Same Sentiment	0.57	0.00	1.00	1.00	0.49	4,399,837
Same Sentiment | Same Val. Channel and Outlook	0.61	0.00	1.00	1.00	0.49	1,630,989
Same Sentiment | Different Val. Channel and Outlook	0.52	0.00	1.00	1.00	0.50	646,889

Panel E: Semantic Similarity Measures	Mean	25^{th} pct.	Median	75^{th} pct.	S. D.	Nb. Obs.
Topic Similarity (Analyst-Pairs):						
Topic Overlap (85% Similarity Cutoff)	0.44	0.35	0.44	0.53	0.13	514,784
$\text{Topic Overlap}^{LongReports}$ (85% Similarity Cutoff)	0.55	0.48	0.55	0.62	0.11	6,623
Topic Overlap (80% Similarity Cutoff)	0.85	0.80	0.87	0.92	0.09	514,784
$\text{Topic Overlap}^{LongReports}$ (80% Similarity Cutoff)	0.91	0.88	0.92	0.95	0.06	6,623

Table II:
Topic Attention to Sales and Costs and 12-Month Price Target Forecasts

This table presents results examining how analysts' attention allocation to sales and costs, interacted with the sentiment of related arguments, affect their 12-month price target forecasts. The dependent variable is the natural logarithm of analyst i's price target forecast for firm j in year t normalized by the firm's current price as of the time of the equity report. The analysis uses our main sample spanning 2000–2025. Sales and cost attention are measured as the share of an analyst's report devoted to sales- and cost-related arguments, respectively. These shares are interacted with indicators for whether the corresponding arguments express, on average, positive or negative sentiment. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Expected Capital Gain ($\ln(\frac{E_{i,j,t}[P_{j,t+1}]}{P_{j,t}})$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Sales Attention$_{i,j,t}$*1(Sales Sentiment$^{+}_{i,j,t}$)	0.08***	0.09***	0.05***			
	(0.00)	(0.00)	(0.00)			
Sales Attention$_{i,j,t}$*1(Sales Sentiment$^{-}_{i,j,t}$)	-0.18***	-0.14***	-0.10***			
	(0.00)	(0.00)	(0.00)			
Cost Attention$_{i,j,t}$*1(Cost Sentiment$^{+}_{i,j,t}$)				0.02**	0.09***	0.07***
				(0.01)	(0.01)	(0.01)
Cost Attention$_{i,j,t}$*1(Cost Sentiment$^{-}_{i,j,t}$)				-0.24***	-0.22***	-0.12***
				(0.01)	(0.01)	(0.01)
Firm*Year FE	No	Yes	Yes	No	Yes	Yes
Firm*Analyst FE	No	No	Yes	No	No	Yes
Observations	293,649	244,272	195,101	247,108	199,252	151,236
F Statistics	2712.02	2241.82	861.88	764.08	925.32	324.68
R^2	0.04	0.56	0.78	0.01	0.54	0.78
Within R^2	0.04	0.05	0.03	0.01	0.02	0.01

Table III:
Variation in Arguments Made and Firm HQ-Analyst Distance

This table presents results examining the drivers of attention allocation to firm- and macro-related arguments as a function of the distance between the firm's headquarters city and the analyst's location. Firm-related and macro-related arguments refer to the share of arguments in an analyst's report that pertain to firm-specific and macroeconomic discussions, respectively. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Firm-Related Arguments (%)				Macro-Related Arguments (%)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance $\leq$ 25km	0.88***							
	(0.15)							
Distance $\leq$ 100km		0.58***						
		(0.13)						
Distance $>$ 75^{th} Pct.					0.18***			
					(0.06)			
Distance $>$ 85^{th} Pct.						0.21***		
						(0.07)		
Log Distance (km)			-0.11***	-0.06**			0.04***	0.04**
			(0.02)	(0.03)			(0.01)	(0.02)
Firm*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	Full	Full	HQ = Analyst Country	Full	Full	Full	HQ = Analyst Country
Observations	239,895	239,895	239,895	170,040	239,895	239,895	239,895	170,040
F Statistics	35.809	19.759	27.662	3.989	8.488	9.059	12.579	6.308
R^2	0.519	0.519	0.519	0.529	0.461	0.461	0.461	0.473
Within R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table IV:
Drivers of Topic Attention – The Case of Inflation Narratives

This table presents results examining the drivers of attention allocation to inflation. The analysis uses our main sample spanning 2000–2025, restricting to analysts located in different countries than the firm's headquarters. *Attention to Inflation* is defined as the share of analysts' reports allocated to discussing the topic "Inflation and CPI," at the analyst *i*, firm *j*, and year *t* level. $\text{Inflation}^{FirmHQCountry}$ and $\text{Inflation}^{AnalystCountry}$ refer to the realized inflation in the firm's HQ country and analyst country, respectively, in year *t*. The remaining variables are indicators denoting whether realized inflation in the firm's HQ and analyst countries falls within specific ranges, respectively. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Attention to $\text{Inflation}_{i,j,t}$				
				Analyst–HQ Distance	
				≤ Median	> Median
	(1)	(2)	(3)	(4)	(5)
$\text{Inflation}_{p,t}^{FirmHQCountry}$	0.04***	0.02***			
	(0.00)	(0.00)			
$\text{Inflation}_{c,t}^{AnalystCountry}$	0.01***	0.01**			
	(0.00)	(0.00)			
$\text{Inflation}_{p,t}^{FirmHQCountry}$ in $[0\%, 2\%]$			-0.02	-0.01	-0.03
			(0.02)	(0.02)	(0.03)
$\text{Inflation}_{p,t}^{FirmHQCountry}$ in $(2\%, 5\%]$			0.02	0.02	0.02
			(0.02)	(0.03)	(0.03)
$\text{Inflation}_{p,t}^{FirmHQCountry} > 5\%$			0.09***	0.07	0.10***
			(0.03)	(0.05)	(0.04)
$\text{Inflation}_{c,t}^{AnalystCountry}$ in $[0\%, 2\%]$			0.02	0.03	0.01
			(0.02)	(0.03)	(0.03)
$\text{Inflation}_{c,t}^{AnalystCountry}$ in $(2\%, 5\%]$			0.04**	0.09**	0.02
			(0.02)	(0.04)	(0.03)
$\text{Inflation}_{c,t}^{AnalystCountry} > 5\%$			0.11***	0.15***	0.12***
			(0.03)	(0.05)	(0.04)
Firm FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Observations	73,077	71,143	71,143	35,874	33,141
F Statistics	111.627	27.527	14.444	9.175	7.955
R^2	0.013	0.232	0.232	0.192	0.305
Within R^2	0.013	0.001	0.002	0.002	0.002

Table V:
Valuation Methods, Time Outlook, and 12-Month Price Target Forecasts

This table presents results examining the role of valuation methods in explaining time outlook and quantitative forecasts. The analysis uses our main sample spanning 2000–2025, focusing on reports that use a DCF-based or multiples-based price-target valuation method. Panel A includes firm-by-year fixed effects. Panel B addes analyst-by-firm fixed effects. Panel C is estimated on the subsample of analysts who have consistently used the valuation method in the current report across all of their previous reports for the same firm. In Columns 1, 2, and 3, the dependent variables capture the extent to which analyst i's equity report for for firm j in year t relies on backward-looking, near-future, and distant-future arguments, respectively. In Column 4, the dependent variable is the natural logarithm of analyst's i's price target forecast for firm j in year t normalized by the firm's current price as of the time of the equity report. The key independent variable, *DCF Usage Only*$_{i,j,t}$, is an indicator variable that equals one if the analyst uses a DCF-based valuation method without any mention of a multiples-based approach, and zero otherwise. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedasticity-consistent estimators clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Within-Firm-Year Variation

Dependent variable:	Past$_{i,j,t}$ ×100	Near-future$_{i,j,t}$ ×100	Distant-future$_{i,j,t}$ ×100	Expected Capital Gain$_{i,j,t}$ ×100
	(1)	(2)	(3)	(4)
DCF Usage Only$_{i,j,t}$	-0.35***	-2.56***	1.78***	0.45***
	(0.10)	(0.11)	(0.05)	(0.12)
Firm*Year FE	Yes	Yes	Yes	
Observations	221,958	221,958	221,958	221,958
F Statistics	12.30	531.57	1530.34	13.90
R^2	0.36	0.38	0.38	0.54

Panel B: Within-Analyst-Firm Variation

Dependent variable:	Past$_{i,j,t}$ ×100	Near-future$_{i,j,t}$ ×100	Distant-future$_{i,j,t}$ ×100	Expected Capital Gain$_{i,j,t}$ ×100
	(1)	(2)	(3)	(4)
DCF Usage Only$_{i,j,t}$	0.05	-1.77***	1.22***	0.69***
	(0.12)	(0.15)	(0.06)	(0.18)
Firm*Year FE	Yes	Yes	Yes	Yes
Analyst*Firm FE	Yes	Yes	Yes	Yes
Observations	209,704	209,704	209,704	209,704
F Statistics	0.21	138.34	397.84	13.85
R^2	0.62	0.61	0.58	0.68

Panel C: Consistent Valuation Method Usage

Dependent variable:	Past$_{i,j,t}$ ×100	Near-future$_{i,j,t}$ ×100	Distant-future$_{i,j,t}$ ×100	Expected Capital Gain$_{i,j,t}$ ×100
	(1)	(2)	(3)	(4)
DCF Usage Only$_{i,j,t}$	0.12	-2.41***	1.53***	1.17***
	(0.16)	(0.19)	(0.07)	(0.21)
Firm*Year FE	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes
Observations	144,902	144,902	144,902	144,902
F Statistics	0.60	164.78	474.49	29.83
R^2	0.54	0.54	0.52	0.64

Table VI:
Mental Model Rigidity in the Time Series

This table presents results examining the persistence of analysts' mental models over time. The analysis uses our main sample spanning 2000–2025, pairing each observation (analyst i forecasting firm j in year t) with all forecasts for the same firm in year $t+1$. The evaluations of firm j in $t+1$ by other analysts ($i' \neq i$) serve as a benchmark for assessing analyst-specific persistence. In Panel A, *Same-Analyst Pair* is an indicator for whether a given pair involves the same analyst ($i = i'$) and is zero otherwise ($i \neq i'$). *Same Valuation Method* is an indicator variable equal to one if both analysts in the pair align in their valuation method, i.e., align whether they use a DCF-based price target and whether they use a multiples-based one. In Panel B, *Large Forecast Error* is an indicator variable for whether the magnitude of the analyst's forecast error (realized price minus price target, normalized by the firm's price as of the time of the equity report) is in the largest decile. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Time-Series Mental Model Persistence

Dependent variable:	Same Valuation Method		Topic Overlap		Topic Overlap Same-Analyst Pairs Only
	(1)	(2)	(3)	(4)	(5)
Same-Analyst Pair	0.33***	0.33***	0.08***	0.08***	
	(0.00)	(0.00)	(0.00)	(0.00)	
Same Valuation Method					0.01***
					(0.00)
LHS Mean | Same-Analyst Pair	0.81	0.82	0.40	0.40	0.41
Analyst*Firm*Year FE	No	Yes	No	Yes	No
Analyst*Firm FE	No	No	No	No	Yes
Firm*Year FE	No	No	No	No	Yes
Observations	739,748	721,491	739,748	721,491	69,198
F Statistics	6828.79	13016.08	23900.83	35385.18	62.85
R^2	0.07	0.42	0.09	0.47	0.69
Within R^2	0.07	0.10	0.09	0.13	0.00

Panel B: Time-Series Persistence and Forecast Errors

Dependent variable:	Same Valuation Method		Topic Overlap	
	(1)	(2)	(3)	(4)
Large Forecast Error	-0.01**	0.00	-0.01***	-0.01*
	(0.01)	(0.01)	(0.00)	(0.00)
LHS Mean	0.79	0.80	0.40	0.40
Analyst*Firm FE	Yes	Yes	Yes	Yes
Firm*Year FE	No	Yes	No	Yes
Observations	90,118	63,320	113,229	81,773
F Statistics	3.96	0.05	36.99	2.84
R^2	0.45	0.66	0.48	0.67
Within R^2	0.00	0.00	0.00	0.00

Table VII:
Mental Model Rigidity in the Cross Section

This table presents results examining the persistence of analysts' mental models in the cross-section. The analysis uses our main sample spanning 2000–2025, pairing each observation (analyst i forecasting firm j in year t) with all other forecasts made by the same analyst for different firms in the same year ($j' \neq j$), as well as with forecasts by other analysts for those same firms. The evaluations of firms j' by other analysts ($i' \neq i$) serve as a benchmark for assessing within-analyst persistence. *Same-Analyst Pair* is an indicator for whether a given pair involves the same analyst ($i = i'$) and is zero otherwise ($i \neq i'$). Note that for a given analyst i, the pairing of firms j and j' contributes a different part of the constructed sample than the pairing of j' and j, as this ordering affects the set of comparison analysts (i.e., those forecasting j' versus those forecasting j). *Same Valuation Method* is an indicator equal to one if both forecasts in the pair use the same valuation method—that is, if they both either include or omit a DCF-based and a multiples-based price target. *Same Valuation Method* is an indicator variable equal to one if both forecasts in the pair use the same valuation method—specifically, if they both employ or both omit a DCF-based and a multiples-based price target. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and double-clustered at the level of both firms in the pair. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Same Valuation Method		Topic Overlap		Topic Overlap Same-Analyst Pairs Only
	(1)	(2)	(3)	(4)	(5)
Same-Analyst Pair	0.23***	0.24***	0.03***	0.04***	
	(0.00)	(0.00)	(0.00)	(0.00)	
Same Valuation Method					0.03***
					(0.00)
LHS Mean | Same-Analyst Pair	0.71	0.72	0.30	0.30	0.30
Analyst*Firm-Pair FE	No	Yes	No	Yes	No
Analyst*Firm$_A$ FE	No	No	No	No	Yes
Analyst*Firm$_B$ FE	No	No	No	No	Yes
Observations	12,836,984	12,807,012	12,836,984	12,807,012	1,483,108
F Statistics	2726.39	5057.61	5079.04	13324.16	3163.31
R^2	0.02	0.29	0.01	0.28	0.40
Within R^2	0.02	0.03	0.01	0.02	0.01

Table VIII:
Valuation Methods, Topic Attention, and Forecast Disagreement

This table presents results examining the role of valuation methods and topic attention in explaining forecast (dis-)agreement. The analysis uses our main sample spanning 2000–2025, creating pairs of analysts evaluating the same firm in the same year. The dependent variable, *Forecast Disagreement*$_{A,B,j,t}$, is defined as the absolute difference between forecaster A and B's 12-month price targets normalized by the firm's current price as of the time of the equity reports, multiplied by 100. *Same Valuation Method*$_{A,B,j,t}$, is an indicator variable equal to one if both analysts in the pair align in their valuation method, i.e., align whether they use a DCF-based price target and whether they use a multiples-based one. *Topic Overlap*$_{A,B,j,t}$ captures the number of overlapping topics between analyst A and B's reports divided by the total number of distinct topics mentioned in either report, multiplied by 100. *Argument Overlap*$_{A,B,j,t}$ is the share of overlapping arguments, i.e., topics for which both analysts assign the same time outlook, sentiment, and valuation channel, multiplied by 100. When present, (z) indicates that independent variables are normalized to have mean zero and standard deviation one to allow for direct comparability. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedasticity-consistent estimators clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Forecast Disagreement$_{A,B,j,t}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Same Valuation Method$_{A,B,j,t}$	-0.44***	-0.14				
	(0.05)	(0.09)				
Topic Overlap$_{A,B,j,t}$			-0.06***	-0.04***		
			(0.00)	(0.00)		
Argument Overlap$_{A,B,j,t}$			-0.05***	-0.03***		
			(0.00)	(0.00)		
Same Valuation Method (z)$_{A,B,j,t}$					-0.15***	-0.04
					(0.03)	(0.04)
Topic Overlap (z)$_{A,B,j,t}$					-0.63***	-0.40***
					(0.02)	(0.03)
Argument Overlap (z)$_{A,B,j,t}$					-0.59***	-0.38***
					(0.02)	(0.03)
Firm*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst-A*Firm FE	No	Yes	No	Yes	No	Yes
Analyst-B*Firm FE	No	Yes	No	Yes	No	Yes
Observations	492,345	311,604	492,028	311,341	492,028	311,341
F Statistics	65.65	2.68	665.40	187.07	450.54	124.96
R^2	0.33	0.67	0.34	0.67	0.34	0.67

Appendix

Mental Models and Financial Forecasts

Francesca Bastianello, Paul H. Décaire, Marius Guenzel

Appendix A Proofs and Derivations

A.1 Derivation of Objective Function for Endogenous Valuation Method

The agent solves the following optimization problem:

$$\min_{m,\tau_{s1},\tau_{s2}} \quad \mathbb{E}\left[\left(p_t^{true} - \mathbb{E}_t[p_t^m]\right)^2\right] \qquad s.t. \qquad c_1\tau_{1s} + c_2\tau_{2s} = C \tag{25}$$

We can re-write the objective function as:

$$\mathbb{E}\Big[\left(p_t^{true} - \mathbb{E}[p_t^m]\right)^2\Big] = \mathbb{V}\Big[p^{true} - \mathbb{E}_t[p_t^m]\Big] + \mathbb{E}\Big[p_t^{true} - \mathbb{E}_t[p_t^m]\Big]^2 \tag{26}$$

$$= \mathbb{V}\left[\sum_{j=1,2}\Big[(v_j - m_j)\mathbb{E}_t[x_{jt}] + v_j\left(x_{jt} - \mathbb{E}_t[x_{jt}]\right)\Big]\right] + \mathbb{E}\Big[p_t^{true} - \mathbb{E}[p_t^m]\Big]^2 \tag{27}$$

$$= \sum_{j=1,2}(v_j - m_j)^2\mathbb{V}\left[\frac{\tau_{js}}{\tau_{js}+\tau_{0j}}\left(x_{jt} + u_{jt}\right)\right] + \sum_{j=1,2} v_j^2\left(\frac{1}{\tau_{sj}+\tau_{0j}}\right) + \mathbb{E}\Big[p_t^{true} - \mathbb{E}_t[p_t^m]\Big]^2 \tag{28}$$

$$= \sum_{j=1,2}(v_j - m_j)^2\left(\frac{\tau_{js}}{\tau_{js}+\tau_{0j}}\right)^2\left(\frac{1}{\tau_{0j}} + \frac{1}{\tau_{js}}\right) + \sum_{j=1,2} v_j^2\left(\frac{1}{\tau_{sj}+\tau_{0j}}\right) + \mathbb{E}\Big[p_t^{true} - \mathbb{E}_t[p_t^m]\Big]^2 \tag{29}$$

$$= \sum_{j=1,2}\frac{v_j^2}{\tau_{0j}+\tau_{js}} + \sum_{j=1,2}(v_j - m_j)^2\left(\frac{\tau_{js}}{\tau_{0j}(\tau_{0j}+\tau_{sj})}\right) + \left(\sum_{j=1,2}(v_j - m_j)\mu_j\right)^2 \tag{30}$$

$$= \sum_{j=1,2}\frac{v_j^2 - (v_j - m_j)^2}{\tau_{0j}+\tau_{js}} + \sum_{j=1,2}(v_j - m_j)^2\left(\frac{1}{\tau_{0j}}\right) + \left(\sum_{j=1,2}(v_j - m_j)\mu_j\right)^2 \tag{31}$$

where we used the fact that $\mathbb{V}[x_{jt} - \mathbb{E}_t[x_{jt}]] = (\tau_{js} + \tau_{0j})^{-1}$.[56] The final expression gives us the objective function in (13). Notice that when agents use the correct model, the second and third term cancel out and we are back to the objective function in (7).

