From Numbers to Words: Breaking Down Institutional Beliefs*

Andrea Andolfatto† *Federico Bastianello*‡

[Link to the latest version]

June 20, 2026

Abstract

We examine how large asset managers form and justify long-horizon beliefs using their Capital Market Assumptions (CMAs). Our evidence points to a common architecture: managers decompose equity return expectations into shared building blocks, populate them using heterogeneous modeling assumptions, and process information through causal narratives, while peer consensus anchors forecast revisions. Valuation change and growth explain 77% of cross-sectional dispersion and are most strongly linked to equity allocations. Valuation-change expectations are countercyclical, whereas growth expectations are procyclical, generating countercyclical return expectations overall, with substantial heterogeneity across managers. Disclosed modeling assumptions matter too: mean-reversion and historical calibration predict systematic deviations from peer consensus. Using a new LLM methodology, we extract directed, signed causal networks from CMA narratives. Greater network complexity and attention to valuation change are associated with underreaction to positive earnings news, whereas attention to dividend yield and downturns is associated with overreaction. Comparisons with N-CSR shareholder letters show that CMA narratives reflect persistent institution-specific investment views. Volatility and correlation forecasts, by contrast, vary less across managers and remain closely tied to historical realizations.

*We thank Svetlana Bryzgalova, Leland Bybee, John Y. Campbell, Carlo Favero, Nicola Gennaioli, Katrin Gödker, Francisco Gomes, Sam Hanson, Spencer Kwon (discussant), Max Miller, Florian Nagler, Anna Pavlova, Marco Sammon, Andrei Shleifer, Adi Sunderam, Jakob Ahm Sørensen, and Claudio Tebaldi for thoughtful comments and discussions. We also thank the participants of the Harvard Behavioral Reading Group (2025), the Bocconi Workshop on Machine Learning and Financial Decision Making (2026), the Bocconi Ph.D. Workshop in Behavioral Economics and Finance (2026), and FIRS (2026).

†Bocconi University: andrea.andolfatto@phd.unibocconi.it

‡London Business School: fbastianello@london.edu

How do investors form beliefs, and what drives heterogeneity in investor expectations? A growing literature studies how investors form, communicate, and act on expectations along three complementary lines. First, surveys of self-reported beliefs document systematic biases—such as extrapolation and both over- and underreaction—that challenge rational expectations and motivate behavioral models of belief formation (Greenwood & Shleifer 2014, Bouchaud et al. 2019, Bordalo et al. 2019, 2024*b*, Bastianello 2025). Second, recent work links self-reported beliefs to portfolio choices and trading, showing that pass-through of beliefs to actions varies across investor types (Giglio et al. 2021, Dahlquist & Ibert 2024, Andonov et al. 2026). Third, text-based approaches use investor-authored narratives and news to extract expectations and the underlying economic channels that shape them (Bybee 2023, Décaire & Graham 2024, Bastianello et al. 2024, Sarkar 2025).

We draw on all three tools—numeric expectations, portfolio data, and text—as well as news-based measures of information flow to analyze how large institutional asset managers form and justify the long-horizon expectations for equity returns, volatilities, and correlations reported in their Capital Market Assumptions (CMAs) (Couts et al. 2023, 2024, Dahlquist & Ibert 2024, 2025, Begenau et al. 2025).[1] We "read" their CMAs along two complementary dimensions: in numbers (from tables and figures with forecasts and their decompositions, labeled as "building blocks") and in words (assumptions and narratives that explain the underlying mechanisms), linking beliefs to portfolio allocations, the arrival of news, and managers' shareholder-letter outlooks in N-CSR filings.

Our research question is straightforward: How does the joint analysis of CMA forecasts, their building blocks, their underlying assumptions, and the narratives in CMA documents reveal the mechanisms that drive cross-sectional and time-series variation in asset managers' beliefs? Do these beliefs matter for investment decisions, and do they exhibit systematic biases?

Previewing our findings, we show that combining granular data on the building blocks of

[1]CMAs are not generic surveys; they are inputs that anchor strategic asset allocation and risk budgeting—indeed, leading institutions describe their CMA framework as "a critical foundation" of multi-asset processes, making them a uniquely informative window into professional belief formation. For example, J.P. Morgan's 2022 Long-Term Capital Market Assumptions (LTCMAs) describes its CMA methodology as a cornerstone of its strategic asset-allocation process.

CMA forecasts with the assumptions used to construct those blocks and with managers' narratives reveals how asset managers build, update, and rationalize long-horizon beliefs. On the quantitative side, cross-sectional variation in return expectations is explained by four main building blocks: expected changes in valuation (45%), expected real growth (32%), dividend yield (21%), and inflation (3%). Over time, valuation expectations move countercyclically, whereas growth expectations are procyclical. We further show that the relationships managers embed among these building blocks differ from those observed in historical data: real growth and valuation change are negatively correlated within managers' subjective decompositions but positively correlated in realized data, while inflation and valuation change display the opposite sign pattern. This motivates our investigation of systematic deviations in asset managers' forecasts relative to objective machine-learning-based return forecasts. The machine-learning benchmark systematically outperforms CMA forecasts, and the comparison suggests that managers place relatively more weight on macro-financial variables, especially inflation, and less weight on valuation ratios than the objective benchmark does. Forecast dynamics also point to a consensus anchor: managers revise toward the cross-sectional consensus and deviate from the benchmark in the direction of the recent CMA consensus. By showing that the growth and valuation building blocks systematically map into portfolio allocations, we highlight the drivers of the pass-through from beliefs to actions. By contrast, second-moment expectations are history-anchored: volatility forecasts cluster around long-run averages and adjust only weakly to realized volatility, and pairwise correlation assumptions for equity-equity and equity-bond pairs exhibit limited disagreement, small year-to-year revisions, and close alignment with realized historical correlations.

We then show that disagreement across asset managers is not driven only by different numerical inputs, but also by the assumptions used to construct those inputs. We measure these assumptions by extracting modeling descriptions from CMA equity sections, embedding and clustering them into recurring forecasting approaches. The resulting clusters reveal that historical calibration is associated with more optimistic forecasts, while mean reversion is associated with lower forecasts and more negative valuation-change assumptions, both relative to peers and to the objective benchmark. Thus, disagreement reflects not only different reported building blocks, but also different modeling anchors behind them.

We then turn to managers' broader narratives. We extract causal networks from CMA text to measure heterogeneity in asset managers' explanations, capturing both the complexity of their reasoning and their attention to different economic topics. These narrative structures help explain forecast levels, belief updating, and systematic deviations from objective benchmarks. In particular, ex-ante forecast errors are predictably related to causal-network complexity and topic attention in CMA narratives. We also validate the economic content of CMA narratives by comparing them with shareholder letters filed in N-CSR reports. Same-institution letter–CMA pairs are systematically more similar than cross-institution pairs, indicating that the narratives embedded in CMAs also appear in affiliated fund communications. This supports the interpretation that CMAs capture persistent institution-specific investment views.

Taken together, these results provide a unifying view of how asset managers form long-horizon equity-return expectations, summarized in Figure 1. The evidence points to a common building-block architecture, but not to a common forecasting rule. Asset managers decompose expected returns into valuation change, real growth, inflation, and income, while differing in how they construct and update these components. Disclosed modeling assumptions and calibration methods are associated with the levels of the individual building blocks. Causal narratives capture the topics and transmission channels through which managers organize economic information, while narrative attention and complexity condition how their forecasts respond to external earnings-news sentiment. Peer consensus provides an additional anchor for forecast revisions and benchmark-relative deviations. The final return forecast can therefore be interpreted as combining a common accounting structure with heterogeneous modeling assumptions, narrative representations, and degrees of consensus anchoring. Volatility and correlation beliefs provide a contrasting case, as they are considerably more stable and predominantly anchored to historical estimates.

[Figure 1 here]

We begin by constructing a novel dataset by extracting structured text and return decompositions from Capital Market Assumptions (CMA) reports using natural-language-processing (NLP) and text-parsing techniques. We show that these decompositions can be systematically interpreted through the lens of an approximate accounting identity, revealing that asset managers share mental models grounded in fundamental drivers such as income (dividends and net buybacks), real growth, inflation, and expected valuation change.

We use granular data on the building blocks provided by asset managers to quantify the components of cross-sectional variation in return expectations and find that four main components dominate, with expected valuation change explaining 45% and expected growth in fundamentals another 32%. This component-level variance decomposition is new relative to Dahlquist & Ibert (2025). While they identify valuation change as a source of heterogeneity using external proxies for the building blocks, our direct use of CMA disclosures quantifies the relative importance of each component and highlights that growth, in addition to valuation change, contributes substantially to the cross-sectional variation in return expectations.

We then interpret the building blocks through a Campbell–Shiller lens at the asset manager level. Our contribution is to show explicitly the sources of the countercyclicality in CMA total return expectations (Dahlquist & Ibert 2024). We show that two components primarily drive the cyclical properties of asset managers' return expectations: growth expectations are procyclical, whereas expected valuation changes are countercyclical and, on average, dominate, resulting in overall countercyclical return expectations. We also use the building-block representation to compare the subjective relationships among return components with their objective counterparts. In managers' subjective decompositions, real growth and valuation change are negatively correlated, whereas the corresponding realized correlation is positive; inflation and valuation change display the reverse sign pattern. These differences motivate our subsequent benchmark-based analysis of systematic deviations in CMA forecasts.

We next construct an ex-ante objective benchmark using an equal-weighted ensemble of regularized machine-learning return-forecasting models. The benchmark forecasts systematically outperform CMA forecasts. A feature-importance comparison shows that the variables emphasized by managers differ from those most informative for realized returns: subjective forecasts place relatively more weight on macro-financial predictors, especially inflation, and less weight on valuation ratios such as the earnings–price and dividend–price ratios.

We then study forecast dynamics. Managers whose prior forecasts are far from the leave-one-out consensus revise toward that consensus, even after controlling for contemporaneous changes in peer forecasts and lagged realized fundamentals. In addition, forecast deviations from the machine-learning benchmark move in the direction of the recent CMA consensus. Together, these results suggest that the consensus is an economically meaningful anchor for long-horizon institutional beliefs.

We also examine how asset managers incorporate incoming textual news into long-horizon expectations. Using text-based earnings sentiment, we show that return expectations and valuation-change expectations increase following positive earnings news.

We show the relevance of the building blocks by relating them to the equity allocations of asset managers, which load more strongly on the growth and valuation-change components than on the income/dividend component. This reveals that the components that account for most cross-institution disagreement are also those most closely associated with portfolio allocations. In matched fund-level regressions, the coefficients on nominal growth and valuation change are both significant and economically meaningful, with the coefficient on expected growth nearly twice as large as that on expected valuation change and larger than the univariate coefficient on total return expectations. Taken together, these results indicate that CMA expectations matter for asset allocations and show which components of return expectations are most closely associated with portfolio weights.

Having established the properties of return expectations, we then turn to beliefs about volatility. We find that long-horizon volatility expectations are highly conservative and display limited responsiveness to market conditions. Manager identity explains the vast majority of cross-sectional variation in volatility forecasts, indicating that institutions rely on stable, idiosyncratic frameworks for assessing volatility. Reported forecasts cluster tightly around 15–17% per year and adjust only minimally to large swings in realized volatility. Consistent with this anchoring, an EWMA analysis implies that managers place disproportionate weight on very long histories, with an estimated half-life of 36.82 quarters, or about 9.2 years. We find a similar pattern for equity-equity and equity-bond pairs correlation beliefs. Pairwise correlation assumptions exhibit limited disagreement across managers and small year-to-year revisions, suggesting that institutions also treat correlations as relatively stable second-moment inputs. Taken together, the evidence indicates that both volatility and correlation beliefs are more backward-looking and history-anchored than return expectations.

We complement the numerical evidence by "reading" CMAs in words. We examine the construction of the building blocks themselves. CMA reports often disclose whether assumptions are based on historical or long-run averages, mean-reversion arguments, trend assumptions, simulations, regressions, or other modeling inputs. Using both broad indicators of disclosed techniques and a more granular clustering of modeling phrases, we show that these implementation choices are systematically associated with where managers stand relative to peers and to the objective benchmark. In particular, mean-reversion approaches are associated with lower forecasts and more negative valuation-change assumptions, while historical calibration is associated with higher forecasts. This evidence shows that the assumptions used to populate otherwise similar building-block frameworks are a distinct source of disagreement across institutions.

Using topic modeling on equity-focused segments of each report, we show that the documents cluster into a small number of narrative groups, indicating systematic differences in how institutions frame the macroeconomic environment. We then develop a new empirical framework to extract causal networks from asset managers' explanations of their return expectations. Specifically, we use a large language model (LLM) pipeline to encode directed and signed links among macro variables and the building-block drivers of return expectations. Relative to existing applications of LLMs to financial text, our contribution is not only to classify narratives or summarize their content, but to reconstruct the causal structure through which asset managers link macroeconomic beliefs to return expectation components. This network representation allows us to study heterogeneity in the structure of belief formation, rather than only heterogeneity in topics, sentiment, or textual similarity. We document systematic heterogeneity in both causal-network complexity (path distance, indirect connection ratio, and network transitivity) and topic attention (the normalized number of causal connections to each topic node), with average coefficients of variation exceeding 0.7 and 1, respectively—indicating that asset managers differ substantially in both the richness of their reasoning and the focus of their narratives.

We finally use the text data to further identify the sources of heterogeneity and dynamics in institutional investors' return expectations. The structure and content of asset managers' causal narratives, quantified by our extracted causal networks, predict systematic and predictable deviations from objective benchmarks. Narrative complexity does not shift forecast levels directly, but persistent differences in complexity shape how managers process news: managers with structurally more complex causal networks underreact to positive fundamentals and produce more pessimistic forecasts relative to the benchmark. Topic-specific attention is an additional source of bias. Attention to valuation-change mechanisms is associated with underreaction to positive news, whereas attention to dividend-yield and downturn narratives is associated with forecasts that move more strongly with positive news relative to the benchmark. Overall, both complexity and topic emphasis in causal narratives help explain systematic forecast bias in asset managers' beliefs.

Literature review

A rapidly growing set of papers has pinned down several first-order facts about institutional CMA beliefs. First, subjective premia co-move closely with simple objective benchmarks and are countercyclical; disagreement across institutions is large and persistent, and different priors about long-run valuation levels play an important role (Dahlquist & Ibert 2024, 2025). Second, across assets there is a strong positive subjective risk–return trade-off; both the dispersion of beliefs and the strength of the trade-off are greater across asset classes than across institutions, and aggregated subjective returns predict realized returns largely through the risk-premia component (Couts et al. 2023, 2024).

We contribute to this literature by addressing our central question—how managers form and justify long-horizon expectations on returns and second moments, which are key inputs in asset-pricing and portfolio-allocation models. Four gaps matter. First, prior work reads CMAs "in numbers" and leaves the mechanism—how managers translate assumptions about growth, inflation, productivity, policy, and valuations into forecasts—implicit. We open that box in two steps: by showing that disclosed modeling assumptions are systematically associated with the building blocks of return expectations, and by extracting manager-authored causal networks from CMA narratives. Second, the literature emphasizes first moments, but has not examined whether, how, and why institutional beliefs depart systematically from objective benchmarks. We document systematic deviations from objective benchmarks and show that they are related to forecast dynamics, modeling anchors, and the structure of managers' narratives. Third, existing studies relate beliefs to state variables but not to directly measured news. We address this link empirically by combining text-based earnings-sentiment news with CMA forecasts and narratives, showing that managers update return and valuation-change expectations after earnings news and that this updating varies systematically with narrative complexity and topic attention. Finally, while prior work documents a positive pass-through of institutional beliefs to their actions (Dahlquist & Ibert 2024, Andonov et al. 2026), we further document that the source of variation in return expectations driving asset managers' equity allocations originates from their valuation-change and growth-expectation components.

Beyond the CMA literature, our paper speaks to several broader strands of research. First, it contributes to work documenting that beliefs of both retail and professional investors are heterogeneous and systematically biased (Malmendier & Nagel 2011, Greenwood & Shleifer 2014, Bouchaud et al. 2019, Andre et al. 2022, Nagel & Xu 2023, De la O & Myers 2021, 2024, Bordalo et al. 2024*a*,*b*, Bastianello 2024, 2025) by providing novel evidence on systematic deviations of asset managers' return expectations from objective benchmarks.[2] Second, it relates to the literature at the intersection of machine learning and finance, particularly work using textual information to study beliefs (Asquith et al. 2005, Baker et al. 2021, Bastianello et al. 2024, Bianchi et al. 2024, Bybee et al. 2024, Décaire & Graham 2024, Lopez-Lira 2024, van Binsbergen et al. 2024, Bastianello & Peng 2025, Bianchi et al. 2025, Laarits et al. 2025, Sarkar 2025), by studying the textual content of CMA narratives and validating their manager-specific content against N-CSR shareholder letters rather than relying only on newspapers or analyst reports. Third, it connects with the narrative economics literature, which highlights the role of stories in shaping expectations in financial markets: narratives propagate, frame causal reasoning, and influence asset prices (Shiller 2019, Goetzmann et al. 2022, Kendall & Charles 2022, Bybee 2023). We contribute to this

[2] See Barberis (2018) and Adam & Nagel (2023) for comprehensive reviews.

literature by documenting substantial heterogeneity in asset managers' narratives and mental models. Fourth, it contributes methodologically to the "text-as-data" literature, which has progressed from dictionary-based approaches to modern NLP and large language models capable of extracting forward-looking content and causal structure from financial texts (Gentzkow et al. 2019, Israel et al. 2020, Ke et al. 2021, Li et al. 2023, Chen et al. 2024, Sarkar & Vafa 2024), by developing an approach to extract directed, signed causal networks from CMA narratives. Finally, our work relates to the growing literature on attention and complexity (Veldkamp 2011, Schwartzstein 2014, Bordalo et al. 2019, Gabaix 2019, Bordalo et al. 2022, Gormsen & Huber 2025, Gabaix & Graeber 2024, Bastianello & Imas 2025, Gonçalves et al. 2025), by constructing text-based measures of attention and complexity from CMA causal networks and showing that these measures help explain heterogeneity in forecast levels, belief updating, and benchmark-relative forecast errors.