[56] To derive this explicitly, we see that:

$$\mathbb{V}\left[x_{jt} - \mathbb{E}_t[x_{jt}]\right] = \mathbb{V}\left[x_{jt} - \frac{\tau_{sj}}{\tau_{sj}+\tau_{0j}}s_{jt} - \frac{\tau_{0j}}{\tau_{sj}+\tau_{0j}}\mu_j\right] = \mathbb{V}\left[x_{jt} - \frac{\tau_{sj}}{\tau_{sj}+\tau_{0j}}x_{jt} - \frac{\tau_{sj}}{\tau_{sj}+\tau_{0j}}u_{jt}\right] \tag{32}$$

$$= \mathbb{V}\left[\frac{\tau_{0j}}{\tau_{sj}+\tau_{0j}}x_{jt} - \frac{\tau_{sj}}{\tau_{sj}+\tau_{0j}}u_{jt}\right] = \left(\frac{\tau_{0j}}{\tau_{sj}+\tau_{0j}}\right)^2\frac{1}{\tau_{0j}} + \left(\frac{\tau_{sj}}{\tau_{sj}+\tau_{0j}}\right)^2\frac{1}{\tau_{sj}} = \frac{1}{\tau_{sj}+\tau_{0j}} \tag{33}$$

A.2 Proof of Proposition 1

We can either inspect the first order conditions in the text, or the optimal allocation of attention to variable x_k, which is given by the expression in (15):

$$\tau^*_{sk} = \frac{C + c_j\tau_{0j} - \sqrt{\frac{v_j^2}{v_k^2}}\sqrt{c_k c_j}\tau_{0k}}{c_k + \sqrt{\frac{v_j^2}{v_k^2}}\sqrt{c_k c_j}} \tag{34}$$

Ceteris paribus, this expression is clearly increasing in v_k (as the numerator is increasing in v_k and the denominator is decreasing in v_k), decreasing in c_k (as the numerator is decreasing in c_k and the denominator is increasing in c_k), and decreasing in τ_{0k} (as the numerator is decreasing in τ_{0k}). □

A.3 Proof of Proposition 2

The marginal benefit evaluated at $\tau_{sk} = 0$ is given by: $v_k^2(\tau_{0k})^{-2}$. The marginal cost at $\tau_{sk} = 0$ λc_k. Since $\tau_{sk} \geq 0$, it follows that if the marginal cost is greater than the marginal benefit, then agents will optimally set $\tau_{sk} = 0$. Setting $\tau_{sk} = 0$ in the budget constraint then delivers $\tau_{sj} = \frac{C}{c_j}$.

Therefore for the optimal $\tau_{sk} = 0$, it must be that: $v_k^2(\tau_{0k})^{-2} < \frac{c_k v_j^2}{c_j\left(\tau_{02} + \frac{C}{c_j}\right)}$. Ceteris paribus, this condition is more likely satisfied when v_k is low, τ_{0k} is high, c_k is higher. This is also more likely satisfied when v_j is high, τ_{0j} is low,c_j is low, and when C is high. □

A.4 Proof of Proposition 3

We are interested in proving that $\tau^m_{sk} > \tau^*_{sk} \iff \left(\frac{v_k - m_k}{v_k}\right)^2 < \left(\frac{v_j - m_j}{v_j}\right)^2$. Using the expressions in (11) and (15), we have:

$$\tau^m_{sk} > \tau^*_{sk} \iff \frac{C + c_j\tau_{0j} - \sqrt{\frac{v_j^2 - (v_j - m_j)^2}{v_k^2 - (v_k - m_k)^2}}\sqrt{c_k c_j}\tau_{0k}}{c_k + \sqrt{\frac{v_j^2 - (v_j - m_j)^2}{v_k^2 - (v_k - m_k)^2}}\sqrt{c_k c_j}} > \frac{C + c_j\tau_{0j} - \sqrt{\frac{v_j^2}{v_k^2}}\sqrt{c_k c_j}\tau_{0k}}{c_k + \sqrt{\frac{v_j^2}{v_k^2}}\sqrt{c_k c_j}} \tag{35}$$

This holds true if and only if:

$$\frac{v_j^2 - (v_j - m_j)^2}{v_k^2 - (v_k - m_k)^2} < \frac{v_j^2}{v_k^2} \iff \left(\frac{v_k - m_k}{v_k}\right)^2 < \left(\frac{v_j - m_j}{v_j}\right)^2 \tag{36}$$

□

A.5 Proof of Proposition 4

Starting from a model with no wedges, we are interested in understanding when it might be optimal to introduce a wedge along a given dimension. Formally, we are interested in understanding conditions for which the following relationship holds:

$$MSE(w_1^2, 0) < MSE(0, 0) \tag{37}$$

where $w_1^2 = (v_1 - m_1)^2$. Using the definition of MSE in (16), the above condition holds if:

$$\frac{\left(\sqrt{c_1(v_1^2 - w_1^2)} + \sqrt{c_2 v_2^2}\right)^2}{C + c_1\tau_{01} + c_2\tau_{02}} + w_1^2\left(\frac{1}{\tau_{01}} + \mu_1^2\right) < \frac{\left(\sqrt{c_1 v_1^2} + \sqrt{c_2 v_2^2}\right)^2}{C + c_1\tau_{01} + c_2\tau_{02}} \tag{38}$$

$$w_1^2\left(\frac{1}{\tau_{01}} + \mu_1^2\right) < \frac{c_1 w_1^2 + 2v_2\sqrt{c_1 c_2}\left(\sqrt{v_1^2} - \sqrt{v_1^2 - w_1^2}\right)}{C + c_1\tau_{01} + c_2\tau_{02}} \tag{39}$$

For a given wedge w_1^2 and $v_1, v_2 \geq 0$, $v_1^2 \geq w_1^2$, this expression is easier to satisfy when the variable of interest is less valuable (lower v_1, μ_1), when the value of learning about the other variables is higher (lower τ_{02}, c_2, or higher v_2), or when the total budget on attention is lower (lower C). □

Appendix B LLM Prompt to Extract Reasoning

Step 1: Identify key drivers of valuation exercise

Prompt input: Equity report.

Prompt output: Main mental model.

1. Main Task Analyze the provided equity report excerpt to identify all topics that clearly and directly explain the analyst's price target. For each topic:

- Categorize it into one of the following "entity": "firm," "industry," or "macroeconomy."
- Categorize it into one of the following "channel": "earnings/cash flow: sales", "earnings/cash flow: costs", "earnings/cash flow: margins", "earnings/cash flow: earnings", "discount rate", "valuation", or "other".
- Identify its associated sentiment: "positive", "negative", "neutral", or "unclear."
- Identify the best keyword that captures the sentiment associated with the topic.
- Determine the time outlook: "past", "present", "near-future", "distant-future", or "unclear."
- Identify a 5 word snippet from the text to capture where the topic is first discussed.

2. Equity Report

```
<equityreport>
{EquityReport}
</equityreport>
```

3. Topic Identification

- Focus exclusively on topics directly tied to the justification of the price target.
- Each identified topic should be specific and clearly related to a single factor influencing the price target.
- Disregard general statements or any information not directly relevant to the price target justification.

4. Entity Identification

- Assign each topic to the most appropriate entity: "firm" (company-specific factors), "industry" (sector or industry trends), or "macroeconomy" (broader economic conditions).
- If a topic could reasonably belong to more than one entity, select the entity that best fits the context provided in the report.

5. Channel Identification Assign each topic to one of the following channels based on its context:

- "earnings/cash flow: sales": If related to sales or revenues.
- "earnings/cash flow: costs": If related to costs or operating expenses.
- "earnings/cash flow: margins": If related to profitability margins.
- "earnings/cash flow: earnings": If related to earnings, EBITDA, EBIT, or cash flows.
- "discount rate": If related to terms like "discount rate", "weighted average cost of capital/wacc", "equity beta", "idiosyncratic risk", "credit rating", "systematic risk", "equity risk premium (erp)", "market risk premium (mrp)", "interest rate", "risk-free rate (rf)", or "treasury yield".
- "valuation": If related to valuation terms such as "undervalued," "overvalued," or "fair-valued."
- "other": If none of the above apply.

6. Sentiment Identification Determine the sentiment associated with each topic, selecting one of the following:

- "positive": Indicates optimism or favorable conditions such as "beat expectations", "surpass expectations", "confidence", "bullish", "optimistic", ...
- "negative": Indicates pessimism or unfavorable conditions such as "uncertainty", "challenges", "caution", "concerns", "bearish", "underperform expectations", ...
- "neutral": Indicates neutrality or conditions in line with expectations such as "in line with", "meet consensus", ...

7. Keyword Identification

- Identify one or two keywords used by the analyst that best characterize the "Sentiment" variable.
- If no keywords can clearly characterize the sentiment, use "".

8. Time Outlook Identification Assign the time outlook based on the report's content. Remember that the report was written in the year {Year}. Possible values:

- "past": Events that have already occurred.
- "present": Events happening now.
- "near-future": Events expected within 1-3 years or short-term guidance.
- "distant-future": Events expected in more than 3 years or long-term guidance.
- "unclear": When the time frame is not clear.

9. Snippet Identification

- Identify a 5 word snippet from the text to capture where the topic is first discussed.

10. Output Format For each topic, provide your answer in the following format:

`{Topic}//{Entity}//{Channel}//{Sentiment}//{Keyword}//{Timeoutlook}//{Snippet}`

- `{Topic}` is the specific topic.
- `{Entity}`: "firm," "industry," or "macroeconomy."
- `{Channel}`: "earnings/cash flow: sales", "earnings/cash flow: costs", "earnings/cash flow: margins", "earnings/cash flow: earnings", "discount rate", "valuation", or "other".
- `{Sentiment}`:"positive", "negative", "neutral", or "unclear."
- `{Keyword}`: the best keyword that captures the sentiment associated with the topic.
- `{Timeoutlook}`:"past", "present", "near-future", "distant-future", or "unclear."
- `{Snippet}`: 5 word snippet.

11. Example Here is an example of a text with the associated output to help you: "Investment summary: Inflation will raise the cost of inputs while also reducing demand. We have a negative outlook. Despite lagging behind its peers in the EV market, we remain optimistic about the firm's ability to increase its market presence over the next 5 years, as reflected in our hold recommendation. Pricing pressure in the company's main segment has compressed the firm's margins." In this example, the output should be:

```
inflation//macroeconomy//earnings/cash flow: costs//negative//negative
↪  outlook//near-future//Inflation will raise the cost
market share//firm//earnings/cash flow: sales//positive//remain
↪  optimistic//distant-future//Despite lagging behind its peers
pricing pressure//industry//earnings/cash flow:
↪  margins//negative//pressure//past//Pricing pressure in the company
```

This should be the only output you provide.

Step 2: Review the selection of valuation drivers and identify any missing relevant drivers

Prompt input: Equity report and previous analysis.

Prompt output: Omitted mental model.

1. Main Task Review the previous equity report analysis and identify all topics that clearly and directly explain the analyst's price target and that were omitted in the previous analysis, based on the provided equity report. For each omitted topic:

- Categorize it into one of the following "entity": "firm," "industry," or "macroeconomy."
- Categorize it into one of the following "channel": "earnings/cash flow: sales", "earnings/cash flow: costs", "earnings/cash flow: margins", "earnings/cash flow: earnings", "discount rate", "valuation", or "other".
- Identify its associated sentiment: "positive", "negative", "neutral", or "unclear."
- Identify the best keyword that captures the sentiment associated with the topic.
- Determine the time outlook: "past", "present", "near-future", "distant-future", or "unclear."
- Identify a 5 word snippet from the text to capture where the topic is first discussed.

Importantly, do not output topics already included in the previous analysis.

2. Equity Report

```
<equityreport>
{EquityReport}
</equityreport>
```

3. Previous Analysis

```
<Previous Analysis>
{Initial Topics}
</Previous Analysis>
```

Previous analysis output structure is:

```
{Topic}//{Entity}//{Channel}//{Sentiment}//{Keyword}//{Timeoutlook}//{Snippet}
```

Where

- `{Topic}` is a topic.
- `{Entity}`: "firm," "industry," or "macroeconomy."
- `{Channel}`: "earnings/cash flow: sales", "earnings/cash flow: costs", "earnings/cash flow: margins", "earnings/cash flow: earnings", "discount rate", "valuation", or "other".
- `{Sentiment}`:"positive", "negative", "neutral", or "unclear."
- `{Keyword}`: the best keyword that captured the sentiment associated with the topic.
- `{Timeoutlook}`:"past", "present", "near-future", "distant-future", or "unclear."
- `{Snippet}`: 5 word snippet.

4. Omitted Topic Identification

- Focus exclusively on omitted topics directly tied to the justification of the price target.
- Each identified omitted topic should be specific and clearly related to a single factor influencing the price target.
- Disregard general statements or any information not directly relevant to the price target justification.

5. Entity Identification

- Assign each omitted topic to the most appropriate entity: "firm" (company-specific factors), "industry" (sector or industry trends), or "macroeconomy" (broader economic conditions).
- If an omitted topic could reasonably belong to more than one entity, select the entity that best fits the context provided in the report.

6. Channel Identification Assign each omitted topic to one of the following channels based on its context:

- "earnings/cash flow: sales": If related to sales or revenues.
- "earnings/cash flow: costs": If related to costs or operating expenses.
- "earnings/cash flow: margins": If related to profitability margins.
- "earnings/cash flow: earnings": If related to earnings, EBITDA, EBIT, or cash flows.
- "discount rate": If related to terms like "discount rate", "weighted average cost of capital/wacc", "equity beta", "idiosyncratic risk", "credit rating", "systematic risk", "equity risk premium (erp)", "market risk premium (mrp)", "interest rate", "risk-free rate (rf)", or "treasury yield".
- "valuation": If related to valuation terms such as "undervalued," "overvalued," or "fair-valued."
- "other": If none of the above apply.

7. Sentiment Identification Determine the sentiment associated with each omitted topic, selecting one of the following:

- "positive": Indicates optimism or favorable conditions such as "beat expectations", "surpass expectations", "confidence", "bullish", "optimistic", ...
- "negative": Indicates pessimism or unfavorable conditions such as "uncertainty", "challenges", "caution", "concerns", "bearish", "underperform expectations", ...
- "neutral": Indicates neutrality or conditions in line with expectations such as "in line with", "meet consensus", ...

8. Keyword Identification

- Identify one or two keywords used by the analyst that best characterize the "Sentiment" variable.
- If no keywords can clearly characterize the sentiment, use "".

9. Time Outlook Identification Assign the time outlook based on the report's content. Remember that the report was written in the year {Year}. Possible values:

- "past": Events that have already occurred.
- "present": Events happening now.
- "near-future": Events expected within 1-3 years or short-term guidance.
- "distant-future": Events expected in more than 3 years or long-term guidance.
- "unclear": When the time frame is not clear.

10. Snippet Identification

- Identify a 5 word snippet from the text to capture where the topic is first discussed.

11. Output Format For each topic that is not already included in the previous analysis, provide your answer in the following format:

`{Topic}//{Entity}//{Channel}//{Sentiment}//{Keyword}//{Timeoutlook}//{Snippet}`

- `{Topic}` is the specific omitted topic.
- `{Entity}`: "firm," "industry," or "macroeconomy."
- `{Channel}`: "earnings/cash flow: sales", "earnings/cash flow: costs", "earnings/cash flow: margins", "earnings/cash flow: earnings", "discount rate", "valuation", or "other".
- `{Sentiment}`:"positive", "negative", "neutral", or "unclear."
- `{Keyword}`: the best keyword that captures the sentiment associated with the topic.
- `{Timeoutlook}`:"past", "present", "near-future", "distant-future", or "unclear."
- `{Snippet}`: 5 word snippet.

12. Example Here is an example of a text with the associated output to help you: "Investment summary: Inflation will raise the cost of inputs while also reducing demand. We have a negative outlook. Despite lagging behind its peers in the EV market, we remain optimistic about the firm's ability to increase its market presence over the next 5 years, as reflected in our hold recommendation. Pricing pressure in the company's main segment has compressed the firm's margins." In this example, the output should be:

```
inflation//macroeconomy//earnings/cash flow: costs//negative//negative
↪  outlook//near-future//Inflation will raise the cost
market share//firm//earnings/cash flow: sales//positive//remain
↪  optimistic//distant-future//Despite lagging behind its peers
pricing pressure//industry//earnings/cash flow:
↪  margins//negative//pressure//past//Pricing pressure in the company
```

This should be the only output you provide.