At a high level, our contribution is to integrate these strands of literature in a single CMA-centered empirical design: we read beliefs "in numbers", via forecasts in levels and component decompositions from tables and figures, and "in words", via disclosed modeling assumptions and an LLM pipeline that extracts the causal networks embedded in CMA narratives. This joint treatment moves us beyond documenting first- and second-moment levels and their co-movements with state variables, to identifying the assumptions and narrative mechanisms through which managers construct, update, and rationalize their expectations.

Roadmap of the paper

The remainder of this paper proceeds as follows. In Section 1 we discuss data and methodologies. In Section 2 we study equity return and second-moment (volatility and correlation) expectations "in numbers", starting from accounting identities and valuation models and examining how the resulting return-expectation components relate to asset managers' equity allocations. In Section 3 we study expectations "in words", first examining the assumptions and methods used to construct CMA inputs, then extracting the causal content of CMA narratives with an LLM pipeline, and finally comparing CMA narratives with N-CSR shareholder-letter outlooks. Section 4 concludes.

1 Data and Methodology

This section outlines our dataset construction. We hand-collect publicly available Capital Market Assumptions (CMAs) from major financial institutions covering the period 2010–2025. These documents contain both qualitative narratives and quantitative return forecasts. To enrich this primary dataset, we incorporate news sentiment data from RavenPack and mutual fund allocation data from Morningstar Direct.

1.1 Capital Market Assumptions

We document the Capital Market Assumptions (CMA) produced by large institutional investors, beginning with an overview of their role and followed by a detailed account of our dataset construction and coverage. A key innovation of our dataset is that it captures both the qualitative inputs—economic models, causal narratives, and analytical frameworks—and the corresponding quantitative outputs—explicit return expectations and second moments—within the same documents for each asset manager. This dual structure approximates a supervised learning setting, allowing us to study how different institutions map a common set of inputs into heterogeneous forecasts.

Capital Market Assumptions are forward-looking projections of return expectations, volatilities, and correlations across asset classes and serve as essential inputs for strategic asset allocation and long-term investment planning. Their institutional significance is well established (Dahlquist & Ibert 2024, 2025, Couts et al. 2023, 2024, Andonov et al. 2026). CMAs are not passive survey responses, but integral business tools produced independently by institutions—typically annually or quarterly—by dedicated teams of economists, quantitative analysts, and senior strategists who synthesize macroeconomic trends, asset-pricing models, and academic research. Forecasting horizons usually range from 3 to 30 years, with 10-year projections the most common.[3]

Many CMA reports reference prominent asset-pricing theories and empirical research, often resembling comprehensive white papers rather than simple forecast tables. This an-

[3] Some institutions focus on shorter horizons (less than 10 years), while others emphasize longer-term projections (20–30 years).

alytical depth suggests that CMAs capture more sophisticated beliefs than those typically elicited in household or professional investor surveys.

We hand-collect a dataset of long-term CMAs from both asset managers and institutional investment consultants.[4] For simplicity, we use the term asset managers to refer to both throughout the paper. Including both reflects their complementary roles in shaping institutional investment decisions. Pension funds and endowments frequently rely on consultants for strategic asset allocation guidance, while asset managers use their published CMAs to inform internal portfolio strategies and provide advisory support to external clients.[5]

[Table 1 here]

Our final sample comprises 83 distinct institutions: 59 asset managers and 24 consultants. Table 1 provides detailed information on the organizations included in the dataset along with their respective coverage periods. The CMA documents span the years 2008 to 2026, coverage varies across institutions due to differences in data availability and publication frequency. For temporal analysis, each CMA report is assigned a reference date corresponding to the data cutoff used in the institution's forecasting process, as explicitly stated in the report.[6]

[Table 2 here]

Each institution-year CMA provides return forecasts across a range of asset classes. We focus on major asset classes, selected based on three criteria: (i) relevance to institutional portfolios (e.g., U.S. Equity, Developed ex-U.S. Equity, Emerging Market Equity), (ii) consistency of appearance within institutions over time, and (iii) broad coverage across institutions. Because asset class labels vary across managers (e.g., "U.S. Large Cap Equity" vs. "U.S. Equity"), we standardize them using a consistent naming convention.[7]

Our baseline empirical analysis uses expected geometric returns as reported in the CMA

[4]CMAs are manually collected from publicly available sources, including institutional websites and archived versions retrieved via archive.org.

[5]See Andonov et al. (2026) for empirical evidence that consultant recommendations directly affect pension fund investment decisions.

[6]This reference date captures the most recent information incorporated into the institution's subjective expectations and typically precedes the publication date by weeks or months. When a report does not specify a data cutoff date, we use the publication date as a proxy.

[7]If a CMA specifies a benchmark (e.g., S&P500, MSCI EM), we use that information to guide the mapping. When no benchmark is provided, we rely on textual descriptions of regional or style exposures. If a U.S. Cash forecast is not reported, we substitute short-term U.S. Treasury bills as a proxy.

documents, without imposing a uniform forecast horizon (the median forecast horizon equal to 10 years across all asset classes). Table 2 summarizes coverage across 10 asset classes, reporting the number of institutions and years with available forecasts.

1.1.1 Return Decomposition

CMA reports typically include sections that describe the forecasting frameworks and component assumptions underlying asset managers' equity return expectations, as illustrated in Figure 2. This qualitative evidence serves as the foundation for our subsequent quantitative analysis, in which we integrate additional datasets to examine what drives these component-level beliefs and how they are formed. The widespread use of the building-block approach—first documented by Dahlquist & Ibert (2025)—motivates our investigation. We hand-collect and construct a new CMA dataset at the asset manager–year level that disaggregates forecasts into seven components: growth, dividends, valuation change, buybacks, share issuance, margin mean-reversion, and currency/inflation adjustments. This structure enables us to measure each component's contribution to belief dispersion and to relate asset manager beliefs to their narratives and subsequent portfolio allocations.

Figure 2 summarizes the equity model formulations employed by asset managers across comparable time periods, revealing substantial heterogeneity that we explore in detail in later sections. Despite institutional differences, most managers rely on a common set of foundational elements - nominal growth, valuation adjustments, and dividend yields. This convergence reflects a shared methodological framework in which return expectations are expressed as linear combinations of fundamental drivers. While these three elements form the analytical core, many institutions incorporate additional components, such as buybacks, share issuance, profit margin normalization, and currency adjustments. This additive decomposition offers a transparent mapping from macro-financial assumptions to return expectations and naturally reflects an approximate return accounting identity. At the same time, the precise structure of the decomposition varies by asset class and across managers, highlighting key sources of disagreement in return expectations.

[Figure 2 here]

1.1.2 Beliefs on Volatility and Correlations

Capital Market Assumptions report not only return expectations but also the second moments managers use for portfolio construction, namely the volatility of each asset class and the matrix of pairwise correlations across classes. We hand-collect these correlation matrices from the same reports, standardizing each report's asset-class labels to the universal classification used for the return forecasts, and obtain a panel of subjective volatilities and pairwise correlations at the asset-manager and vintage-year level. To benchmark the subjective correlations against realized comovement, we match each asset class to a traded total-return index and compute realized correlations from trailing ten- and fifteen-year monthly windows. [8]

We focus on correlation estimates for equities and bonds. The mapping from managers' asset-class labels to traded indices is tighter for equities than for fixed income. Each equity class corresponds to a small set of near-equivalent broad-market indices, so the realized benchmark is essentially unique: U.S. small-cap equity maps to the Russell 2000 or the S&P 600, and developed ex-US equity to MSCI EAFE or MSCI World ex-US. Several fixed-income classes instead span materially different instruments. "U.S. Bonds" covers both the aggregate index and mortgage-backed securities, "U.S. Government Bonds" ranges from intermediate through long and twenty-year-plus Treasuries, and "Global Bonds ex-US" combines developed-market investment grade with hard- and local-currency emerging-market debt. The realized comparison is therefore inherently more ambiguous for bonds, and this is also where subjective correlations sit furthest from history: Global Bonds ex-US has both the most heterogeneous set of candidate indices and the largest deviation from realized correlation.

[8] Indices are total-return series obtained from Bloomberg. Where a CMA specifies a benchmark we follow it; otherwise we map each class to the most representative liquid index. We exclude classes whose only available proxies are appraisal-based or illiquid, such as private equity, private real estate, hedge funds, and direct lending.

1.1.3 Text

We convert each CMA report into a structured digital format with an Optical Character Recognition and machine-learning-based workflow that preserves reading order, recovers document hierarchy, and extracts embedded visuals. The resulting structured data undergoes preprocessing. For topic modeling with Latent Dirichlet Allocation (LDA), we perform additional text preprocessing and construct a document-term matrix, as detailed in Appendix A.1. In contrast, the LLM pipeline relies solely on the initial preprocessing stage and requires no further text-specific cleaning (Appendix A.2.1). We also analyze word counts over time and across institutional investors, as described in Appendix A.3, revealing disparities in reporting practices. Finally, Appendix A.4 presents word clouds for key financial categories, highlighting the language asset managers use when discussing different topics.