Step 3: Generate Standardized Labels for the Topics

Prompt input: Equity report, combined previous analysis, standardized topic lists.

Prompt output: standardized labels.

1. Main Task Identify the standardized labels that best match **each** topic from the list below.

- It is crucial to label every topic from the Topic List.

3. Previous Analysis

<topic_list>
{Topic List}
</topic_list>

Previous analysis output structure is:

{Topic}//{Sentiment}//{Timeoutlook}//{Snippet}

Where

- {Topic} is a topic.
- {Sentiment}:"positive", "negative", "neutral", or "unclear."
- {Timeoutlook}:"past", "present", "near-future", "distant-future", or "unclear."
- {Snippet}: is a 5 word snippet to help you identify where in the report this topic is discussed.

3. Equity Report

<equityreport>
{EquityReport}
</equityreport>

4. Instructions 1

- Use the "Standardized Label List 1" to label each topic if there is a clear and direct match in the following list:

```
<Standardized Label List 1>
{Standardized Label 1}
</Standardized Label List 1>
```

To apply the standardized labels "undervalued," "overvalued," and "fair-valued," the topic must involve a comparison of the current valuation with either the firm's peers or a historical reference point.

- Undervalued: Use this label when the topic suggests the stock is priced lower than its peers or historical benchmarks. Look for expressions such as "inexpensive," "undervalued," "at a discount," "cheap," or "trades below."
- Overvalued: Apply this label when the topic indicates the stock is priced higher than its peers or historical norms. Key expressions include "expensive," "overvalued," "trades at a premium," or "trades above."
- Fair-Valued: Use this label when the topic implies the stock's valuation is in line with its peers or historical standards. Look for phrases like "compares to," "similar to" or "trades on par."

If none of the labels are a clear and direct match for the topic, try this list instead as a last resort:

```
<Standardized Label List 2>
{Standardized Label 2}
</Standardized Label List 2>
```

Important: It is crucial to only use the labels provided in those two lists.

5. Refinement

- Organize the final list logically.

6. Output Format For each topic, provide your answer in the following format:

```
{Topic}//{Standardized Label}
```

- `{Topic}` is the specific omitted topic.
- `{Standardized Label}`: is the label that best matches the topic.

This should be the only thing that you output!

12. Example

```
Increased sales in Asia//sales and revenues
New competitors//new entrants
US GDP growth//gross domestic product (gdp)
```

make sure that you labeled all the topics contained in topic list.

Appendix C LLM Prompt To Extract Valuation Methods

Analyze the following text from an equity analyst report and extract:

1. **Valuation Methods Mentioned**
2. **Valuation Methods Explicitly Used to Derive the Price Target (PT, TP, PO, OP)**

STEP 1: Identify All Valuation Methods Mentioned

From the report text, determine whether any of the following valuation methods are mentioned, either by name or through a recognized variant:

- **DCF**: Discounted Cash Flow (DCF) / Net Present Value (NPV) / Dividend Discount Model (DDM)
- **PE**: Price-to-Earnings (P/E or PER) / PEG Ratio
- **PB**: Price-to-Book (P/B)
- **PS**: Price-to-Sales (P/S) / Price-to-Revenue
- **PCF**: Price-to-Cash Flow (P/CF) / Price-to-Free Cash Flow (P/FCF) / Price-to-Unlevered Free Cash Flow (P/UFCF)
- **EVEBITDA**: Enterprise Value to EBITDA (EV/EBITDA) / EV to EBIT (EV/EBIT)
- **EVSales**: Enterprise Value to Sales (EV/Sales) / EV to Revenue (EV/Revenue)
- **EVCF**: Enterprise Value to Cash Flow (EV/CF) / Enterprise Value to Free Cash Flow (EV/FCF) / Enterprise Value to Unlevered Free Cash Flow (EV/UFCF)
- **EVBV**: Enterprise Value to Book Value (EV/BV)
- **NAV**: Net Asset Value (NAV)
- **Liquidation Value**
- **Replacement Cost Method**
- **SOTP**: Sum-of-the-Parts
- **EVA**: Economic Value Added / Residual Income Model (RIM) / Economic Profit
- **Dividend Yield**
- **Other Multiples**: Comparables-based multiples not listed above
- **Other Valuation Method**: Any other method not captured by the categories above

For each method listed in this list, return `1` if present and `0` if absent.

STEP 2: Identify Primary Valuation Methods Used for Price Target

Objective: Identify valuation methods that are explicitly and directly used to calculate the price target (PT/TP/PO/OP).
Important: Do not collect a method used only to derive "Fair Value" (not price target). Fair Value is not a synonym of Price Target.

Core Requirement: The report must explicitly state that a valuation method was used to derive the price target. Look for language such as:

- "We derive our price target using..."
- "Our target is based on..."
- "We apply [method] to arrive at our price target..."

2.1 Classification Rules

INCLUDE if:

- The method is explicitly described as contributing to the price target calculation
- Multiple methods are combined (e.g., "50% DCF, 50% P/E multiple")
- A multiple is used with growth/discount rates (classify under both the multiple AND DCF)

EXCLUDE if:

- Used only for reference, support, or "sanity check"
- Used only for comparison against the price target
- Used only to calculate terminal value within a larger model
- Used only to derive "Fair Value" (not price target)
- Referenced only as implied market metrics for comparison
- Used only to inform valuation judgment without direct calculation

Important Distinctions:

- Enterprise Value (EV) and Price (P) should be treated as separate methods
- Sum-of-the-Parts (SOTP) must be explicitly stated as used for price target calculation

Evidence Standard:

- Must be literally written in the report

- Do not infer usage from structure or implied language
- Direct statements only (e.g., "We use sum-of-the-parts to value the firm")

Output: List all qualifying valuation methods. If multiple methods apply, list all that meet the criteria above.

STEP 3: Supporting Evidence Snippet

Requirement: For each identified valuation method, provide a **5-word snippet** from the equity analyst report that supports your classification.

If no supporting snippet available: Return "null"

STEP 4: Output Format

Return the results in the following JSON format:

```
{
  "valuation_all_methods": {
    "DCF": 0,
    "PE": 0,
    "PB": 0,
    "PS": 0,
    "PCF": 0,
    "EVEBITDA": 0,
    "EVSales": 0,
    "EVCF": 0,
    "EVBV": 0,
    "NAV": 0,
    "Liquidation Value": 0,
    "Replacement Cost Method": 0,
    "SOTP": 0,
    "EVA": 0,
    "Dividend Yield": 0,
    "Other Multiples (comparables)": 0,
    "Other Valuation Method": 0
```

```
},
"methods_tied_to_target": {
  "DCF": 0,
  "PE": 0,
  "PB": 0,
  "PS": 0,
  "PCF": 0,
  "EVEBITDA": 0,
  "EVSales": 0,
  "EVCF": 0,
  "EVBV": 0,
  "NAV": 0,
  "Liquidation Value": 0,
  "Replacement Cost Method": 0,
  "SOTP": 0,
  "EVA": 0,
  "Dividend Yield": 0,
  "Other Multiples (comparables)": 0,
  "Other Valuation Method": 0
},
"snippets_all_methods": {
  "DCF": "null",
  "PE": "null",
  "PB": "null",
  "PS": "null",
  "PCF": "null",
  "EVEBITDA": "null",
  "EVSales": "null",
  "EVCF": "null",
  "EVBV": "null",
  "NAV": "null",
  "Liquidation Value": "null",
  "Replacement Cost Method": "null",
  "SOTP": "null",
  "EVA": "null",
  "Dividend Yield": "null",
```

```
    "Other Multiples (comparables)": "null",
    "Other Valuation Method": "null"
  },
  "snippets_tied_to_target": {
    "DCF": "null",
    "PE": "null",
    "PB": "null",
    "PS": "null",
    "PCF": "null",
    "EVEBITDA": "null",
    "EVSales": "null",
    "EVCF": "null",
    "EVBV": "null",
    "NAV": "null",
    "Liquidation Value": "null",
    "Replacement Cost Method": "null",
    "SOTP": "null",
    "EVA": "null",
    "Dividend Yield": "null",
    "Other Multiples (comparables)": "null",
    "Other Valuation Method": "null"
  }
}
```

JSON Value Guidelines:

- Use `1` for methods that are present/used
- Use `0` for methods that are not present/not used
- Use `"null"` (with quotes) for snippets when no supporting text is available
- Use actual 5-word snippets in quotes when available (e.g., `"target based on DCF model"`)
- Ensure proper JSON formatting with commas and closing braces

Important: Return only valid JSON. Do not include any additional text or explanations.

5. Classification Examples

Primary Valuation Method Examples:

INCLUDE only methods used to generate the price target directly:

Example 1:

> "Our DCF-based PT implies 3.2x/2.7x '23/'24E EV/sales"
> *Classification: DCF only (EV/Sales is implied comparison, not generation method)*

Example 2:

> "Our valuation is based on DCF and supported by ∼16x our 2014 EPS"
> *Classification: DCF only (P/E provides support, not generation)*

Example 3:

> "Our $165 price target is derived from a blend of two valuation techniques, equally weighted: 1) relative P/E valuation, which yields a value of $177 per share, and 2) our discount cash flow (DCF) model, which yields a value of $153"
> *Classification: Both DCF and P/E*

DCF Classification Guidelines:

Tag as DCF when mentioning NPV or DCF inputs:

Example 4:

> "We added $15 per share of additional value to reflect the estimated net present value (NPV)"
> *Classification: DCF*

Example 5:

> "We apply a long-term (post-tax) weighted average cost of capital (WACC) of 9.7%"
> *Classification: DCF*

Key Principle: Distinguish between methods that **generate** the target versus those that **support**, **compare**, or **validate** it.

Appendix D LLM Prompt to Extract Inflation Narratives

Analyze the following text from an equity analyst report and extract all arguments that capture the analyst's reasoning about the causes and consequences of inflation.

`"EquityReport"`

STEP 1: Output Format

If you find explicit discussion of inflation, its causes, or its consequences for firm valuation, report each instance using the following format:

`{raw_node}::{cause}::{valuation_channel}::{sentiment}::{snippet}::{confidence}`

If no relevant discussion is found, output a single space character: `" "` — do not provide any other text, explanation, or formatting in that case.

STEP 2: Field Definitions

- **raw_node**: The key topic or concept mentioned (e.g., "Cost inflation", "Production costs inflation").
- **cause**: The analyst's stated or implied cause of inflation. Choose one of the following:
  - Demand/Government Spending
  - Demand/Monetary Policy
  - Demand/Pent-up demand
  - Demand/Demand shift
  - Demand/Residual
  - Supply/Supply chain issues
  - Supply/Labor
  - Supply/Energy crisis
  - Supply/Residual
  - Miscellaneous/Pandemic
  - Miscellaneous/Government mismanagement
  - Miscellaneous/Russia-Ukraine war
  - Miscellaneous/Inflation expectations
  - Miscellaneous/Base effect

  - Miscellaneous/Government debt
  - Miscellaneous/Tax increases
  - Miscellaneous/Price-gouging
  - Missing
  - Other

- **valuation_channel**: The channel through which inflation affects firm valuation. Choose one of:
  - earnings/cash flow: sales
  - earnings/cash flow: costs
  - earnings/cash flow: margins
  - earnings/cash flow: earnings
  - discount rate
  - valuation
  - missing
  - other

- **sentiment**: The tone of the analyst's discussion. Choose one of:
  - positive
  - negative
  - neutral
  - other

- **snippet**: A concise extract from the report text that supports the mapping. Use exact phrasing from the report.
- **confidence**: A rating from 1 (low confidence) to 5 (high confidence) indicating how certain you are that the snippet reflects analyst reasoning about inflation.

STEP 3: Example Output

Input Snippet:

> "Cost inflation was also a key theme in all meetings (such as labour, energy), although we sensed most Telcos also saw measures to manage this through continued cost control focus. For 2022, inflation in production costs should be largely offset by price increases (+4/+5% on 01/2/2022, a possible further increase this coming July)."

Expected Output:

```
{Cost Inflation}::{Supply/Labor}::{earnings/cash flow: costs}::{negative}::{"Cost
↪ inflation was also a key theme in all meetings (such as labour, energy), although we
↪ sensed most Telcos also saw measures to manage this through continued cost control
↪ focus."}::{5}

{Cost Inflation}::{Supply/Energy crisis}::{earnings/cash flow:
↪ costs}::{negative}::{"Cost inflation was also a key theme in all meetings (such as
↪ labour, energy), although we sensed most Telcos also saw measures to manage this
↪ through continued cost control focus."}::{5}

{Production costs inflation}::{Supply/Supply chain issues}::{earnings/cash flow:
↪ costs}::{negative}::{For 2022, inflation in production costs should be largely
↪ offset by price increases (+4/+5\% on 01/2/2022, a possible further increase this
↪ coming July).}::{5}

{Production costs inflation}::{Supply/Supply chain issues}::{earnings/cash flow:
↪ earnings}::{positive}::{For 2022, inflation in production costs should be largely
↪ offset by price increases (+4/+5\% on 01/2/2022, a possible further increase this
↪ coming July).}::{5}

{Production costs inflation}::{Supply/Supply chain issues}::{earnings/cash flow:
↪ margins}::{neutral}::{For 2022, inflation in production costs should be largely
↪ offset by price increases (+4/+5\% on 01/2/2022, a possible further increase this
↪ coming July).}::{5}
```

Reminder: This is the only output that should be returned—no additional explanation or summary should be provided.

Appendix E List of Topics and Associated Categories

Table E1: Main topics This table presents all topics included in *Standardized Label 1* used in the prompt in Step 3 of Appendix B.

Category	Topic	Category	Topic	Category	Topic
Category	Topic	Category	Topic	Category	Topic
Pricing	Pricing Power	Corp. Investment	Investment Opp.	Envr.	Energy Transition
Pricing	Pricing Pressure	Corp. Investment	Automation And Robotics	Legal And Reg. Risk	Legal Envr.
Pricing	Pricing Strategy	Prod/Utilization	Oper. Leverage And Scale	Legal And Reg. Risk	Lawsuits And Settlement
Mkt Structure	Barriers To Entry	Prod/Utilization	Utilization	Legal And Reg. Risk	Reg. Envr. And Compliance
Mkt Structure	Collusion	Financing	Credit Rating	Legal And Reg. Risk	Cybersec. And Data Protection
Mkt Structure	Customer Concentration	Financing	Credit Avail. And Access	Legal And Reg. Risk	Regulatory Approval
Mkt Structure	Mkt Share	Financing	Debt Lvl /Leverage	Dom. Govrn.	Political Envr.
Mkt Structure	Mkt Size	Financing	Cash And Liquidity	Dom. Govrn.	Fiscal Envr.
Mkt Structure	Monopoly	Financing	Stock Offerings And Issuances	Dom. Govrn.	Govrn. Debt
Mkt Structure	Oligopoly	Financing	Pe/Vc/Alt. Financing	Dom. Govrn.	Taxation
Mkt Structure	Threat Of Substitute	Financing	Risk Mgmt/Hedging	Dom. Govrn.	Subsidies
Comp. Advantage	Cost Structure	Financing	Covenants	Dom. Govrn.	Govrn. Procurement
Comp. Advantage	Econ. Of Scale	Payout	Stock Buybacks	Dom. Govrn.	Infrastructure
Comp. Advantage	Vertical Integration	Payout	Dividend Policy And Strategy	Housing Mkt	Housing Supply
Mkt Development	New Entrants	M&A	Spin-Off And Divestitures	Housing Mkt	Mortgage Rates
Mkt Development	Mkt Expansion (Industry)	M&A	Synergy And Integration	Housing Mkt	Housing Demand
Mkt Development	Mkt Expansion (Geography)	M&A	Mergers And Acquisitions	Housing Mkt	Home Prices
Customer	Customer Acquisition	M&A	Partnership/Joint-Venture	Demog. And Social	Demog. Changes
Customer	Customer Demand/Spending	M&A	Restructuring	Demog. And Social	Income Inequality
Customer	Customer Loyalty And Retention	M&A	Antitrust	Macro Disrupt.	Natural Disasters
Brand	Brand Equity And Reputation	Contracting	Licensing And Royalty	Macro Disrupt.	Pandemic And Public Health
Brand	Marketing And Advertising	Contracting	Contract Terms And Duration	Macro Disrupt.	Wars And Terrorism
Product	Product/Service Mix	Governance	CEO/Senior Mgmt.	Macro Disrupt.	Economic Sanctions
Product	Product/Service Quality	Governance	Mgmt. Guidance	Raw Input	Commodity/Raw Material
Product	Product/Service Positioning And Diff.	Governance	Corp. Governance	Raw Input	Oil And Gas
Product	Product Life Cycle And Obsolescence	Governance	Corp. Ownership Structure	GDP And Econ. Cycle	Gross Domestic Product (GDP)
Product	New Products/Services	Governance	Corp. Social Resp. (CSR)	GDP And Econ. Cycle	Economic Recovery
Product	Product Recall	Labor Mkt.	Employee Safety	GDP And Econ. Cycle	Economic Recession
Product	Product Pipeline	Labor Mkt.	Hiring And Layoff	Inflation	Inflation And CPI
Sup. Chain	Production Capacity And Planning	Labor Mkt.	Labor Strikes	Foreign Exchange	Foreign Exchange
Sup. Chain	Transportation And Logistics	Labor Mkt.	Pensions	Intl.	Intl. Trades/Import And Exports
Sup. Chain	Inventory Mgmt.	Labor Mkt.	Employee Satisfaction	Intl.	Globalization
Sup. Chain	Order Fulfil./Backlog/Backorder	Labor Mkt.	Corporate Culture	Intl.	Tariffs
Sup. Chain	Sup. Chain Agility And Flexibility	Labor Mkt.	Outsourcing	Intl.	Geopolitical Envr.
Sup. Chain	Supplier Rel. And Bargaining Power	Labor Mkt.	Wages	Industry	Industry Life Cycle
Sup. Chain	Supplier Concentration	Labor Mkt.	Unions	Discount Rate	Discount Rates/WACC/CoC
Innovation / R&D	Innovation And R&D	Labor Mkt.	Unemployment Rates	Systematic Risk	Risk Premium/Systematic Risk
Innovation / R&D	Clinical Trials	Labor Mkt.	Labor Force Participation	Monetary Policy	Interest Rate/Treasury Yield
Innovation / R&D	Trade Secrets And Patents	Labor Mkt.	Skills Shortages	Monetary Policy	Monetary Policy
Innovation / R&D	Intellectual Property (IP)	Labor Mkt.	Immigration	Valuation	Undervalued
Innovation / R&D	AI And Cognitive Applications	Envr.	Envr. Policies And ESG	Valuation	Overvalued
Innovation / R&D	Technology Lifecycle	Envr.	Pollution And Toxic Waste	Valuation	Fair-Valued
Corporate Investment	Capital Expenditure (CAPEX)	Envr.	Weather And Climate Change		

Table E2: Alternative topics This table presents all topics included in *Standardized Label 2* used in the prompt in Step 3 of Appendix B.