1.1.4 Summary Statistics

We analyze return forecasts that asset managers publish across multiple asset classes and investment horizons. Figure 3 presents subjective return forecasts for U.S. equity, revealing two key patterns. Panel (a) and (b) show that expectations and volatilities forecasts dispersion remains consistently high throughout our sample period, suggesting that disagreement stems from fundamental differences in methodology rather than transient information asymmetries around specific events. Panel (c) examines a subset of representative investors selected for having the most comprehensive time series in our sample, documenting substantial heterogeneity in both belief persistence and baseline optimism across institutions. In Appendix A.6, Figure A9 displays the time-series variation in asset managers' return expectations across Developed Market Equity, Emerging Market Equity, U.S. Investment Grade, U.S. High Yield, U.S. Government Bond, and U.S. Cash. These patterns are consistent with our findings for U.S. Equity return expectations.

[Figure 3 here]

1.2 Other Data

We complement our primary Capital Market Assumptions dataset by merging several external data sources. First, we incorporate news-sentiment data from RavenPack to capture market information flow and investor sentiment, which we match with market capitalization data from CRSP. Second, we use actively managed U.S. mutual fund data from Morningstar Direct to study how stated market expectations relate to actual portfolio allocation decisions. Third, we obtain monthly return-predictor data from Amit Goyal's webpage to construct objective expected-return forecasts, and we supplement these with earnings, dividends, the CAPE ratio, and stock return data retrieved from Robert Shiller's webpage.

1.2.1 Ravenpack: Earnings Sentiment

Our news sentiment analysis leverages RavenPack, a platform that processes corporate and macroeconomic news from multiple sources in real time, generating quantitative sentiment scores and event classifications.[9] We focus specifically on earnings-related news events in our analysis. Earnings announcements represent the most frequent and systematically scheduled firm-specific information releases in RavenPack, providing regular updates on corporate performance. Using this earnings-focused approach, we assess news sentiment impact on the single U.S. Equity asset class rather than individual firms, aggregating firm-level news to the S&P 500 index level, including only companies identified through CRSP monthly market cap data. We define unweighted and weighted sentiment measures:

$$\bar{S}_\tau = \frac{1}{|\mathcal{I}_\tau|} \sum_{\ell \in \mathcal{I}_\tau} S_{\ell,\tau} \tag{1}$$

$$\bar{S}_\tau^{(\mathrm{wtd})} = \frac{\sum_{\ell \in \mathcal{I}_\tau} w_{\ell,\tau} S_{\ell,\tau}}{\sum_{\ell \in \mathcal{I}_\tau} w_{\ell,\tau}} \tag{2}$$

where $\mathcal{I}_\tau$ represents news items within interval τ, $S_{\ell,\tau}$ is the sentiment score of news item ℓ, and $w_{\ell,\tau}$ represents market-cap weights. In the empirical specifications below, S_t denotes

[9] We utilize RavenPack's structured metadata including sentiment scores, event classifications, and relevance indicators. Details of our processing methodology are provided in Appendix A.7.

the annual value-weighted earnings-news sentiment measure constructed from $\bar{S}_{\tau}^{(\mathrm{wtd})}$ and matched to CMA reference year t.

1.2.2 Morningstar: Equity Allocations

We investigate portfolio-allocation decisions by analysing U.S.-domiciled open-end mutual funds drawn from Morningstar Direct and matched to the asset managers in our CMA database, using the "BrandingName" variable. The resulting sample comprises funds that are classified as Equity, Fixed Income, Allocation, or Alternative in Morningstar's Global Broad Category Group. Portfolio holdings are reported monthly or quarterly per SEC requirements.

We assume CMA narratives influence fund managers' decisions within defined temporal windows. Our baseline specification considers a three-month influence window following the CMA embedded date rather than publication date. We also examine alternative windows starting before the CMA embedded date, as managers may process information early. We track net fund allocations in U.S. equity markets.

1.2.3 Certified Shareholder Reports of Registered Investment Companies

We augment the analysis with shareholder communications from mutual fund annual and semi-annual reports filed with the U.S. Securities and Exchange Commission through EDGAR. These materials are drawn primarily from Forms N-CSR, N-CSRS, and related amendments, which are periodic SEC filings used by registered management investment companies to file shareholder reports and related disclosures. They contain both standardized regulatory material and narrative sections, such as portfolio-manager discussions, performance commentary, market outlooks, and descriptions of portfolio positioning. We use these documents to compare the language of CMA reports with the language managers use when communicating investment views to fund shareholders.

We construct the filing universe from the SEC files over the period 2010–2026. We retain shareholder-report form types, primarily N-CSR, N-CSR/A, and N-CSRS, and record filing-level identifiers and metadata, including the Central Index Key (CIK), registrant name, form type, filing date, accession number, EDGAR filing-detail page, and source document URL.

For each filing, we retrieve the corresponding EDGAR filing-detail page and identify the relevant filing document. Whenever possible, we select the document whose SEC Type exactly matches the target form type. If no such document is available, we use the complete submission file and isolate the document block corresponding to the target form type before extracting text. This step reduces the amount of unrelated material processed from exhibits or other filing components.

We then extract the shareholder letter or manager-discussion section from the selected document. The extraction routine searches for textual markers associated with portfolio-manager commentary and removes standardized material such as financial statements, auditor reports, tables, and regulatory disclosures. We identify each filer using its CIK and match funds to the asset managers in our CMA database.

The resulting corpus consists of fund-specific communications directed at mutual fund investors. These texts typically discuss recent fund performance, market conditions, portfolio positioning, and forward-looking views. They therefore complement the CMA dataset by allowing us to compare managers' long-horizon capital-market assumptions with the narratives they provide to fund shareholders.

1.3 Extracting Causal Networks with LLMs

Asset managers approach complex forecasting problems by decomposing them into components and forming beliefs about how these components interact. Institutional asset managers make these reasoning chains explicit in Capital Market Assumption reports when discussing equity return expectations, articulating how economic variables transmit to returns. We extract these stated causal mechanisms to quantify heterogeneity in managers' mental models, measuring both which topics receive attention and how complex the causal reasoning structure is.

Our approach shares the intuition of directed graph frameworks that formalize causal reasoning, while adapting to the specific context of extracting stated relationships from text. The networks we extract preserve cycles and feedback loops as managers articulate them, capturing heterogeneity in how different institutions conceptualize the same forecasting problem.

Our extraction pipeline, summarized in Figure 4, transforms unstructured narratives into quantifiable networks through systematic text processing. We segment CMA reports to isolate discussions of equity returns and their drivers. We use the large language model Anthropic's Claude Sonnet 4.5 to identify explicit causal claims, extracting three elements: the causing concept, the affected concept, and the directional relationship. We harmonize these concepts across reports using a standardized financial taxonomy derived from RavenPack event categories, assigning each concept to consistent topics and subtopics. Each relationship receives a sign classification—positive when an increase in the cause produces an increase in the effect, and negative for inverse relationships. The final dataset contains, for each causal link, the source text, the standardized cause and effect nodes with their topic classifications, and the relationship sign. To isolate the mental model underlying equity forecasts, we retain only nodes connected through causal paths to U.S. equity return expectations, since managers may discuss economic relationships tangentially without incorporating them into their return expectations.[10]

[Figure 4 here]

The network representation allows us to construct two families of text-based measures. The first captures *narrative complexity*: how indirect and interconnected the manager's causal reasoning is. The second captures *topic attention*: which economic topics receive the most emphasis in the manager's reasoning. In the empirical analysis below, we denote a generic complexity measure by $C_{i,t}$ and a generic topic-attention measure by $A_{i,t}$ for asset manager i in year t.

We focus on two broad dimensions of narrative complexity. The first is *chain complexity*, which captures whether managers explain equity returns through direct links or through multi-step causal chains. We measure this dimension in two closely related ways. *Average path length* measures the average number of causal steps connecting topics in the network. A higher value means that the manager's reasoning relies on longer causal chains. For example,

[10] Additional implementation details and prompt methodologies are documented in Appendix A.2.

a direct narrative such as:

$$A \to R$$

is less chain-complex than a multi-step narrative such as:

$$A \to B \to C \to D$$

The *indirect connection ratio* captures the same economic idea from a more return-focused perspective: it measures the share of links that belong to multi-step causal paths leading to, or from, U.S. equity return. Although average path length and the indirect connection ratio are not mathematically identical, both capture the extent to which the manager's reasoning relies on intermediate mechanisms rather than direct relationships.

The second dimension is *interconnected complexity*, measured by *network transitivity*. Transitivity captures the extent to which the manager's causal narrative forms closed triangles of relationships. A high-transitivity narrative is not necessarily longer, but it is more locally interconnected. For example, if the network contains:

$$A \to B, \qquad B \to C, \qquad A \to C$$

then the relationship between A, B, and C forms a transitive triangle. High transitivity can reflect a richer and more internally connected mental model, but it need not imply more accurate forecasts; whether such complexity improves or distorts belief formation is an empirical question.

Figure 5 provides a schematic illustration of these two dimensions of narrative complexity and the three measures we use to capture them.

[Figure 5 here]

Topic attention is conceptually distinct from narrative complexity. For each economic topic, such as valuation change, economic growth, or dividend yield, we count the normalized number of causal links connected to that topic. These edge-count measures capture how much attention a manager devotes to each topic in the causal narrative, regardless of whether the overall network is simple or complex.