Category	Topic	Category	Topic	Category	Topic
Undetermined	Profitability Margins	Undetermined	Cash Flows	Undetermined	Discounted Cash Flow (DCF)
Undetermined	Sales And Revenues	Undetermined	Industry Trends	Undetermined	Sum-Of-The-Parts (SOP)
Undetermined	Costs And Expenses	Undetermined	Macroeconomic Trends	Undetermined	Other
Undetermined	Earnings, EBITDA, And EPS	Undetermined	Valuation Multiples		

Appendix F Equity Report Excerpt and LLM Output

Excerpt (Deutsche Bank, The Walt Disney Company, 2017-02-09): "**Theme Parks performance remains strong**, with **cost efficiencies in the US Parks** coming through as we expected. Mgt seems to be **planning to increase prices in the domestic parks**, or expand demand-based pricing to multi-day tickets. They hinted at some pricing action, without specificity. **International Parks' performance improved**, generating OI growth."

Output structure:

- {Theme park performance}//{firm}//{earnings/cash flow: earnings}//{positive}//{strong performance}//{present}//{**Theme park performance remains strong**}
- {Cost efficiencies}//{firm}//{earnings/cash flow: costs}//{positive}//{cost savings}//{present}//{**cost efficiencies in the US**}
- {Park pricing}//{firm}//{earnings/cash flow: sales}//{positive}//{price increase}//{near-future}//{**planning to increase prices in**}
- {International parks performance}//{firm}//{earnings/cash flow: earnings}//{positive}//{improved results}//{past}//{**International Parks' performance improved**}

Appendix G Multiples Reference Points

Firm Performance as a Reference Point

Forward-Looking Interpretation of Historical Data

1. **Equita, LUX.MI^C19, 2004-02-03:**
   *"**[We apply a]** P/E of 20.5x vs. 22.3x historical average in the last 5 years, despite the potential synergies that can be achieved in the following years."*
2. **UBS Equities, URI.N, 2014-01-21:**
   *"The 6.1x **[we use]** is within the historical range, which we think is supported by improvement in the business (execution and mix have driven margins higher, and cash flow and returns are improving)."*
3. **Jefferies, TXN.OQ, 2020-02-05:**
   *"A 24x P/E **[used in our valuation]** is in the top half of the 5-yr range of 16x–25x and we think warranted at the bottom of inventory correction and given its exposure to secular trends (IoT, Auto, Industrial, Consolidation)."*
4. **Morgan Stanley, EIGE.L^C20, 2008-02-27:**
   "P/E Take the market 2010 P/E of 10.5x and assume a 20% discount for below-market long-term growth."

Backward-Looking Interpretation of Historical Data

1. **UBS Equities, HES.N, 2013-04-24:**
   "$75 PT assumes 4.8x normalized 2013E DACF, in-line with its historical average."
2. **Deutsche Bank, GLW.N, 2023-04-26:**
   "P/E valuation ($35, using 18x 2023E core EPS, in line with the historical average)."
3. **Kepler Chevreux, BDTG.DE, 2023-01-13:**
   "At our TP [target price], BDT would be trading in line with the average of its five-year EV/EBITDA multiple (6x)."

Peer Group and Industry as a Reference Point

1. **UBS, BBVA.MC, 2009-03-04:**
   "We believe Spirent is cheap at current levels and based on our forecasts Spirent trades on 7x 2010E P/E and an EV/sales of 0.8x, compared to Agilent on 9.4x and 0.9x respectively."

2. **Bank of America, TT.N, 2017-02-03:**

   "Our PO of $85 is based on applying a 17x P/E to our '18 EPS forecast, broadly in line with the average for the peer group."

3. **Nomura, SCC.BK, 2013-01-03:**

   "We assign a multiple of 10x, in line with its regional chemical peers."

4. **Credit Suisse, PER.L^L14, 2014-03-14:**

   "The target price is derived by putting Perform on 28.5x, a 20% discount to internet peers (on 35.6x 2014e)."

Appendix H Appendix Figures and Tables

(A) **Credit-Suisse.**



(B) **Deutsche Bank.**



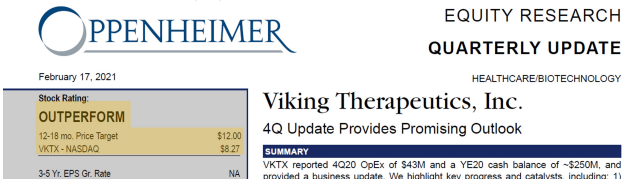
(C) **Barclays.**



(D) **Scotia Bank.**



(E) **Oppenheimer.**



(F) **Bank of Montreal.**



(G) **Jefferies.**



(H) **Evercore.**



Figure H.1:
Examples of Equity Report First Pages

This figure displays the first-page headers from eight equity reports issued by various brokerage houses. Across reports, two pieces of information, (1) the investment recommendation (buy/hold/sell) and (2) the price target, are consistently emphasized, as highlighted in yellow. Their visual prominence relative to other valuation elements, such as earnings forecasts, underscores their first-order importance in equity research. This consistency aligns with the emphasis placed on these outputs by external ranking agencies such as the *Wall Street Journal* and *Institutional Investor*, which evaluate analysts primarily on the basis of recommendations and price target performance.

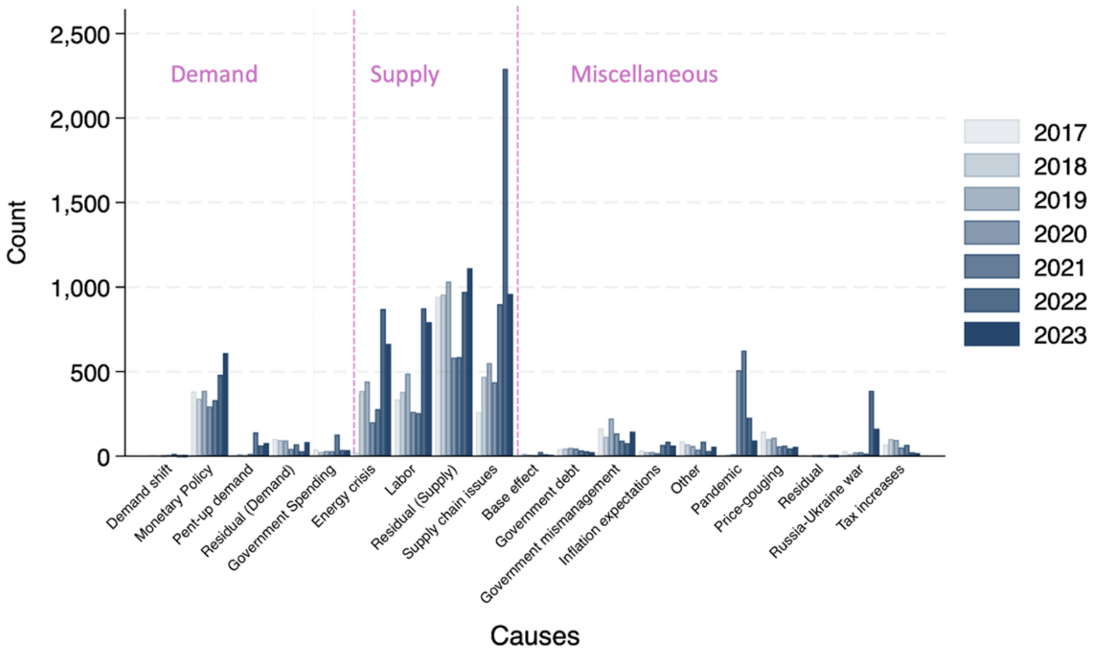
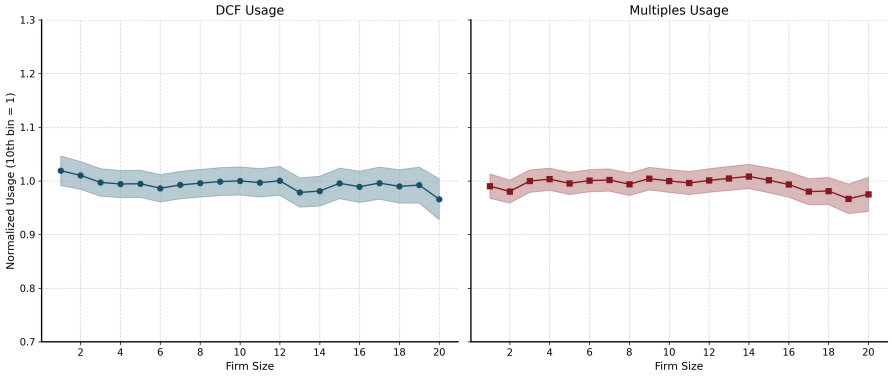


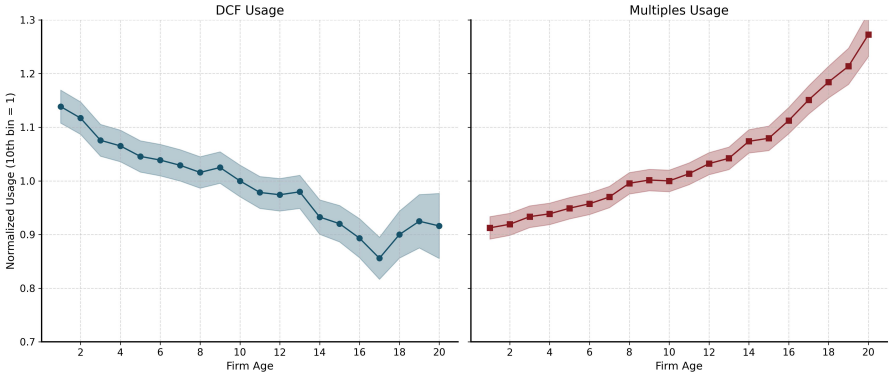
Figure H.2: Inflation Narratives

This figure shows analysts' reasoning related to inflation, extracted using the prompt provided in Appendix D, which imposes and elicits the narratives in Andre et al. (2024a).

(A) Size.



(B) Firm Age.



(C) Market-to-Book.

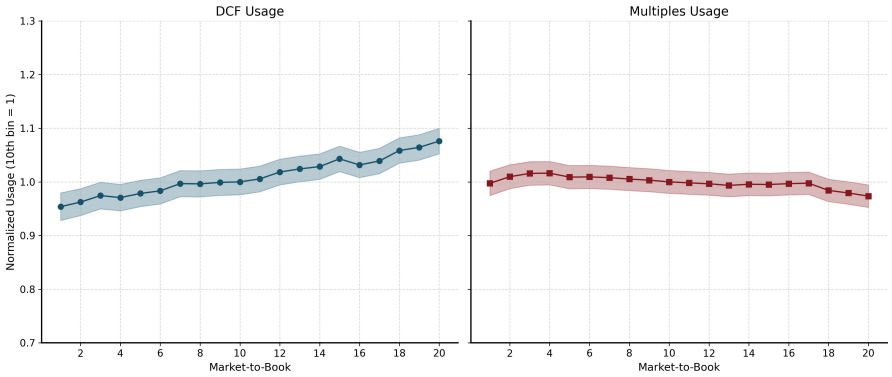


Figure H.3:
Valuation Method Choice and the Cross-Section of Firms – Robustness

This figure repeats Figure V, estimated with analyst-firm fixed effects. The plots the propensity to use either a DCF method (blue) or a multiples approach (red) as a function of firm characteristics. Panels A, B, and C show results by firm size (total assets), firm age (measured from the year of incorporation), and market-to-book ratio, respectively. For each characteristic, firms are sorted into 20 equal-sized bins (5%) within each year of the sample. See Figure V for additional details. All series are normalized to equal 1 at ventile 10 to allow for comparison of relative trends. Confidence intervals are at the 90th percentile, and standard errors are heteroskedasticity-robust.

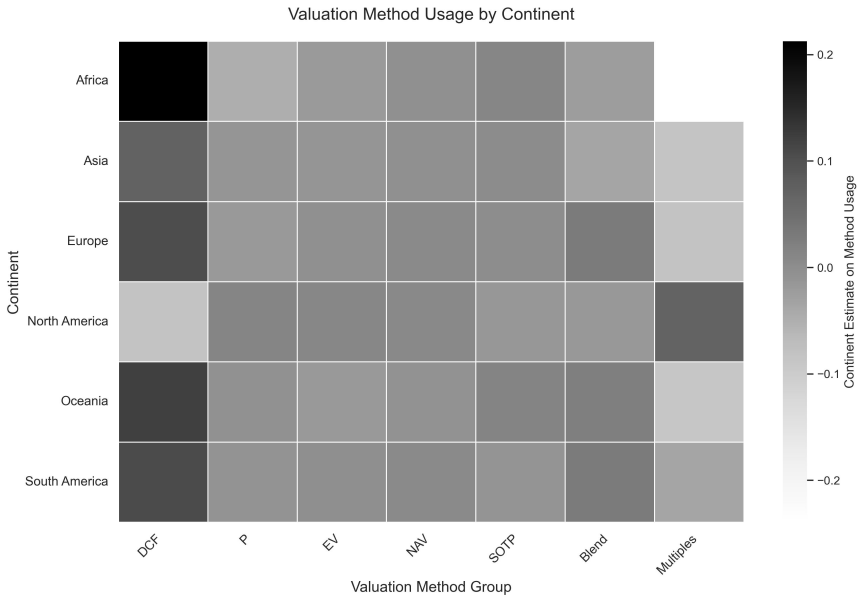


Figure H.4: Valuation Methods – Heterogeneity by Geography

This figure shows additional evidence on heterogeneity in valuation methods by geography, replicating Panel C of Figure IV while absorbing firm-year fixed effects.

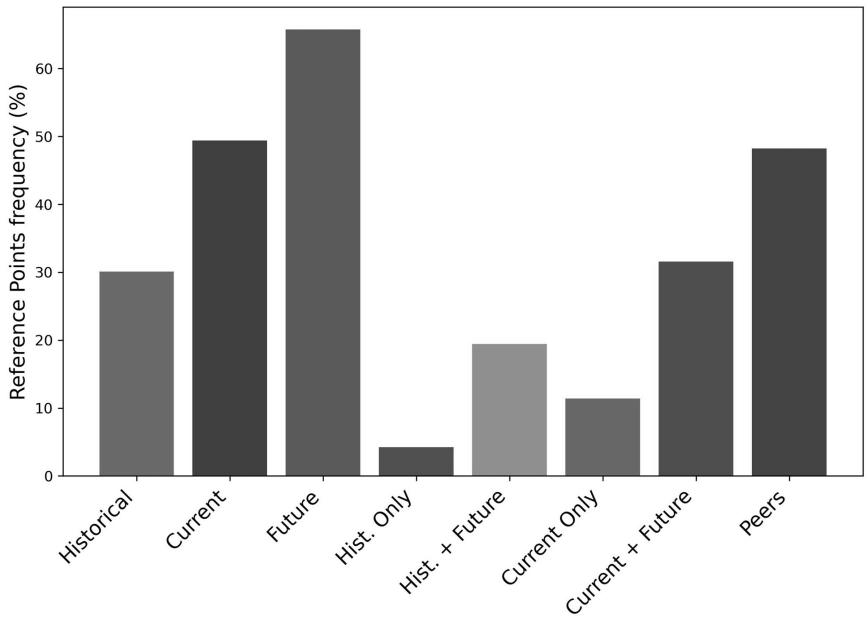


Figure H.5: Multiples Reference Points

This figure shows the distribution of reference point types used in multiples-based valuation methods, including firm-specific, peer-based, historical, current, and forward-looking benchmarks.

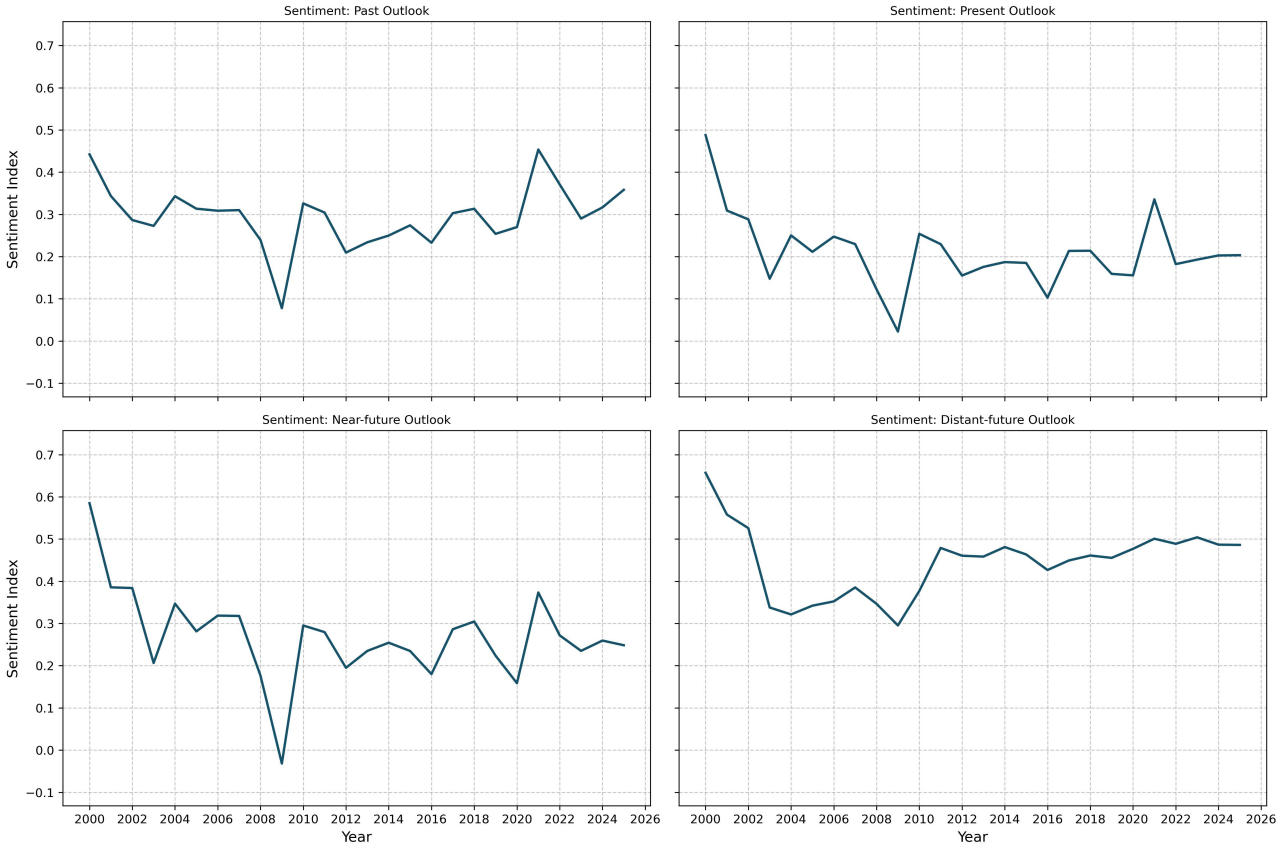


Figure H.6:
Average Sentiment by Time Outlook

This figure plots average sentiment over time by time outlook—past, present, near future (1–3 years), and distant future (>3 years)—for the period 2000–2025. The x-axis denotes years. The y-axis represents the average sentiment in reports written in a given year, measured on a scale from −1 (negative) to +1 (positive), with 0 indicating neutral sentiment.

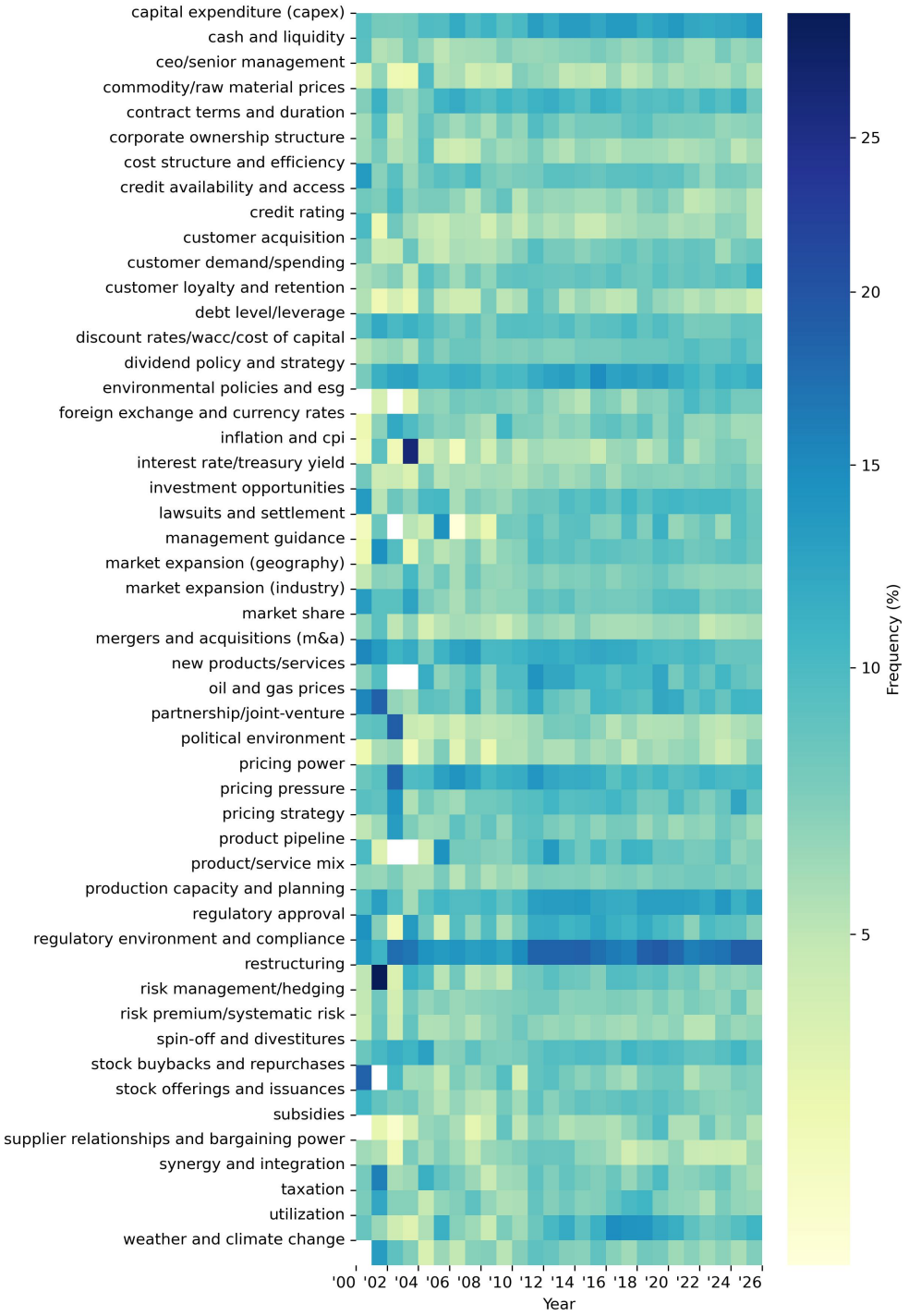


Figure H.7:
Topic Attention Allocation Over Time – Utilities

This figure plots attention allocation to the 50 most frequently discussed topics across utility firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.

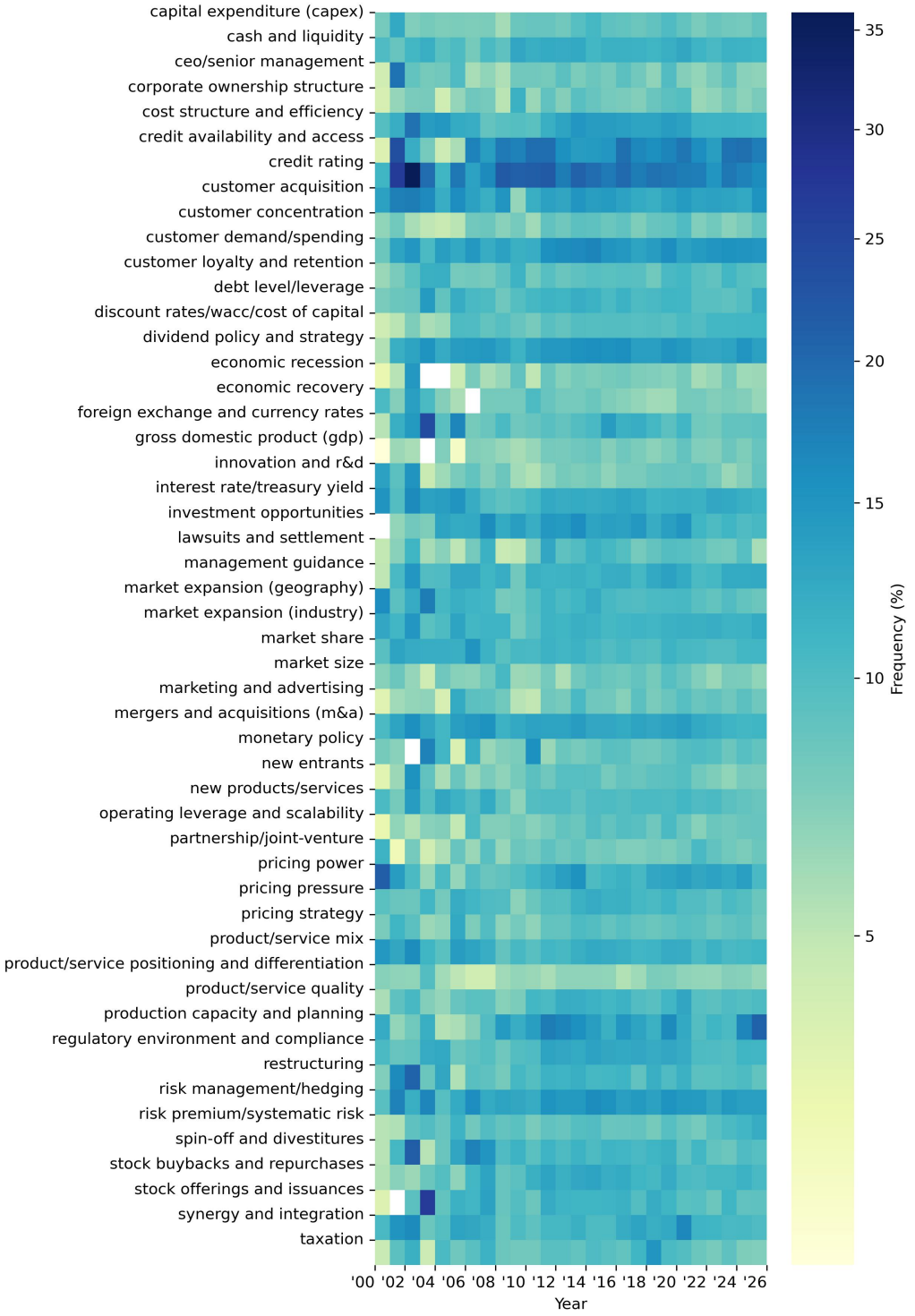


Figure H.8:
Topic Attention Allocation Over Time – Finance and Insurance

This figure plots attention allocation to the 50 most frequently discussed topics across finance and insurance firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.

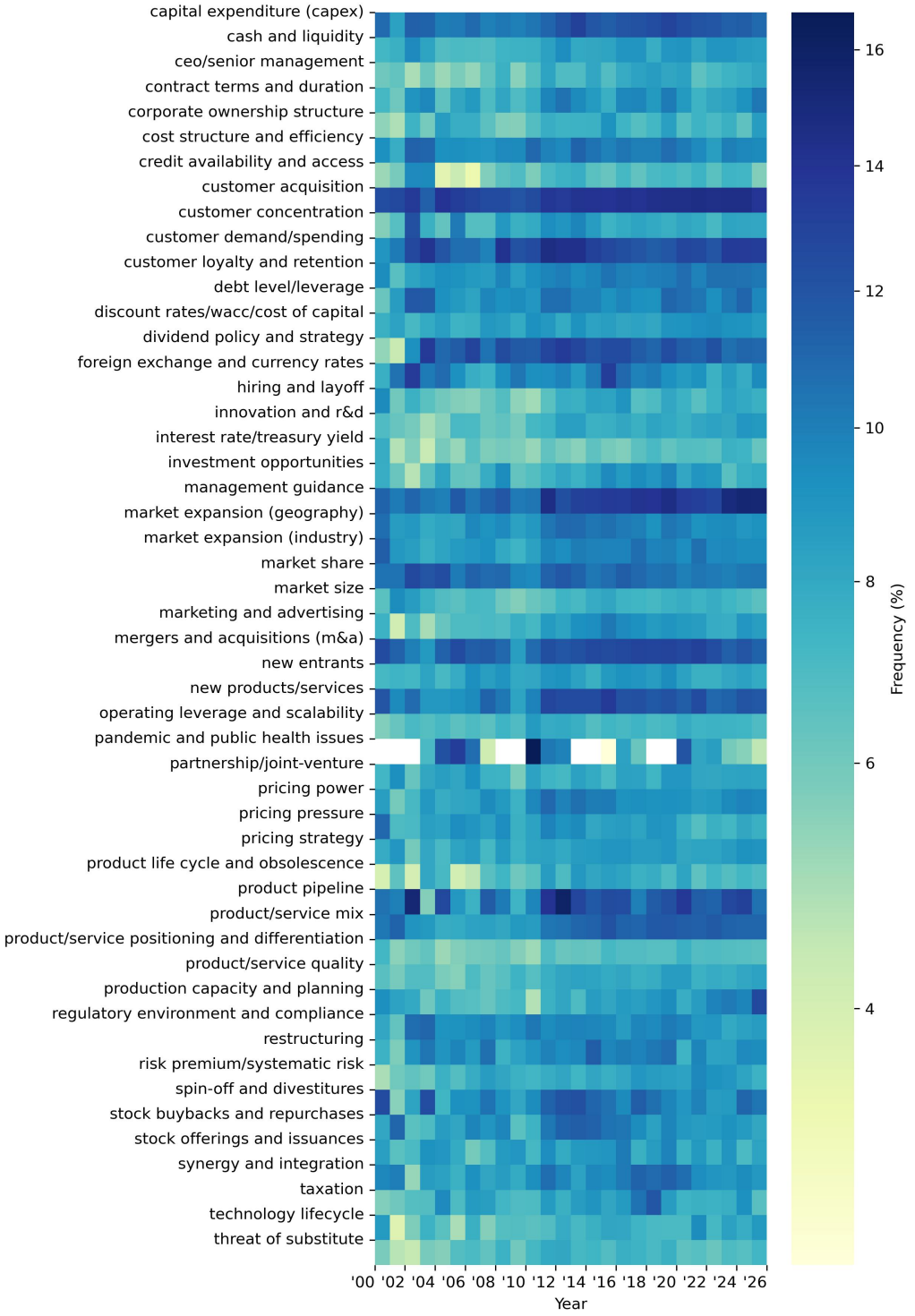


Figure H.9:
Topic Attention Allocation Over Time – Information

This figure plots attention allocation to the 50 most frequently discussed topics across information firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.



Figure H.10:
Topic Attention Allocation Over Time – Professional, Scientific, and Technical Services

This figure plots attention allocation to the 50 most frequently discussed topics across professional services firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.



Figure H.11:
Topic Attention Allocation Over Time – Mining / Oil & Gas

This figure plots attention allocation to the 50 most frequently discussed topics across mining and oil & gas firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.

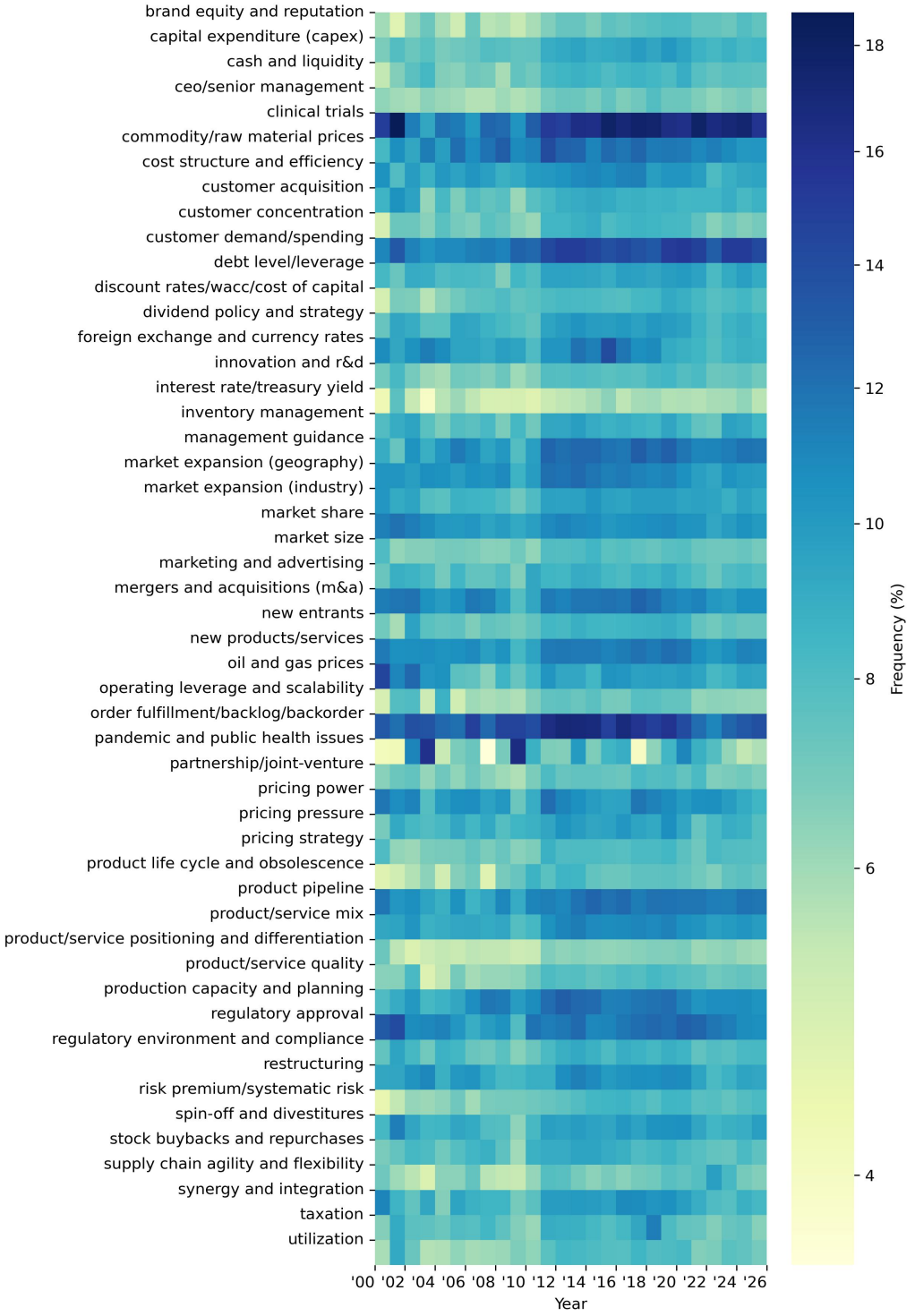


Figure H.12:
Topic Attention Allocation Over Time – Manufacturing

This figure plots attention allocation to the 50 most frequently discussed topics across manufacturing firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.