Table 3 reports summary statistics for the network-complexity and topic-attention measures. We construct these statistics by first computing the time-series average of each measure for each asset manager, and then calculating cross-sectional moments across managers. The coefficients of variation show substantial heterogeneity across managers. The indirect connection ratio and network transitivity have coefficients of variation close to 0.9, while topic-attention measures have coefficients of variation exceeding 1. This indicates that asset managers differ substantially both in how they structure causal reasoning and in which economic topics they emphasize.

[Table 3 here]

The extracted networks therefore transform qualitative narratives into measurable representations of managers' mental models. Differences in network architecture—the length of causal chains, the prevalence of indirect mechanisms, the degree of local interconnection, and the topics receiving attention—allow us to study why sophisticated investors disagree despite observing similar macro-financial information. The approach provides field evidence on how institutions translate complex economic environments into structured forecasting frameworks and where their conceptual models diverge.

2 Framework-Based Analysis

The substantial heterogeneity in institutional return expectations raises fundamental questions about how sophisticated investors construct their beliefs. We decompose these expectations through an approximate return-identity framework, revealing which economic fundamentals—growth, valuation changes, and shareholder distributions—drive cross-sectional disagreement among asset managers. Despite diverse proprietary models, return expectations map systematically to these common building blocks. We examine these components through a Campbell–Shiller lens to understand how beliefs about fundamentals and valuations vary across institutions and time, then link them to actual portfolio allocations to identify which components most strongly influence investment decisions. We also examine how long-horizon volatility beliefs are formed, how persistent they are, and how informative they are for realized volatility.

2.1 An Approximate Identity Framework

Text analysis of CMA reports reveals frequent references to building-block approaches in the discussion of equity return assumptions. In contrast, mentions of multi-factor risk models are comparatively rare, as most institutions emphasize fundamental drivers such as yields and growth. Consistent with this narrative, the forecast tables published by these institutions typically decompose return expectations into components closely aligned with the Gordon Growth Model—such as income yield, growth, and valuation change.

Dahlquist & Ibert (2025) note that CMA return expectations are commonly decomposed into building blocks and, using external proxies for these components, infer that valuation change is the primary source of heterogeneity in subjective equity premia.

We take a complementary route: rather than inferring components residually, we directly extract the components managers explicitly publish in their CMA tables, enabling like-for-like comparisons across asset managers and years. We find that valuation change is the single largest component of cross-sectional variation in return expectations, but it accounts for only 45% of that variation. The remaining components therefore play a substantial role, with growth expectations contributing meaningfully to disagreement across asset managers.

This approach disaggregates total equity return expectations into key components: real growth, inflation, income from shareholder distributions, and valuation adjustments. We examine the structure of these beliefs and their consistency with the standard Campbell–Shiller present-value logic in Section 2.2. The novelty is that our component-level data are directly observed from the CMA building-block tables rather than imputed from auxiliary sources, which allows us to quantify component-specific variation and pass-through with minimal modeling assumptions.

We begin by noting that asset managers provide return expectations and return decompositions in raw, non-log units. To align with this convention, we derive a simple approximate return identity that reflects the underlying logic of these reported decompositions. Starting with a one-period identity:

$$1 + R_{t+1} = \frac{D_{t+1} + P_{t+1}}{P_t} = \frac{D_{t+1}}{P_t} + \frac{P_{t+1}}{P_t} \tag{3}$$

where P_{t+1} and D_{t+1} are the price and the dividend paid at time $t+1$. We further decompose the capital and dividend yield terms into nominal growth in earnings, X_{t+1}/X_t, and growth in price-to-earnings ratios, $(P/X)_{t+1}/(P/X)_t$:

$$
\begin{aligned}
1 + R_{t+1} &= \frac{D_{t+1}}{P_{t+1}} \frac{(P/X)_{t+1}}{(P/X)_t} \frac{X_{t+1}}{X_t} + \frac{(P/X)_{t+1}}{(P/X)_t} \frac{X_{t+1}}{X_t} && (4) \\
&= (1 + DY_{t+1})(1 + G_{px,t+1})(1 + G_{x,t+1}) && (5) \\
&= (1 + DY_{t+1})(1 + G_{px,t+1})(1 + G_{real\,x,t+1})(1 + \pi_{t+1}) && (6)
\end{aligned}
$$

where $DY_{t+1} = D_{t+1}/P_{t+1}$, $1 + G_{px,t+1} = (P/X)_{t+1}/(P/X)_t$, and $1 + G_{x,t+1} = X_{t+1}/X_t$. The last expression decomposes gross nominal earnings growth into its real component and inflation:

$$
1 + G_{x,t+1} = (1 + G_{real\,x,t+1})(1 + \pi_{t+1})
$$

Some asset managers distinguish between earnings growth driven by revenue expansion and earnings growth driven by changes in profit margins. To capture this distinction, we further decompose real earnings growth as:

$$
1 + G_{real\,x,t+1} = (1 + G_{real\,adj.\,x,t+1})(1 + G_{margin,t+1})
$$

where $G_{real\,adj.\,x,t+1}$ is the real earnings growth component net of margin changes and $G_{margin,t+1}$ captures the contribution of profit margins. Incorporating this refinement leads us to:

$$
1 + R_{t+1} = (1 + DY_{t+1})(1 + G_{px,t+1})(1 + G_{real\,adj.\,x,t+1})(1 + G_{margin,t+1})(1 + \pi_{t+1}) \quad (7)
$$

Taking a first-order approximation by dropping higher-order interaction terms yields:

$$
R_{t+1} \approx DY_{t+1} + G_{px,t+1} + G_{real\,adj.\,x,t+1} + G_{margin,t+1} + \pi_{t+1} \quad (8)
$$

This identity highlights how the raw return can be expressed as a linear function of raw dividend yields, raw growth rates, and inflation. Some asset managers note that D represents the full cash flow to investors—including dividends and net share repurchases—rather than dividends alone. Following Boudoukh et al. (2007), we therefore refine the payout term as

$$DY_{t+1} = DY_{div,t+1} + DY_{\text{buyback},t+1} - DY_{\text{issuance},t+1}$$

where $DY_{div,t+1}$ is the dividend yield, $DY_{\text{buyback},t+1}$ is the yield from share repurchases, and $DY_{\text{issuance},t+1}$ is the dilution yield from net share issuance.

Taking subjective expectations and combining all terms, we express the return expectation in a form consistent with the decomposition observed in CMA reports: This identity highlights how the raw return can be expressed as a linear function of dividend yields, growth rates, and inflation. Some asset managers note that D represents the full cash flow to investors—including dividends and net share repurchases—rather than dividends alone. Following Boudoukh et al. (2007), we therefore refine the payout term as:

$$DY_{t+1} = DY_{div,t+1} + DY_{\text{buyback},t+1} - DY_{\text{issuance},t+1}$$

where $DY_{div,t+1}$ is the dividend yield, $DY_{\text{buyback},t+1}$ is the yield from share repurchases, and $DY_{\text{issuance},t+1}$ is the dilution yield from net share issuance. Taking subjective expectations from the perspective of asset manager i at CMA reference date t, we write the manager-specific return expectation as:

$$\begin{aligned} \widetilde{\mathbb{E}}_{i,t}^{CMA}[R_{t+1}] &= \widetilde{\mathbb{E}}_{i,t}\left[DY_{div,t+1} + DY_{buyback,t+1} - DY_{issuance,t+1}\right] \\ &\quad + \widetilde{\mathbb{E}}_{i,t}\left[G_{px,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[G_{real\,adj.\,x,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[G_{margin,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[\pi_{t+1}\right] \end{aligned} \tag{9}$$

where

$$\widetilde{\mathbb{E}}_{i,t}[G_{real\,x,t+1}] = \widetilde{\mathbb{E}}_{i,t}[G_{real\,adj.\,x,t+1}] + \widetilde{\mathbb{E}}_{i,t}[G_{margin,t+1}]$$

and $\widetilde{\mathbb{E}}_{i,t}$ denotes asset manager i's subjective expectation conditional on the information and assumptions embedded in its CMA at time t. This standardized identity allows us to compare levels and component contributions for the same components across asset managers and years. In practice, these components are labeled differently across CMA reports. The term $\widetilde{\mathbb{E}}_{i,t}[DY_{div,t+1}]$ is commonly referred to as "income", "net payout yield", "carry", or "dividend yield" in simpler decompositions. The valuation term $\widetilde{\mathbb{E}}_{i,t}[G_{px,t+1}]$ is often labeled as "valuation adjustment" or "change in valuation." For most asset managers, $\widetilde{\mathbb{E}}_{i,t}[G_{real\,x,t+1}]$ is captured under "EPS growth" or combined with inflation under "nominal growth", though some distinguish between the revenue-driven component $\widetilde{\mathbb{E}}_{i,t}[G_{real\,adj.\,x,t+1}]$ and the margin component $\widetilde{\mathbb{E}}_{i,t}[G_{margin,t+1}]$. Finally, $\widetilde{\mathbb{E}}_{i,t}[\pi_{t+1}]$ represents the "inflation" component.