Figure H.13:
Topic Attention Allocation Over Time – Retail Trade

This figure plots attention allocation to the 50 most frequently discussed topics across retail trade firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.



Figure H.14:
Topic Attention Allocation Over Time – Transportation

This figure plots attention allocation to the 50 most frequently discussed topics across transportation firms in our sample from 2000 to 2025. The x-axis denotes the year in which reports were produced, and the y-axis lists the top 50 topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.

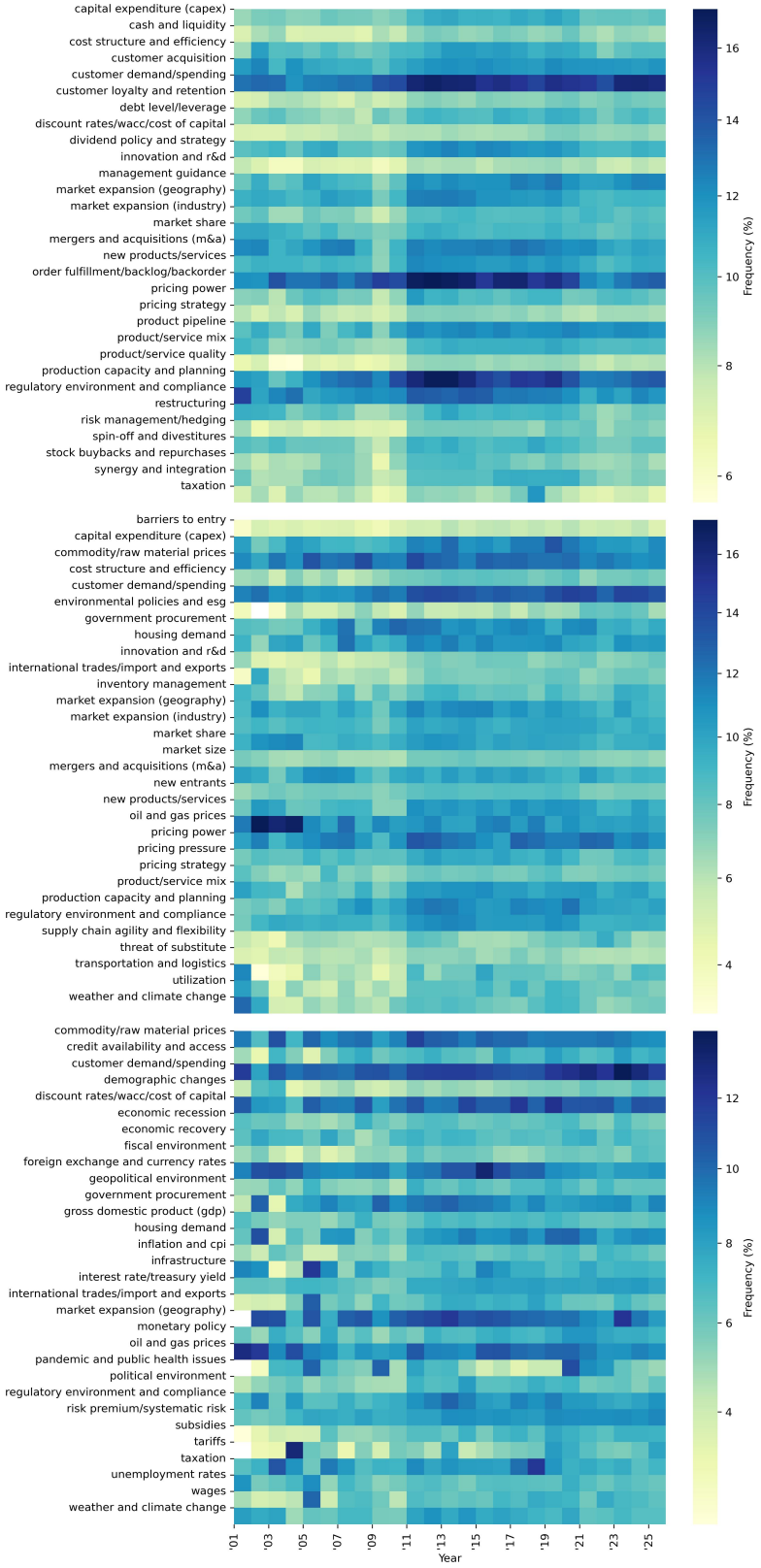


Figure H.15:
Topic Attention Allocation Over Time – Entities

This figure plots the attention allocation to the 20 most important topics for the three types of entities we focus on in the sample–firm, industry, and macroeconomy–for the period 2000–2025. The first block denotes firm topics. The second block denotes industry topics. The third block denote the macroeconomy topics. The x-axis denotes the year in which reports were produced, and the y-axis lists the topics. The color-coded heatmap is shown on the right, with yellow shades indicating lower values and dark blue shades representing the highest values in the sample. A non-linear gamma scale is used to enhance the visibility of variation across the full range of intensities. Frequency (in percent) corresponds to the share of the average report in a given year that discusses each topic.

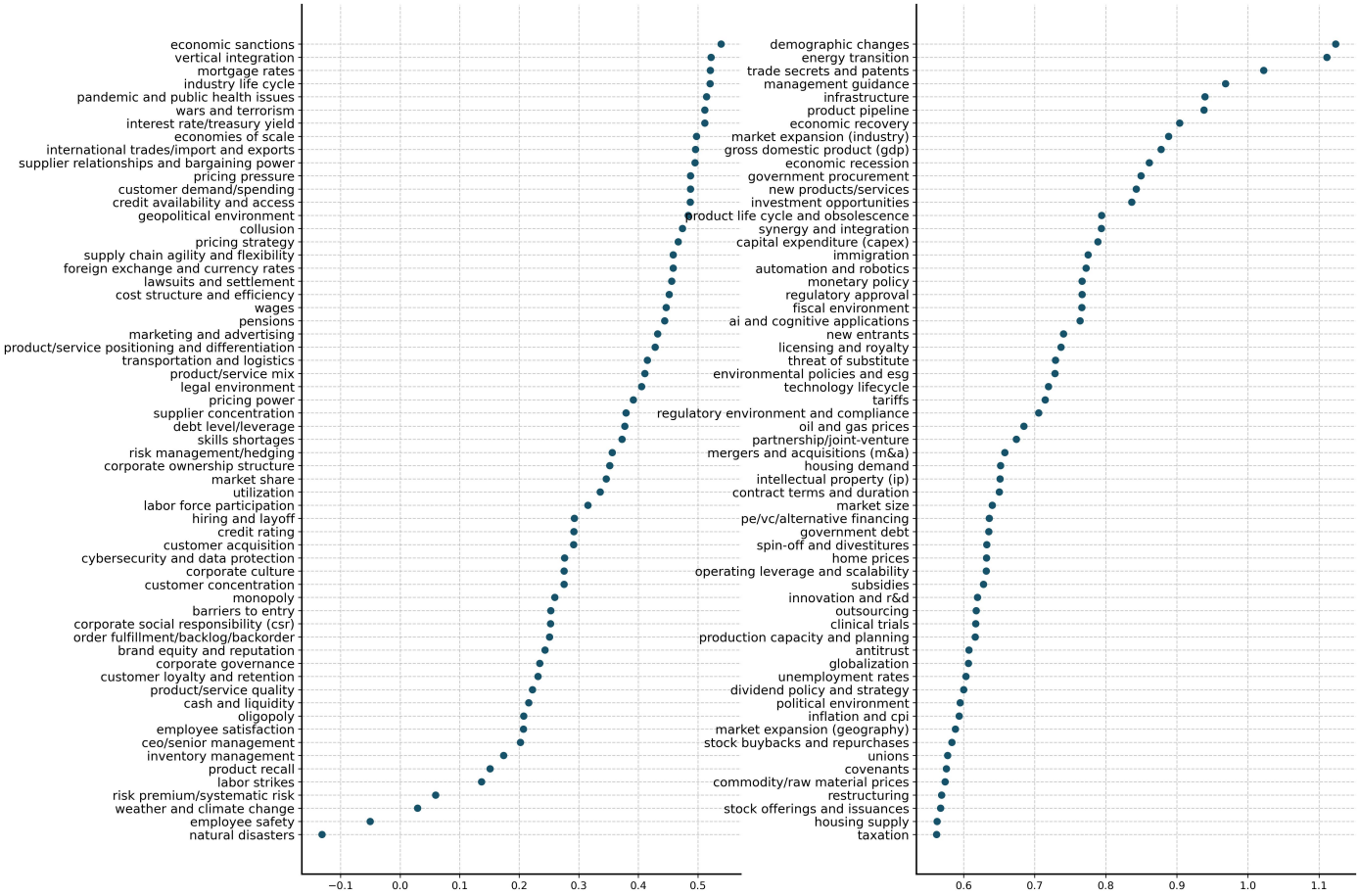


Figure H.16: Topic Outlook

This figure plots the average topic outlook associated with the most common topics in our sample over the period 2000–2025. The x-axis represents the average outlook for each topic, calculated on a scale of −1 (past), 0 (present), and +1 (near-future), and +2 (distant-future). The y-axis displays the corresponding topic labels.

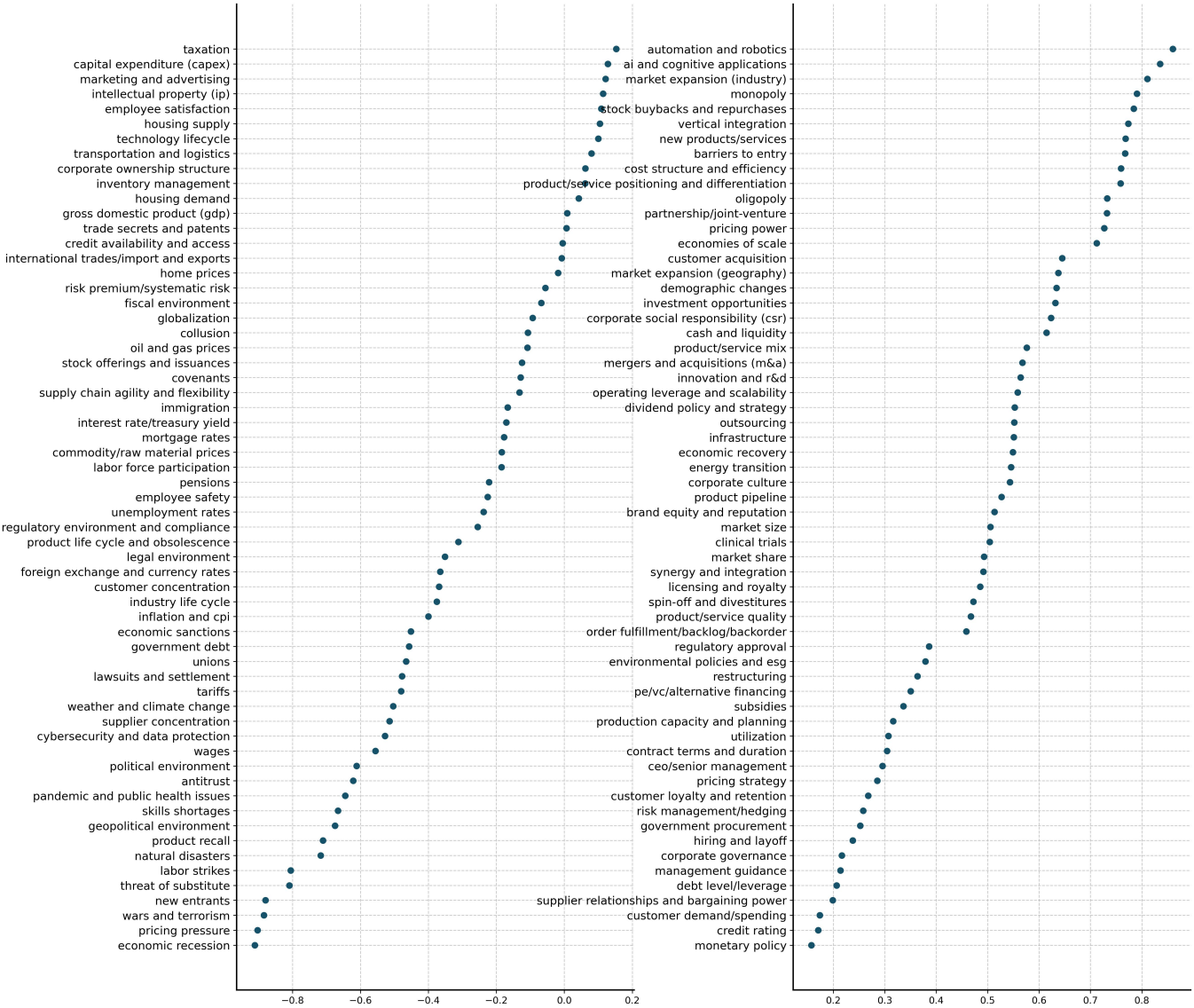


Figure H.17: Topic Sentiment

This figure plots the average sentiment associated with the most common topics in our sample over the period 2000–2025. The x-axis represents the average sentiment for each topic, calculated on a scale of −1 (negative), 0 (neutral), and +1 (positive). The y-axis displays the corresponding topic labels.

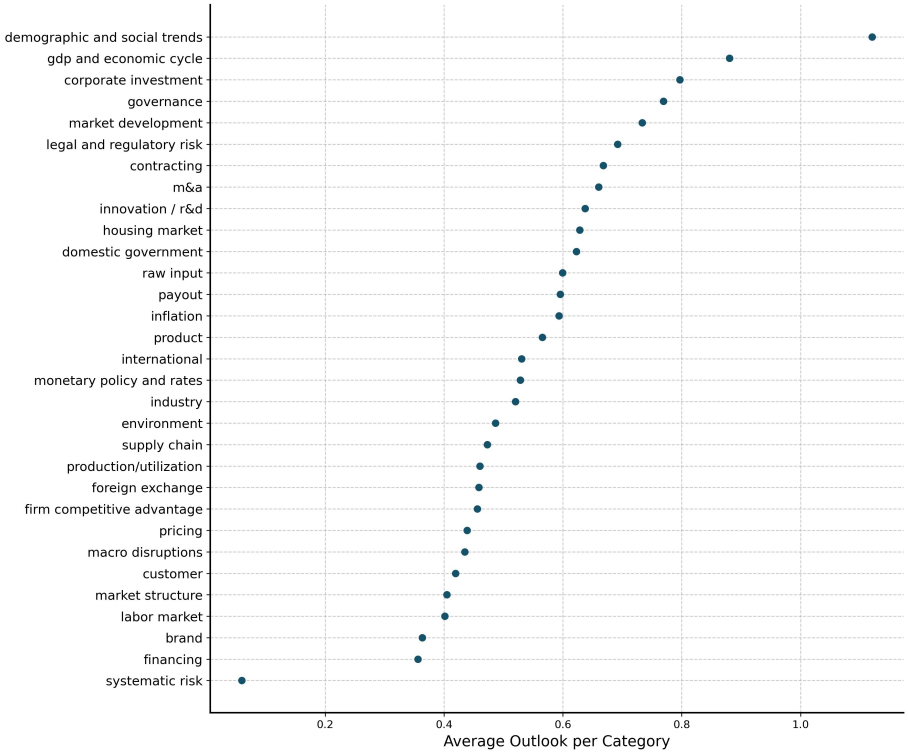


Figure H.18: Topic Category Outlook

This figure plots the average category outlook in our sample over the period 2000–2025. The x-axis represents the average outlook for each topic category, calculated on a scale of −1 (past), 0 (present), and +1 (near-future), and +2 (distant-future). The y-axis displays the corresponding topic labels.