Harmonizing these labels to a common taxonomy is a data contribution that enables the cross-asset-manager variance decompositions and pass-through tests below. This model-based framework for returns is consistent with a broad literature on equity premia, valuation, and return and cash-flow expectations. For example, Fama & French (2002) decompose the historical equity premium into contributions from dividend yields and dividend growth. Campbell & Ammer (1993) show that time variation in stock and bond returns can be attributed to changing expectations of future dividends and discount rates. Cochrane (2011) argues that fluctuations in asset prices are driven primarily by time-varying discount rates rather than by changes in expected cash flows. More recently, a growing behavioral literature has proposed subjective expectations—about cash flows as well as discount rates—as a key source of variation in asset prices (De la O & Myers 2021, Nagel & Xu 2022, Bordalo et al. 2024*a*,*b*, Bretscher et al. 2024, Bastianello 2024).

Nearly all asset managers incorporate nominal growth and dividend yield into their frameworks—Figure A12 in Appendix A.8 summarizes which building blocks each asset manager includes in their CMA decomposition. However, there is substantial variation in the inclusion of other components—particularly margin adjustments and share issuance components. Notably, dividend yield, buybacks, and share issuance are conceptually linked, as they jointly determine net shareholder payouts. We quantify this institutional heterogeneity in model structure and show that three blocks—growth, income, and valuation—dominate in practice. This documents the common accounting logic across proprietary frameworks and provides the basis for component-level variance decompositions and pass-through analyses of beliefs to allocations.

Panel (a) in Figure 6 quantifies the diversity of model structures, showing that roughly two-thirds of CMA reports include three building blocks, while fewer reports incorporate only two or more than four components. Panel (b) in Figure 6 provides additional insight by contrasting the inclusion and exclusion rates for each building block across the full sample. Blue bars indicate the percentage of asset managers that include a given factor, while orange bars represent exclusion rates. Nominal growth and dividend yield emerge as near-universal components, valuation change is included by approximately 85% of managers, and fewer than half incorporate margin adjustments or share issuance.

[Figure 6 here]

Figure 7 presents partial R^2 estimates for equity return expectations under three sets of fixed effects: asset manager, year, and investment horizon. The bar chart indicates that asset manager-specific factors explain the largest share of variation in return forecasts, substantially exceeding the contributions of temporal variation and forecast horizon. This suggests that, while market conditions and forecast periods do play a role, the identity of the forecasting institution is the primary driver of return expectation estimates. This motivates our component-level analysis below, which shows that much of this cross-institution dispersion is associated with valuation-change and growth assumptions in the published CMA blocks.

[Figure 7 here]

Figure 8 presents a stacked area chart that decomposes average subjective equity return expectations into their constituent building blocks over the period 2010–2025. The most prominent and stable contributors are inflation (blue), real growth (dark blue), and dividend yield (brown). In contrast, the valuation change component (red) exhibits notable variation. We show directly from the CMA blocks that most of the time-series variation in subjective returns comes through valuation adjustments and growth expectations, whereas inflation is comparatively stable across institutions and over time. The average subjective equity return, shown as a dash-dotted red line, exhibits substantial time variation, with a decline around 2016–2018 followed by a recovery in subsequent years.

[Figure 8 here]

Figure 9 reports the yearly standard deviations of individual building blocks from 2015–2025, capturing the cross-sectional dispersion in beliefs across asset managers. The dashed-dotted dark red line represents the overall standard deviation of subjective equity return forecasts during this period. The thickness of each area segment represents its contribution to overall dispersion. Valuation change (red) consistently accounts for the largest share of dispersion, reflecting substantial cross-sectional disagreement among asset managers. Real growth expectations (dark blue) also contribute substantially, followed by dividend yield (brown), while inflation expectations exhibit relatively low dispersion. The overall standard deviation of subjective equity return expectations (the dashed-dotted red line) remains below the cumulative dispersion of individual components, illustrating diversification effects—disagreements across components partially offset one another. A pronounced rise in cross-sectional disagreement begins in 2021 and persists thereafter, potentially reflecting periods of elevated macroeconomic uncertainty or structural market shifts, during which institutional forecasters exhibit more divergent views on future conditions.

[Figure 9 here]

Figure 10 reports partial R^2 values from a regression with time fixed effects that decomposes institutional return expectations into seven explanatory components: real growth, inflation, dividend yield, valuation change, buybacks, share issuance, and margins. By absorbing common time shocks, this specification isolates cross-sectional heterogeneity across asset managers. Valuation change and real growth emerge as the dominant factors, jointly accounting for the largest share of the variation in return forecasts across institutions (45% and 32%, respectively).

Our variance decomposition shows using the reported CMA blocks directly that although valuation change, or repricing, explains the largest component of the cross-institution variation in subjective equity return expectations, growth and the other blocks together account for the majority of the remaining variation, approximately 55%.

[Figure 10 here]

2.2 Campbell-Shiller Decomposition

In line with works studying the relationship between return expectations, cash-flow expectations, and valuation ratios (De la O & Myers 2021, Dahlquist & Ibert 2024, Bordalo et al. 2024*b*), we consider a subjective Campbell–Shiller decomposition to study the building blocks of large institutional asset managers' return expectations. We denote asset manager i's subjective expectation operator at time t by $\widetilde{\mathbb{E}}_{i,t}$. We estimate asset-manager-level Campbell–Shiller regressions using each institution's own component series.

We first show that the countercyclicality of CMA return expectations arises from two component-level patterns: negative loadings of valuation change expectations on valuation ratios and positive, procyclical loadings of growth expectations. We then compare the subjective relationships among return components with their objective counterparts in realized data.

2.2.1 Cyclicality of Return Expectations and their Building Blocks

Following Campbell & Shiller (1987, 1988), log-linearizing around the one-period return identity, the log price–dividend ratio pd_t can be written as:

$$pd_t = \kappa + \Delta d_{t+1} - r_{t+1} + \rho \, pd_{t+1} \tag{10}$$

where Δd_{t+1} is future dividend growth, r_{t+1} is the future log return, and pd_{t+1} is the future price–dividend ratio. Taking subjective expectations from the perspective of asset manager i, taking the covariance of the left- and right-hand sides with respect to pd_t, and dividing by the variance of pd_t, we obtain:

$$1 \;=\; \underbrace{\frac{\mathrm{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}], pd_t\Big)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{\Delta d},pd)}} - \underbrace{\frac{\mathrm{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[r_{t+1}], pd_t\Big)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{r},pd)}} + \rho \underbrace{\frac{\mathrm{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[pd_{t+1}], pd_t\Big)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{pd},pd)}} \tag{11}$$

where $\beta^i_{(\widetilde{\Delta d},pd)}$ is the regression coefficient of asset manager i's subjective dividend-growth expectations on the price–dividend ratio, $\beta^i_{(\widetilde{r},pd)}$ is the regression coefficient of asset manager i's subjective return expectations on the price–dividend ratio, and $\beta^i_{(\widetilde{pd},pd)}$ is the regression coefficient of asset manager i's subjective expected future price–dividend ratio on the current price–dividend ratio. To express (11) in terms of valuation changes, we add and subtract

pd_t from the first argument of the last covariance term in (11) and obtain:

$$1-\rho \quad = \quad \beta^i_{(\widetilde{\Delta d},pd)} - \beta^i_{(\widetilde{r},pd)} + \rho \underbrace{\frac{\text{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}], pd_t\Big)}{\text{Var}(pd_t)}}_{\beta^i_{(\widetilde{\Delta pd},pd)}} \tag{12}$$

where Δpd_{t+1} is the one-period change in the log price–dividend ratio and $\beta^i_{(\widetilde{\Delta pd},pd)}$ is the regression coefficient of asset manager i's subjective expected change in the future price–dividend ratio on the current price–dividend ratio. Equation (12) highlights how the regressions of $\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]$, $\widetilde{\mathbb{E}}_{i,t}[r_{t+1}]$, and $\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]$ on pd_t are connected through the accounting identity.

To introduce earnings into the identity, in the spirit of De la O & Myers (2021), we use $pd_t = px_t - dx_t$, equivalently $px_t = pd_t + dx_t$, where x_t denotes smoothed log earnings and $dx_t = d_t - x_t$ is the smoothed log payout ratio. Substituting into (10) gives:

$$px_t \quad = \quad \kappa + \Delta x_{t+1} - r_{t+1} + (1-\rho)dx_{t+1} + \rho\, px_{t+1} \tag{13}$$

Since $1-\rho$ is close to zero, movements in the smoothed payout ratio do not play a large role in explaining variation in px_t, and we approximate (13) by:

$$px_t \quad \approx \quad \kappa + \Delta x_{t+1} - r_{t+1} + \rho\, px_{t+1} \tag{14}$$

Following the same procedure, taking subjective expectations from the perspective of asset manager i, taking covariances with px_t, and dividing by $\text{Var}(px_t)$, we obtain:

$$1-\rho \quad = \quad \underbrace{\frac{\text{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[\Delta x_{t+1}], px_t\Big)}{\text{Var}(px_t)}}_{\beta^i_{(\widetilde{\Delta x},px)}} - \underbrace{\frac{\text{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[r_{t+1}], px_t\Big)}{\text{Var}(px_t)}}_{\beta^i_{(\widetilde{r},px)}} + \rho \underbrace{\frac{\text{Cov}\Big(\widetilde{\mathbb{E}}_{i,t}[\Delta px_{t+1}], px_t\Big)}{\text{Var}(px_t)}}_{\beta^i_{(\widetilde{\Delta px},px)}} \tag{15}$$

where $\beta^i_{(\widetilde{\Delta x},px)}$ is the regression coefficient of asset manager i's subjective expected change in long-term earnings on the price-to-long-term-earnings ratio, $\beta^i_{(\widetilde{r},px)}$ is the regression coefficient of asset manager i's subjective return expectations on the price-to-long-term-earnings ratio, and $\beta^{i}_{(\widetilde{\Delta px},px)}$ is the regression coefficient of asset manager i's subjective expected change in the valuation ratio on the price-to-long-term-earnings ratio.