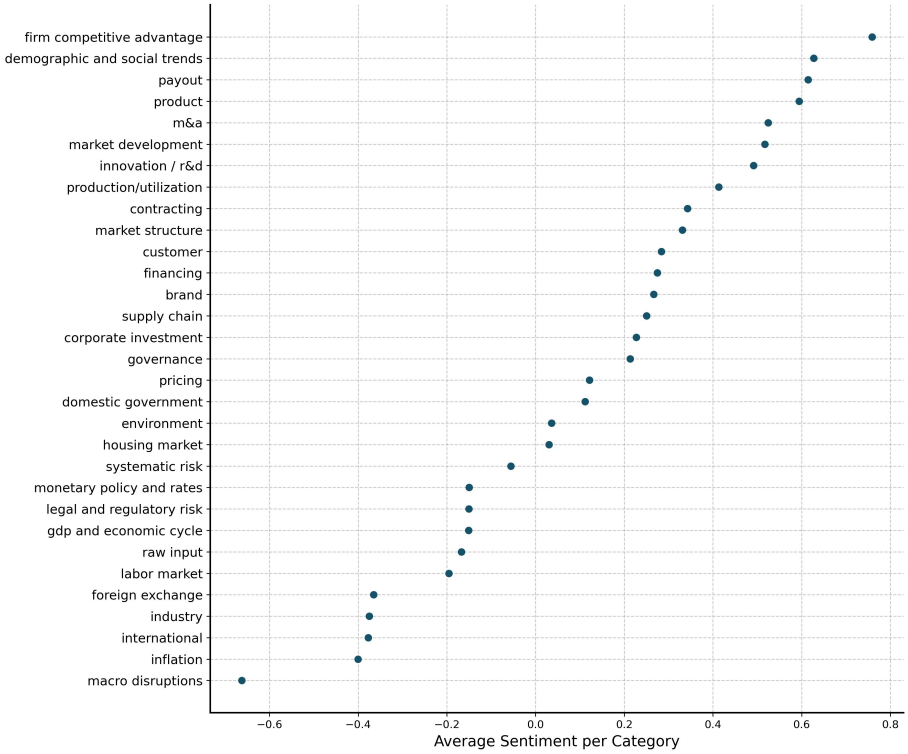
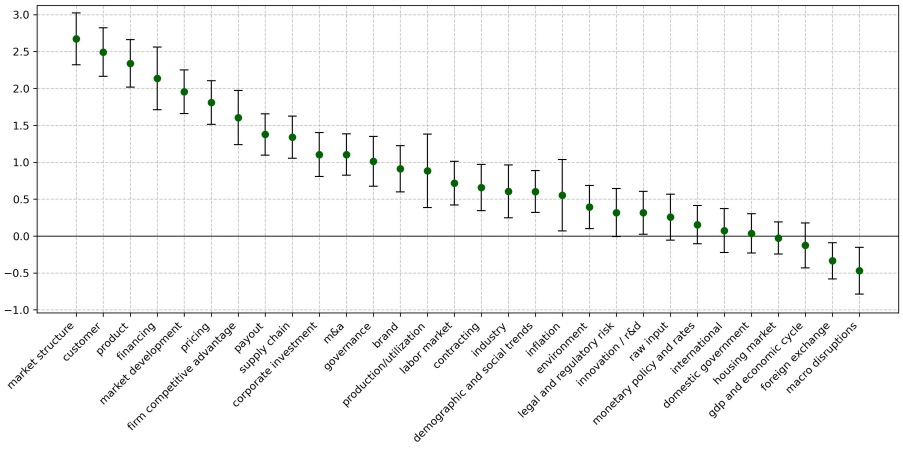


Figure H.19: Topic Category Sentiment

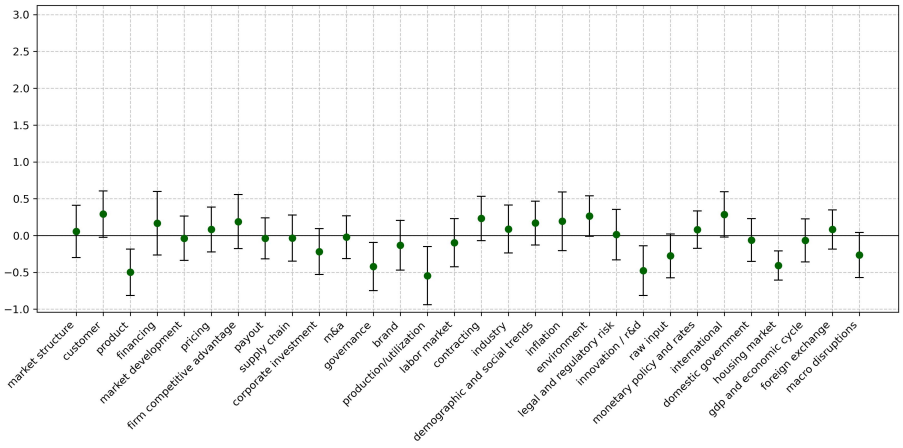
This figure plots the average topic category sentiment in our sample over the period 2000–2025. The x-axis represents the average sentiment for each category, calculated on a scale of −1 (negative), 0 (neutral), and +1 (positive). The y-axis displays the corresponding topic labels.

(A) Contribution from Differences in Representation.



p-value from joint test of all coefficients equal to zero: < 0.0001

(B) Contribution from Differences in Pricing.



p-value from joint test of all coefficients equal to zero: 0.0015

Figure H.20: Sources of Disagreement: Alternative Order of Valuation Methods. This figure shows the decomposition of differences in price targets across analysts evaluating the same firm at the same time into the components specified in equation (22), reflecting differences in analyst attention to topic categories (Panel A) or in how they price different attributes (Panel B). Specifically, we estimate:

$$\mathbb{E}_t^i[\hat{p}_{f,t+1}] - \mathbb{E}_t^{i'}[\hat{p}_{f,t+1}] = \sum_{k=1}^{K} \left(\lambda_k^+ \Delta\alpha_{kft}^{i,+} + \lambda_k^- \Delta\alpha_{kft}^{i,-} + \Delta\lambda_k^+ \alpha_k^{i',+} + \Delta\lambda_k^- \alpha_k^{i',-} \right) + \epsilon_{f,t}$$

with $\alpha_{kft}^{i,+}$ ($\alpha_{kft}^{i,-}$) denoting the fraction of analyst i's report on firm f in year t allocated to topic k when it is discussed with positive (negative) sentiment, and with $\Delta a_{kft}^{i,+}$ ($\Delta a_{kft}^{i,-}$) capturing the difference in the fraction of analyst i's and analyst i''s report on firm f in year t allocated to topic k with positive (negative) sentiment. We run this regression on analyst pairs where one analyst bases their price target on a DCF-based valuation (without any mention of multiples) and the other on a multiples-based valuation (without any mention of DCF) We systematically sort the pairs as (i =DCF, i' =multiples), so that the difference in valuation weights captures the effect of multiples- relative to DCF-based methods. $\alpha_{kft}^{i,+}$, $\alpha_{kft}^{i,-}$, $\alpha_{kft}^{i',+}$, and $\alpha_{kft}^{i',-}$ are scaled by their in-sample standard deviation to enable direct comparisons across topics and sentiment directions. In Panel A, dots represent the average estimates of λ_k^+ and λ_k^-, respectively. In Panel B, dots represent the average differences in estimates of $\lambda_k^+ - \lambda_k'^+$ and $\lambda_k^- - \lambda_k'^-$, respectively. The "undetermined" and "valuation" categories are excluded from the exercise.

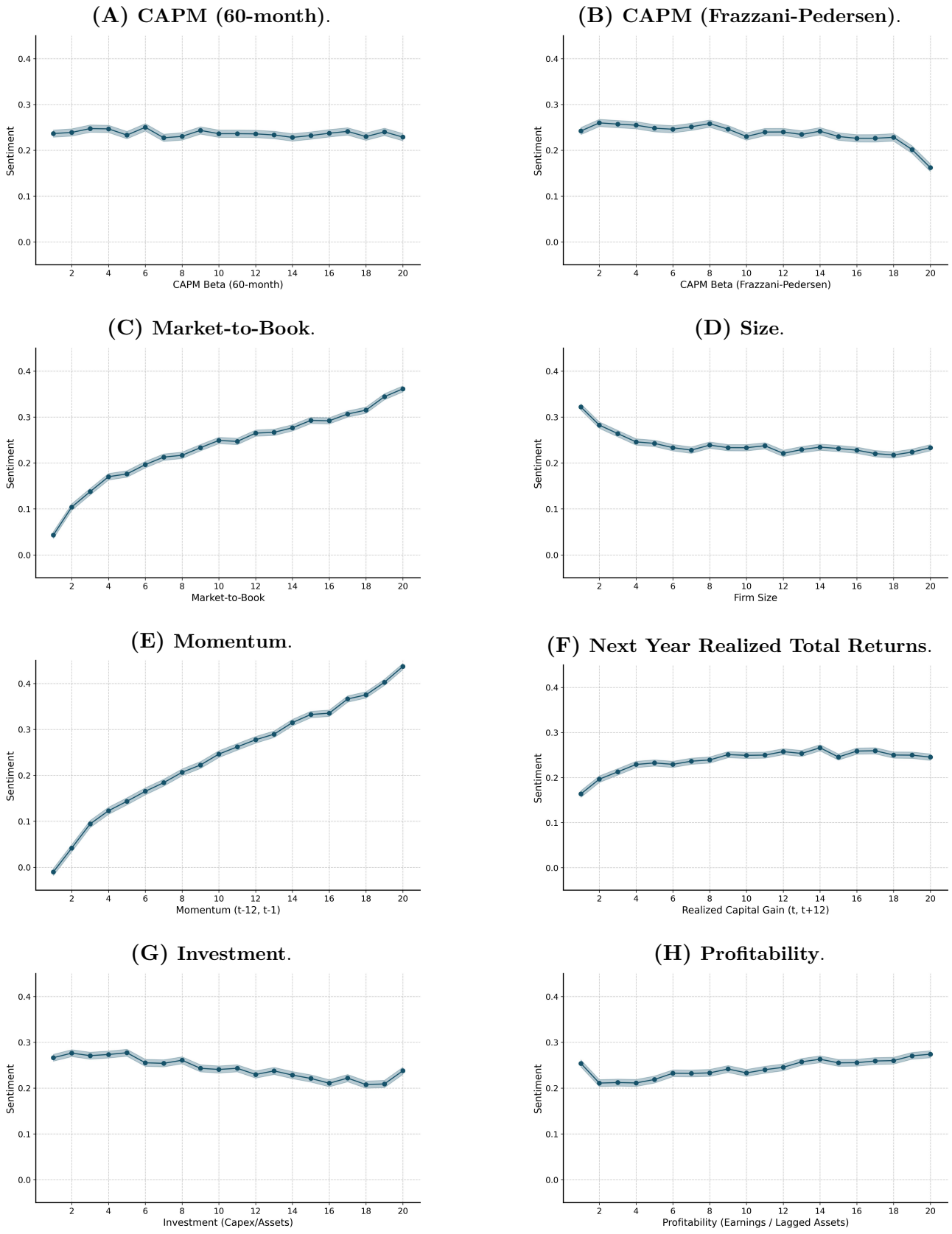


Figure H.21:
Sentiment in the Cross-Section

This figure plots the cross-sectional patterns of sentiment for various key variables: CAPM beta, market-to-book, firm size, momentum, next year realized capital gain, investment, and profitability across all firms in our sample with data on these characteristics. The x-axis corresponds to percentile brackets in 5% increments. In all panels, the y-axis represents sentiment of the average report within each percentile bracket. Report sentiment is calculated as the average sentiment of all topics discussed in the report, with -1 for negative sentiment, 0 for neutral sentiment, and $+1$ for positive sentiment.

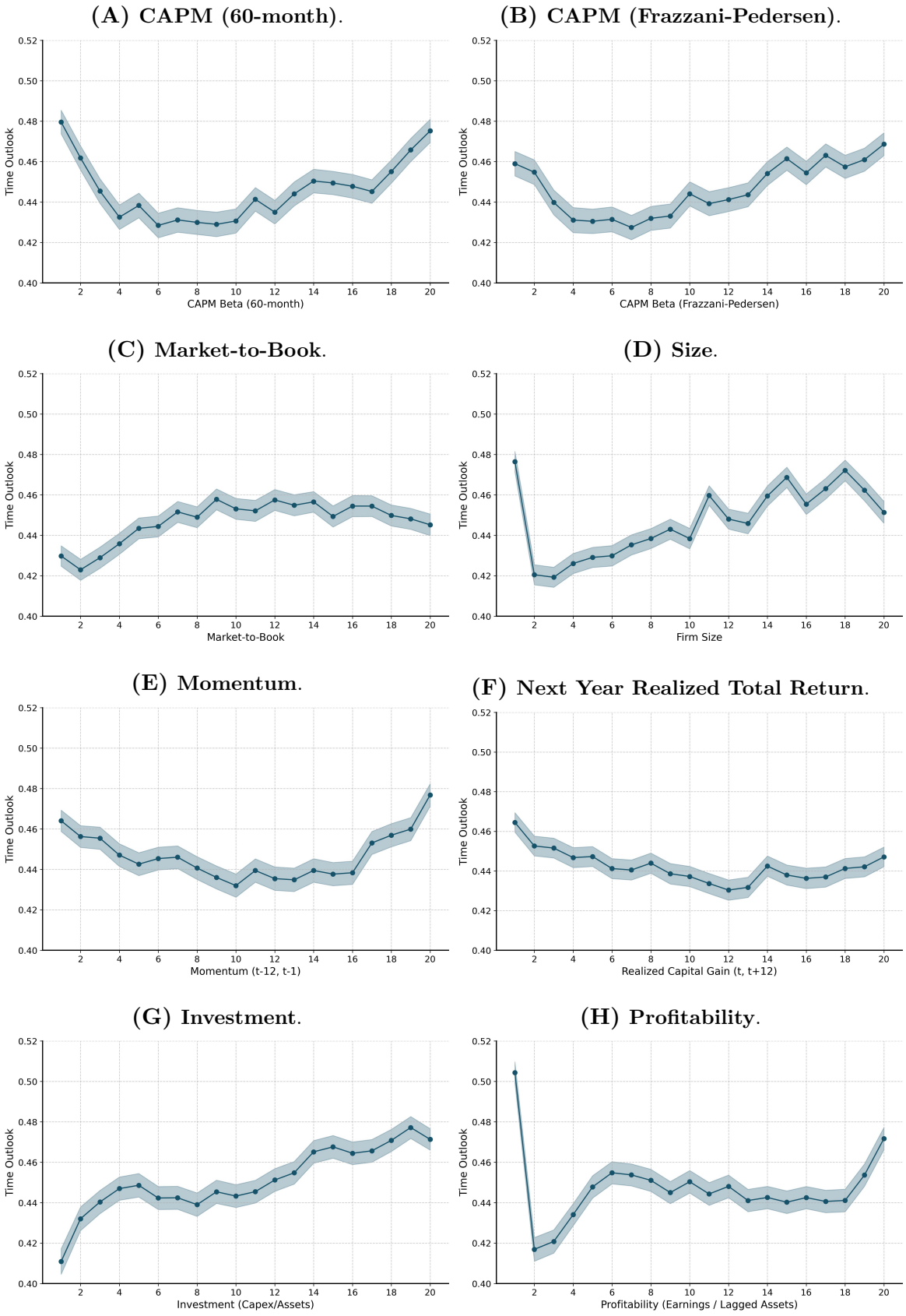


Figure H.22:
Time Outlook in the Cross-Section

This figure plots the cross-sectional patterns of time outlook for various key variables: CAPM beta, market-to-book, firm size, momentum, next year realized capital gain, investment, and profitability across all firms in our sample with data on these characteristics. The x-axis corresponds to percentile brackets in 5% increments. In all panels, the y-axis represents the the time outlook of the average report in each percentile bracket. Report time outlook is calculated as the average outlook of all topics discussed in the report, with -1 for past outlook, 0 for present outlook, $+1$ for near-future outlook, and $+2$ for distant-future outlook.

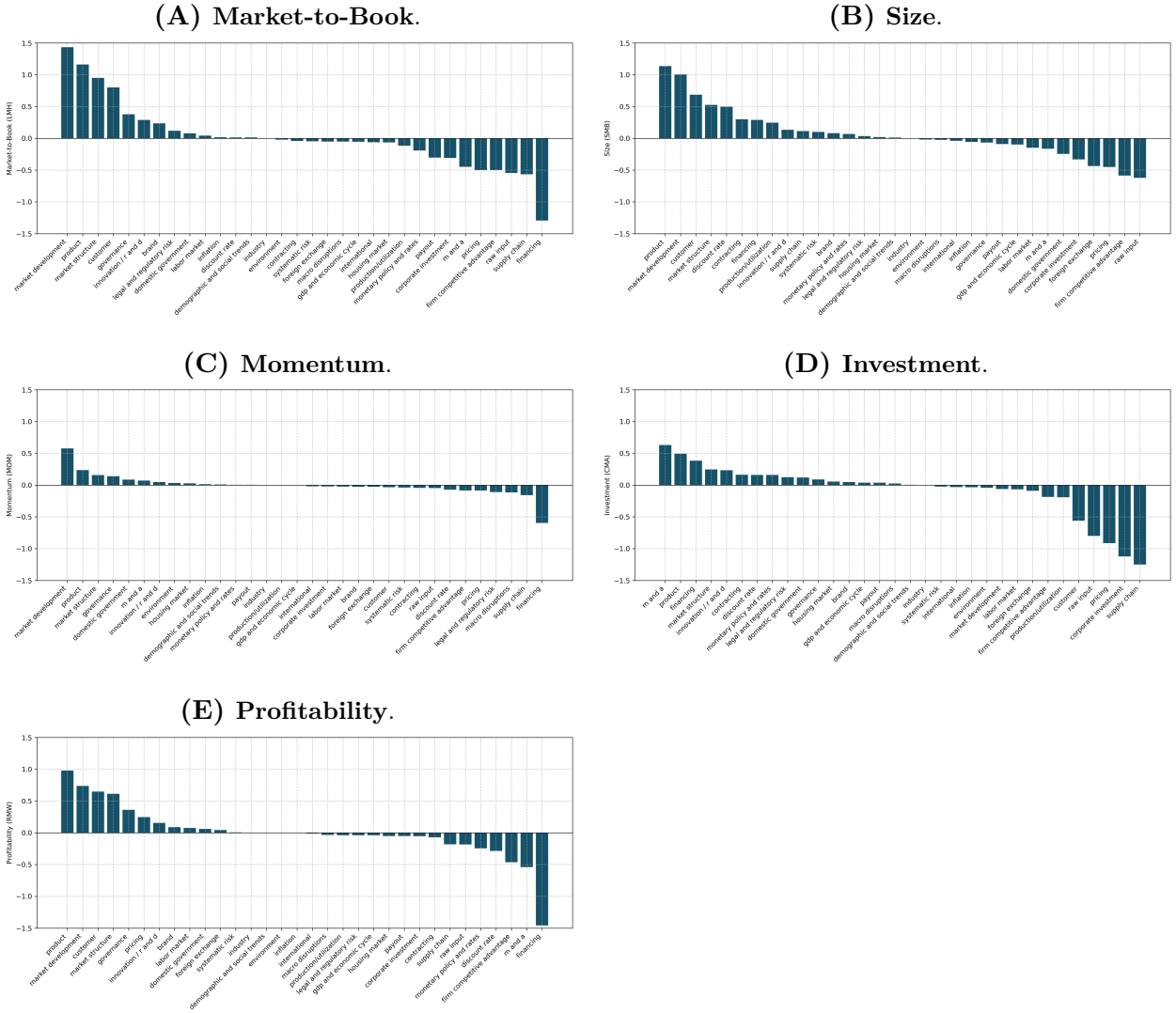


Figure H.23:
Topic Focus in the Cross-Section

This figure plots the cross-sectional patterns of topic category attention for various asset pricing sorts: market-to-book, firm size, momentum, investment, and profitability. The x-axis lists the categories. In all panels, the y-axis shows the average difference in how much a given category is discussed between the long and short ends of the corresponding asset pricing sort, estimated with industry-by-year fixed effects. Positive values indicate greater discussion on the long end.

Table H.1:
Comprehensiveness of Extracted Arguments

This table presents results examining the comprehensiveness of extracted topics and arguments, focusing on cash-flow channels: sales, costs, earnings, and margins. The first panel reports the proportion of topics per report associated with one, two, or at least three cash-flow channels. The second panel repeats this exercise using topic categories. The third panel reports the proportion of topic per report associated with one, two, or three distinct sentiments. The fourth panel repeats this exercise using topic categories.

Number of Valuation Channels per Topic	**Share of Sample**
1	77%
2	10%
3+	13%
Number of Valuation Channels per Topic Category	**Share of Sample**
1	65%
2	13%
3+	23%
Number of Sentiments per Topic	**Share of Sample**
1	76%
2	20%
3	4%
Number of Sentiments per Topic Category	**Share of Sample**
1	67%
2	22%
3	11%

Table H.2:
Informativeness of Extracted Information by Report Location – Sentiment

This table presents results examining the informativeness of extracted information for analysts' 12-month price target forecasts across different sections of the report. $Sentiment^{Q}_{i,j,t}$ denotes the sentiment associated with the portion of the report falling within a given quintile of report length. Coefficients indicate the effect of a one standard deviation change in the independent variable. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedasticity-consistent estimators clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Expected Capital Gain ($\ln(\frac{E_{i,j,t}[P_{j,t+1}]}{P_{i,t}})$)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\text{Sentiment}^{Q1-beginning}_{i,j,t}$	0.06***					0.04***	0.04***	0.04***	0.04***
	(0.00)					(0.00)	(0.00)	(0.00)	(0.00)
$\text{Sentiment}^{Q2}_{i,j,t}$		0.05***				0.03***	0.03***	0.03***	0.03***
		(0.00)				(0.00)	(0.00)	(0.00)	(0.00)
$\text{Sentiment}^{Q3}_{i,j,t}$			0.04***				0.02***	0.02***	0.02***
			(0.00)				(0.00)	(0.00)	(0.00)
$\text{Sentiment}^{Q4}_{i,j,t}$				0.03***				0.02***	0.02***
				(0.00)				(0.00)	(0.00)
$\text{Sentiment}^{Q5-end}_{i,j,t}$					0.02***				0.01***
					(0.00)				(0.00)
Firm*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	248,828	247,145	242,244	234,456	225,325	244,640	236,234	222,921	205,313
F Statistics	9198.117	7619.013	5955.035	4849.901	1526.691	5654.906	4016.012	3111.253	2421.346
R^2	0.585	0.573	0.566	0.559	0.545	0.601	0.611	0.618	0.623
Within R^2	0.109	0.083	0.065	0.049	0.013	0.143	0.162	0.176	0.181

Table H.3:
Time Outlook, Sentiment, and 12-Month Price Target Forecasts

This table presents results examining the role of sentiment and time outlook in shaping analysts' 12-month price target forecasts. The analysis uses our main sample spanning 2000–2025. The dependent variable is the natural logarithm of analyst i's price target forecast for firm j in year t normalized by the firm's current price as of the time of the equity report. $Outlook_{i,j,t}$ denotes the average time outlook of analyst i for firm j in year t. $Sentiment_{i,j,t}$ represents the average sentiment of analyst i for firm j in year t. Positive values for Sentiment and Outlook indicate positive sentiment and forward-looking outlook, respectively. $Sentiment_{i,j,t}^{Distant-future}$, $Sentiment_{i,j,t}^{Near-future}$, $Sentiment_{i,j,t}^{Present}$, and $Sentiment_{i,j,t}^{Past}$ are calculated using only the statements classified as referring to the distant future, near future, present, and past, respectively. Coefficients in Columns (4) to (7) are standardized to reflect one standard deviation changes in the independent variables, allowing for easier comparison across effects. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedasticity-consistent estimators clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Expected Capital Gain ($\ln(\frac{E_{i,j,t}[P_{j,t+1}]}{P_{j,t}})$)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Outlook}_{i,j,t}$	0.04***		0.01***				
	(0.00)		(0.00)				
$\text{Sentiment}_{i,j,t}$		0.15***	0.18***				
		(0.00)	(0.00)				
$\text{Sentiment}_{i,j,t}^{Past}$				0.02***			
				(0.00)			
$\text{Sentiment}_{i,j,t}^{Present}$					0.06***		
					(0.00)		
$\text{Sentiment}_{i,j,t}^{Near-future}$						0.06***	
						(0.00)	
$\text{Sentiment}_{i,j,t}^{Distant-future}$							0.03***
							(0.00)
Firm*Year FE	No	No	Yes	Yes	Yes	Yes	Yes
Observations	301,364	301,364	251,553	204,904	248,946	251,196	113,493
F Statistics	519.52	11262.84	6721.59	1905.90	9724.38	9708.42	1972.61
R^2	0.00	0.10	0.61	0.55	0.59	0.59	0.59

Table H.4:
Valuation Methods and Topic Alignment

This table presents results examining the role of valuation methods in explaining topic and argument alignment. The analysis uses our main sample spanning 2000–2025, creating pairs of analysts evaluating the same firm in the same year. In Columns 1 and 2, the dependent variable, *Topic Overlap*$_{A,B,j,t}$, captures the number of overlapping topics between analyst A and B's reports divided by the total number of distinct topics mentioned in either report, multiplied by 100. In Columns 3 and 4, the dependent variable, *Argument Overlap*$_{A,B,j,t}$, is the share of overlapping arguments, i.e., topics for which both analysts assign the same time outlook, sentiment, and valuation channel, multiplied by 100. *Same Valuation Method*$_{A,B,j,t}$ is an indicator variable equal to one if both analysts in the pair align in their valuation method, i.e., align whether they use a DCF-based price target and whether they use a multiples-based one. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedasticity-consistent estimators clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Topic Overlap$_{A,B,j,t}$		Argument Overlap$_{A,B,j,t}$	
	(1)	(2)	(3)	(4)
Same Valuation Method (%)$_{A,B,j,t}$	2.68***	1.61***	0.68***	0.36***
	(0.06)	(0.05)	(0.05)	(0.05)
Firm*Year FE	No	Yes	No	Yes
Analyst-A*Firm FE	No	Yes	No	Yes
Analyst-B*Firm FE	No	Yes	No	Yes
Observations	514,784	466,481	514,458	466,137
F Statistics	2162.25	1080.57	217.96	45.03
R^2	0.02	0.52	0.00	0.38

Table H.5:
Mental Model Alignment Across Analysts – Outlook, Valuation Channels, and Sentiment

This table presents results relating sentiment, time outlook, and valuation channels. The analysis is conducted at the analyst pair (A, B), topic k, firm j and year t level. The dependent variable, *Same Sentiment*, equals one if both analysts share the same sentiment for the overlapping topic k and zero otherwise. *Same Outlook* equals one if both analysts share the same outlook to topic k and zero otherwise. *Same Valuation Channel* equals one if both analysts assign the same valuation channel to topic k and zero otherwise. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Same Sentiment$_{A,B,k,j,t}$			
	(1)	(2)	(3)	(4)
Same Outlook$_{A,B,k,j,t}$	0.05***		0.04***	0.04***
	(0.00)		(0.00)	(0.00)
Same Valuation Channel$_{A,B,k,j,t}$		0.05***	0.05***	0.04***
		(0.00)	(0.00)	(0.00)
Topic*Firm*Analyst-Pair FE	No	No	No	Yes
Topic*Firm*Year FE	No	No	No	Yes
Observations	4,399,837	4,399,837	4,399,837	1,580,527
F Statistics	3210.48	2494.32	2982.91	707.96
R^2	0.00	0.00	0.00	0.62

Table H.6:
Semantic Similarity

This table provides an example of the semantic similarity between the topic "Antitrust" and the other topics included in the analysis. A higher value of semantic similarity indicates a greater likelihood that analysts are referring to a similar underlying concept when discussing the two topics in their analysis. We measure semantic similarity in two steps. First, for each topic used in the analysis, we generate the associated embedding vector using OpenAI embedding API "text-embedding-ada-002." In a second step, we measure the Cosine similarity for each pair of topics.

Topic	**Semantic similarity with "Antitrust"**
Collusion	86%
Oligopoly	85%
Barriers to entry	82%
Vertical integration	82%
[...]	
Market share	81%
Partnership / Joint-venture	81%
Mergers and acquisitions (M&A)	80%
Intellectual property (ip)	80%
[...]	
Home prices	78%
Monetary policy	78%
Income inequality	78%
Corporate social responsibility (csr)	78%
[...]	
Capital expenditure (capex)	74%
Discount rates / WACC / Cost of capital	73%
Debt level / Leverage	72%
Risk premium / Systematic risk	72%

Table H.7:
Analyst Proximity and Forecast Disagreement

This table presents results examining the relation between analysts' country of location and topic attention alignment as well as forecast disagreement. The analysis uses our main sample spanning 2000–2025, creating pairs of analysts evaluating the same firm in the same year. In Columns 1 and 2, the dependent variable, *Topic Overlap*$_{A,B,j,t}$, captures the number of overlapping topics between analyst A and B's reports divided by the total number of distinct topics mentioned in either report, multiplied by 100. In Columns 3 and 4, the dependent variable, *Argument Overlap*$_{A,B,j,t}$, is the share of overlapping arguments, i.e., topics for which both analysts assign the same time outlook, sentiment, and valuation channel, multiplied by 100. In Columns 5 and 6, the dependent variable, *Forecast Disagreement*$_{A,B,j,t}$, is defined as the absolute difference between forecaster A and B's 12-month price targets normalized by the firm's current price as of the time of the equity reports, multiplied by 100. The key independent variable, $1^{\textit{Analysts Located in the Same Country}}_{A,B,j,t}$, is an indicator variable equal to one if both analysts are located in the same country and zero otherwise. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Topic Overlap$_{A,B,j,t}$		Argument Overlap$_{A,B,j,t}$		Forecast Disagreement$_{A,B,j,t}$	
	(1)	(2)	(3)	(4)	(5)	(6)
$1^{\text{Analysts Located in the Same Country}}_{A,B,j,t}$	1.64***	1.04***	0.57***	0.32***	-1.30***	-0.44***
	(0.09)	(0.07)	(0.07)	(0.06)	(0.15)	(0.11)
Firm*Year FE	No	Yes	No	Yes	No	Yes
Observations	514,784	492,345	514,458	492,028	514,784	492,345
F Statistics	318.37	220.89	66.88	24.53	73.84	16.58
R^2	0.00	0.27	0.00	0.17	0.00	0.33

Table H.8:
Explaining Valuation Methods

This table shows the explanatory power of year fixed effects, contributor fixed effects, firm fixed effects, analyst fixed effects, for the usage of DCF- versus multiples-based valuation methods. The analysis uses universe of equity reports from 2000–2025 with information on valuation methods, price target, and analysts, focusing on reports that use a DCF-based or multiples-based price-target valuation method. *DCF Usage Only*$_{i,j,t}$, is an indicator variable that equals one if the analyst uses a DCF-based valuation method without any mention of a multiples-based approach, and zero otherwise. *Multiple Usage Only*$_{i,j,t}$, is an indicator variable that equals one if the analyst uses a multiples-based valuation method without any mention of a DCF-based approach, and zero otherwise. Panels A and B focus on explaining DCF usage, with Panel B showing the incremental R^2 from adding analyst fixed effects. Panels C and D focus on explaining multiple usage, with Panel D showing the incremental R^2 from adding analyst fixed effects. The regressions are estimated using ordinary least squares.

Panel A: DCF				
Dep. Var.: 1(*DCF Usage Only*)	R^2	N	R^2	N
	Full Sample		$\geq$ 100 Obs. per Firm and Analyst	
Year FE	0.010	1,170,653	0.008	591,823
Contributor FE	0.070	1,170,653	0.075	591,822
Firm FE	0.366	1,170,653	0.305	591,823
Analyst FE	0.575	1,170,653	0.619	591,823
Panel B: DCF – Incremental Analyst R^2				
Dep. Var.: 1(*DCF Usage Only*)	R^2	N	R^2	N
	Full Sample		$\geq$ 100 Obs. per Firm and Analyst	
Year + Contributor + Firm FE	0.389	1,170,653	0.337	591,822
+ Analyst FE	0.653	1,170,653	0.663	591,822
Panel C: Multiples				
Dep. Var.: 1(*Multiple Usage Only*)	R^2	N	R^2	N
	Full Sample		$\geq$ 100 Obs. per Firm and Analyst	
Year FE	0.012	1,170,653	0.009	591,823
Contributor FE	0.090	1,170,653	0.088	591,822
Firm FE	0.423	1,170,653	0.351	591,823
Analyst FE	0.629	1,170,653	0.646	591,823
Panel D: Multiples – Incremental Analyst R^2				
Dep. Var.: 1(*Multiple Usage Only*)	R^2	N	R^2	N
	Full Sample		$\geq$ 100 Obs. per Firm and Analyst	
Year + Contributor + Firm FE	0.445	1,170,653	0.383	591,822
+ Analyst FE	0.695	1,170,653	0.686	591,822

Table H.9:
Robustness to Alternative Topic Classifications

This table shows robustness to alternative topic classifications, grouping topics that are semantically sufficiently close based on either a 85% or 80% similarity cutoff. See Tables VIII and H.4 for additional details. See Table H.6 for details and an example on the semantic similarity calculation. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Topic Overlap$_{A,B,j,t}$		Forecast Disagreement$_{A,B,j,t}$			
	85% Smntc. Sim.	80% Smntc. Sim.	Full Sample		Same Val. Method	
	(1)	(2)	(3)	(4)	(5)	(6)
Same Valuation Method$_{A,B,j,t}$	0.02***	0.01***				
	(0.00)	(0.00)				
Topic Overlap$^{85\%SemanticSim.}_{A,B,j,t}$			-0.02***		-0.02***	
			(0.00)		(0.00)	
Topic Overlap$^{80\%SemanticSim.}_{A,B,j,t}$				-0.01***		-0.01**
				(0.00)		(0.01)
Argument Overlap$_{A,B,j,t}$			-0.03***	-0.03***	-0.03***	-0.03***
			(0.00)	(0.00)	(0.00)	(0.00)
Firm*Analyst-Pair FE	No	No	Yes	Yes	Yes	Yes
Firm*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	492,345	492,345	311,341	311,341	123,287	123,287
F Statistics	1040.66	742.33	131.85	109.13	58.49	46.80
R^2	0.30	0.30	0.67	0.67	0.71	0.71

Table H.10:
Valuation Methods, Topic Attention, and Forecast Disagreement – Robustness

This table shows robustness of the results in Table VIII, estimating the specifications on the subsample of analysts who have consistently used the current valuation method when valuing the firm. See Table VIII for additional details. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	Forecast Disagreement$_{A,B,j,t}$			
	(1)	(2)	(3)	(4)
Same Valuation Method$_{A,B,j,t}$	-0.40***	-0.31***		
	(0.07)	(0.07)		
Topic Overlap$_{A,B,j,t}$			-0.06***	-0.05***
			(0.00)	(0.00)
Argument Overlap$_{A,B,j,t}$			-0.05***	-0.04***
			(0.00)	(0.00)
Firm*Year FE	Yes	Yes	Yes	Yes
Analyst-A*Firm FE	No	Yes	No	Yes
Analyst-B*Firm FE	No	Yes	No	Yes
Observations	250,084	248,912	249,917	248,745
F Statistics	30.62	18.30	351.82	329.81
R^2	0.38	0.44	0.38	0.45

Table H.11:
Mental Model Rigidity in the Cross Section – Heterogeneity

This table presents heterogeneity in cross-sectional mental model rigidity by whether firms are located in same country or industry (Panel A) and by industry distance (Panel B). See Table VII for additional details. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Heterogeneity (i)

Dependent variable:	Same Valuation Method		Topic Overlap	
	(1)	(2)	(3)	(4)
Same Country	0.06***		0.02***	
	(0.01)		(0.00)	
Same Industry		0.05***		0.03***
		(0.00)		(0.00)
LHS Mean	0.72	0.72	0.30	0.30
Analyst*Firm$_A$ FE	Yes	Yes	Yes	Yes
Analyst*Firm$_B$ FE	Yes	Yes	Yes	Yes
Observations	1,438,656	1,439,188	1,438,656	1,439,188
F Statistics	85.04	314.79	141.74	950.76
R^2	0.50	0.50	0.39	0.40
Within R^2	0.00	0.00	0.00	0.01

Panel B: Heterogeneity (ii)

Dependent variable:	Same Val.	Top. Ovlp.	Same Val.	Top. Ovlp.	Same Val.	Top. Ovlp.	Same Val.	Top. Ovlp.
	Far Industries				Similar Industries			
	Mining v. Entertnmnt		Admin v. Agricltre		Prof. Srvcs v. Information		Wholesale v. Retail	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same Industry	0.26***	0.14***	0.63***	0.11***	0.01***	0.01***	-0.00	0.02***
	(0.03)	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
LHS Mean	0.69	0.31	0.80	0.32	0.72	0.31	0.83	0.32
Analyst*Firm$_A$ FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Analyst*Firm$_B$ FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66,088	66,088	4,157	4,157	161,769	161,769	59,985	59,985
F Statistics	84.33	145.35	1208.73	132.99	13.39	23.17	0.34	33.65
R^2	0.47	0.45	0.54	0.42	0.52	0.41	0.49	0.35
Within R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table H.12:
Mental Model Rigidity in the Cross Section – Technical Topics

This table presents results on cross-sectional mental model rigidity estimated from technical topics only. estimated using technical topics only. These include specialized subjects reflecting technical reasoning, such as covenants, inventory management, regulatory environment, and contract terms, with the full list available upon request. See Table VII for additional details. The regressions are estimated using ordinary least squares. Standard errors (in parentheses) are estimated using heteroskedastic-consistent estimators and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable:	All Topics	Technical Topics	
		Incl. No Overlap	Positive Overlap
	(1)	(2)	(3)
Same-Analyst Pair	0.04***	0.05***	0.04***
	(0.00)	(0.00)	(0.00)
LHS Mean | Same-Analyst Pair	0.30	0.30	0.33
Analyst*Firm-Pair FE	Yes	Yes	Yes
Observations	12,807,012	12,807,012	7,033,316
F Statistics	13324.16	7902.07	6415.29
R^2	0.28	0.33	0.25
Within R^2	0.02	0.01	0.01