Equations (12) and (15) hold at the individual or consensus level and motivate the first part of our empirical analysis. We focus on individual-level regressions, which avoid aggregation issues related to changes in sample composition and forecast horizons. Specifically, we focus on CMA asset managers for which we collect at least five years of data. We construct the log price–dividend ratio of the S&P 500 from Robert Shiller's website as a proxy for pd_t and the log cyclically adjusted price–earnings ratio from Robert Shiller's website as a proxy for px_t.

We map the theoretical objects into observed CMA components as follows. The headline U.S. equity return forecast, denoted $F^{ER}_{i,t,h}$, is the empirical counterpart of $\widetilde{\mathbb{E}}_{i,t}[r_{t+1}]$. The nominal-growth component, denoted $F^{g}_{i,t,h}$, is the empirical counterpart of $\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]$ or $\widetilde{\mathbb{E}}_{i,t}[\Delta x_{t+1}]$. When reports provide real growth and inflation separately, we define nominal growth as the sum of the two components. The valuation-change component, denoted $F^{VC}_{i,t,h}$, is the empirical counterpart of $\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]$ or $\widetilde{\mathbb{E}}_{i,t}[\Delta px_{t+1}]$.

In Figure 11, for each asset manager i and forecast component $b \in \{ER, NG, VC\}$, we estimate regressions of the form:

$$F^{b}_{i,t,h} = \alpha^{b}_{i} + \beta^{b}_{i,pd}\, pd_t + \varepsilon^{b}_{i,t,h} \tag{16}$$

$$F^{b}_{i,t,h} = \alpha^{b}_{i} + \beta^{b}_{i,px}\, px_t + \varepsilon^{b}_{i,t,h} \tag{17}$$

Here, $b = ER$ denotes the headline U.S. equity return forecast, $b = NG$ denotes nominal growth, and $b = VC$ denotes valuation change. Figure 11 displays the estimated asset-manager-level slopes. Consistent with prior evidence, the majority of asset managers exhibit a negative relationship between return expectations and either the CAPE or the price–dividend ratio. The relationship changes sign for the nominal-growth component: the majority of asset managers exhibit positive coefficients, implying that expected nominal growth is procyclical. Finally, expected valuation change is negatively related to price-to-fundamentals ratios for most managers, making valuation adjustment the main driver of the countercyclicality of asset managers' total return expectations.

[Figure 11 here]

The countercyclical nature of CMA return expectations suggests that, unlike the expectations of retail investors documented by Greenwood & Shleifer (2014), institutional return expectations display properties closer to those implied by classic asset-pricing theories, in which expected returns are high when valuations are low (e.g., Campbell & Cochrane 1999). We next ask whether the building blocks underlying these expectations also exhibit properties consistent with standard asset-pricing logic.

2.2.2 A Conceptual Network Approach

A natural way to assess whether the building blocks are consistent with accounting identities is to examine whether the relationships among them align with those implied by objective data. We also represent these relationships as networks, which we use throughout the rest of the paper and refer to as "causal networks." We use this terminology loosely. By "causal," we mean that the network encodes directional links implied by the accounting structure or stated by managers in their narratives, not necessarily causal effects in the econometric sense.

Network representations arise naturally in asset-pricing settings because return identities link expected returns, expected cash-flow growth, and expected valuation changes. Below, we first derive the network implied by the Campbell–Shiller identity and then illustrate how additional model assumptions generate further links using a standard consumption-based asset-pricing model. In Appendix A.9, we provide the network representation for the AQR (2017) CMA report as a practical example.

A Consumption-Based Asset Pricing Model. Starting from the Campbell–Shiller approximation,

$$r_{t+1} = \kappa + \Delta d_{t+1} - pd_t + \rho \, pd_{t+1} \tag{18}$$

$$= \kappa + \Delta d_{t+1} - (1-\rho) pd_t + \rho \, \Delta pd_{t+1} \tag{19}$$

Taking subjective expectations from the perspective of asset manager i at CMA reference date t yields

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] \;=\; \kappa + \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] - (1-\rho)\,pd_t + \rho\,\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}] \tag{20}$$

We represent the links among these objects using partial derivatives. Because the Campbell–Shiller approximation is an accounting identity, these links should be interpreted as directional accounting relationships rather than causal effects in the econometric sense.[11] Graphically, these directional accounting relationships can be represented as the network shown in Figure 12, with the black and blue nodes and directed links. The Campbell–Shiller (CS) network provides a baseline representation of the links among subjective expected returns, subjective expected cash-flow growth, valuation ratios, and subjective expected valuation changes.

To illustrate how economic structure adds links to this baseline accounting network, we consider a simple consumption-based asset-pricing model that agent i might use as a mental model. The agent considers a perfectly competitive endowment economy with complete markets. In his mind, the representative agent has lifetime utility over consumption:

$$U_t = \widetilde{\mathbb{E}}_{i,t}\left[\sum_{j=0}^{\infty} \beta^j \frac{C_{t+j}^{1-\gamma}}{1-\gamma}\right] \tag{21}$$

[11] Formally, the directional accounting links implied by equation (20) are obtained by solving the identity for each object in turn. They are given by:

$$\frac{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]} = 1, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]} = \rho, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]}{\partial pd_t} = -(1-\rho),$$

$$\frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]} = 1, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]}{\partial pd_t} = 1-\rho, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]} = -\rho,$$

$$\frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]} = \frac{1}{\rho}, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]} = -\frac{1}{\rho}, \qquad \frac{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]}{\partial pd_t} = \frac{1-\rho}{\rho},$$

$$\frac{\partial pd_t}{\partial \widetilde{\mathbb{E}}_{i,t}[r_{t+1}]} = -\frac{1}{1-\rho}, \qquad \frac{\partial pd_t}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}]} = \frac{1}{1-\rho}, \qquad \frac{\partial pd_t}{\partial \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]} = \frac{\rho}{1-\rho}.$$

subject to the budget constraint:

$$\begin{aligned} W_{t+1} &= \omega_t'(\mathbf{1} + \mathbf{R}_{t+1})(W_t - C_t) \\ \omega_t'\mathbf{1} &= 1 \end{aligned} \tag{22}$$

Here, ω_t is an $N \times 1$ vector of portfolio weights and $\mathbf{R}_{t+1}$ is an $N \times 1$ vector of returns. With no labor income, the equilibrium resource constraint is $C_t = D_t$. The law of motion for log consumption, and hence dividend growth, is:

$$\Delta d_t = \Delta c_t = (1 - \phi)\mu + \phi \Delta c_{t-1} + \epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} N(0, 1) \tag{23}$$

Solving the model by deriving the Euler equation and using a guess-and-verify approach yields a price–dividend ratio that is linear in consumption growth. The model-implied counterparts of the objects in equation (20) can then be written as functions of consumption growth:

$$pd_t = A + B\,\Delta c_t, \tag{24}$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] = (1 - \phi)\mu + \phi \Delta c_t, \tag{25}$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}] = B(1 - \phi)(\mu - \Delta c_t), \tag{26}$$

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] = \kappa + (\rho - 1)A + (1 - \phi)\mu(1 + \rho B) + [\phi(1 + \rho B) - B]\,\Delta c_t \tag{27}$$

where A and B are functions of preference parameters and the consumption-growth process, including β, γ, ρ, κ, μ, and ϕ. These additional relationships are reflected by the red nodes and links in Figure 12. This example illustrates how asset-pricing models enrich the Campbell–Shiller baseline network by adding economically motivated links between state variables and the objects in the return identity.

[Figure 12 here]

We use the network representation to guide two empirical analyses: first, an analysis of the relationships across return-expectation components using the extracted building blocks in Section 2.2.3; and second, an analysis of the textual explanations behind return expectations in Section 3.

2.2.3 A Network Approach: Correlations

The networks displayed in Figure 12 imply relationships among the building-block components. We examine whether the subjective relationships embedded in asset managers' reported building blocks resemble those observed in realized data.

In the empirical analysis, we work with the reported and harmonized CMA components. For asset manager i, reference date t, and forecast horizon h, we write the approximate reported building-block identity as

$$F^{ER}_{i,t,h} \approx \underbrace{F^{DY}_{i,t,h}}_{DY} + \underbrace{F^{g}_{i,t,h}}_{g} + \underbrace{F^{VC}_{i,t,h}}_{VC} + \underbrace{F^{\pi}_{i,t,h}}_{\pi} \tag{28}$$

Here, $F^{ER}_{i,t,h}$ denotes the headline U.S. equity return forecast, $F^{DY}_{i,t,h}$ the dividend-yield component, $F^{g}_{i,t,h}$ the real-growth component, $F^{VC}_{i,t,h}$ the valuation-change component, and $F^{\pi}_{i,t,h}$ the inflation component. When the horizon is fixed or suppressed by the empirical design, we write these components as $F^{b}_{i,t}$. We compare the inter-component relationships embedded in managers' subjective decompositions in (28) with the corresponding relationships estimated from realized data from Robert Shiller's online dataset. This exercise asks whether the subjective co-movement among reported building blocks is consistent with the realized co-movement among the corresponding return components.

Table 4 compares the relationships among the four building blocks in managers' subjective decompositions with the corresponding relationships in historical data. We treat each manager–year observation as one draw from a manager's joint decomposition and ask how the four building blocks co-move within a single forecaster. The subjective within-manager column demeans each component by asset manager before pooling, isolating how a given manager revises growth, inflation, dividend yield, and valuation change.

The main differences are concentrated in the valuation-change pairs. Real growth and valuation change are negatively correlated in the subjective within-manager panel ($r = -0.411$), while the corresponding realized correlation over 1871–2014 is positive ($r = +0.308$); the

Fisher z test rejects equality at $p = 0.020$. Managers similarly embed a positive relationship between inflation and valuation change ($r = +0.279$), whereas the realized counterpart is negative ($r = -0.358$); the Fisher z test rejects equality at $p = 0.042$. The Jennrich omnibus test also rejects equality of the full correlation matrices ($\chi^2 = 14.18$, $p = 0.028$). We interpret these differences as evidence that the subjective relationships among return components differ from their realized counterparts, motivating our subsequent analysis of deviations of return expectations from objective benchmarks (Bianchi et al. 2022, 2025).

[Table 4 here]

2.3 CMA Forecasts and Objective Benchmarks

Given the preliminary evidence that the relationships among subjective building blocks differ from their objective counterparts, we next examine return forecasts relative to an ex-ante objective benchmark constructed using machine-learning methods. We first show that this benchmark outperforms asset managers' forecasts, and then compare it with managers' subjective forecasts to study systematic deviations from the benchmark (Bianchi et al. 2022).

2.3.1 Evaluating CMA Forecasts: An Objective Forecasting Benchmark

Recent machine-learning evidence (Gu et al. 2020, Shen & Xiu 2025) suggests that the predictive content for aggregate equity returns is often diffuse, with information spread across many weak predictors rather than concentrated in a small number of dominant variables. This motivates the use of regularized forecasting methods that are designed to aggregate weak signals in high-dimensional predictor sets while controlling for overfitting. At the same time, there is substantial uncertainty about the appropriate specification, penalty structure, and modeling approach. We therefore construct the objective benchmark as an equal-weighted ensemble of regularized return-forecasting models (see Appendix A.11).[12] This combines the machine-learning insight that many weak predictors may matter with the forecast-combination logic of Stock & Watson (2004) and Rapach et al. (2010), whereby averaging across models can

[12] The models included in the ensemble forecasts are Ridge, Lasso, Elastic Net, Random Forest, Gradient Boosting Regression Trees, and two single-hidden-layer feed-forward neural networks: one with an L_2 weight decay penalty and one with an L_1 subgradient penalty.

improve robustness when no single specification is known to be dominant.

We denote the objective machine-learning forecast of the headline U.S. equity return at time t and horizon h by $G_{t,h}$, and asset manager i's corresponding CMA forecast by $F^{ER}_{i,t,h}$. We test for differences in predictive accuracy between $G_{t,h}$ and $F^{ER}_{i,t,h}$ using the Diebold & Mariano (1995) test. For each horizon h, define the squared-error loss differential:

$$d_{i,t,h} = e^2_{\text{CMA},i,t,h} - e^2_{\text{G},t,h}$$

where

$$e_{\text{CMA},i,t,h} = R_{t,t+h} - F^{ER}_{i,t,h}, \qquad e_{\text{G},t,h} = R_{t,t+h} - G_{t,h}$$

Here, $R_{t,t+h}$ is the annualized realized U.S. equity return from t to $t+h$. Let $\bar{d}_h$ be the average of $d_{i,t,h}$ across the manager–year observations available at horizon h, and let n_h denote the number of forecast pairs. The Diebold–Mariano statistic is

$$\text{DM}_h = \frac{\bar{d}_h}{\sqrt{\widehat{V}(\bar{d}_h)/n_h}} \tag{29}$$

where $\widehat{V}(\bar{d}_h)$ is the Newey–West HAC long-run variance estimator with lag $h-1$ (Newey & West 1987). A positive value of DM_h indicates that the objective benchmark has lower squared forecast errors than the CMA forecasts.

Table 5 reports the results comparing the ensemble forecasts with the CMA forecasts collected in this paper in Panel (a), and with the external CMA dataset of Dahlquist & Ibert (2024) in Panel (b). The results indicate that the ensemble forecasts systematically outperform the CMA forecasts, motivating our subsequent analysis of the systematic differences between asset managers' forecasts and our machine-learning benchmark.

[Table 5 here]

2.3.2 Feature Importance: Objective vs. Subjective

To understand the source of differences between asset managers' forecasts and machine-learning forecasts, we apply the same machine-learning methodology to two prediction problems. First, we use the models to predict annualized realized U.S. equity returns, $R_{t,t+h}$, which generate the objective benchmark $G_{t,h}$. Second, we use the same models to predict asset managers' headline CMA forecasts, $F^{ER}_{i,t,h}$. We then compare the feature importance measures recovered from the two exercises to assess which predictors explain realized return dynamics and which predictors explain CMA forecast dynamics.

Figure 13 reports the corresponding feature importance measures in two panels. Panel (a) shows the normalized permutation importance of each of the 16 predictors across the seven regularized learners used to forecast $R_{t,t+h}$. For each learner, predictor importance is measured using the same model-agnostic permutation procedure: we record the average reduction in in-sample R^2 when a predictor's values are randomly shuffled. Using a common importance measure makes the columns in Panel (a) directly comparable. Across specifications, the machine-learning models assign the largest importance to valuation variables and the investment-to-capital ratio.

Panel (b) compares the ensemble importance recovered from the realized-return models with the ensemble importance recovered from the CMA-forecast models. The horizontal axis reports the equal-weighted ensemble importance from the models fitted to realized U.S. equity returns, $r_{t,t+h}$; the vertical axis reports the equal-weighted ensemble importance from the models fitted to CMA headline forecasts, $F^{ER}_{i,t,h}$. To place the two sides on a common scale, we construct an ensemble importance vector for each target by averaging the normalized permutation importance measures across the same seven base learners. Points above the 45-degree line indicate predictors that receive greater importance in explaining CMA forecasts than in explaining realized returns, while points below the line indicate predictors that receive greater importance in explaining realized returns than in explaining CMA forecasts.

[Figure 13 here]

The comparison suggests that the machine-learning models use different signals when predicting realized returns than when predicting asset managers' forecasts. Although valuation-related variables receive some importance in both exercises, the CMA-forecast models assign relatively greater importance to a broader set of macro-financial predictors, especially inflation, whereas the realized-return models assign relatively greater importance to valuation ratios, particularly the earnings–price and dividend–price ratios. We therefore next examine the dynamics of asset managers' return expectations to better understand the origins of the benchmark-relative deviations, $G_{t,h} - F^{ER}_{i,t,h}$.

2.3.3 Forecast Dynamics: Consensus, Errors, and News

Dynamics and Consensus. We begin by testing whether peer dynamics help explain changes in asset managers' beliefs. Specifically, we ask whether managers who are far from the cross-sectional consensus subsequently revise their forecasts toward that consensus.

For the headline return expectation and each forecast component (growth, dividend yield, valuation change, inflation), $b \in \{ER, g, DY, VC, \pi\}$, we estimate the following forecast-change specification:

$$\Delta F^b_{i,t} \quad = \quad \alpha^b_i + \beta^b Dev^b_{i,t-1} + \delta^b \Delta \bar{F}^b_{-i,t} + \gamma^b X^b_{t-1} + \varepsilon^b_{i,t} \tag{30}$$

where $\Delta F^b_{i,t} = F^b_{i,t} - F^b_{i,t-1}$ denotes the change in asset manager i's forecast for component b, and $Dev^b_{i,t-1}$ is the lagged signed deviation from the peer consensus:

$$Dev^b_{i,t-1} = F^b_{i,t-1} - \bar{F}^b_{-i,t-1}, \qquad \bar{F}^b_{-i,t-1} = \frac{1}{N^b_{t-1} - 1} \sum_{j \neq i} F^b_{j,t-1}$$

Here, N^b_{t-1} denotes the number of asset managers reporting component b in year $t-1$, so $\bar{F}^b_{-i,t-1}$ is the previous year's leave-one-out consensus forecast for component b, formed excluding manager i. We also control for the contemporaneous change in the leave-one-out consensus, $\Delta \bar{F}^b_{-i,t} = \bar{F}^b_{-i,t} - \bar{F}^b_{-i,t-1}$, which captures information arriving between $t-1$ and t that is reflected in the forecasts of other managers. Finally, X^b_{t-1} denotes the lagged yearly realization of the corresponding component.

The coefficient of interest is β^b. Under the null that managers' changes are driven only by new information, lagged deviations from the consensus should not predict subsequent forecast changes once we control for contemporaneous changes in peer beliefs and lagged realized fundamentals. Hence, the null is $\beta^b = 0$. By contrast, $\beta^b < 0$ indicates convergence toward the peer consensus: managers whose forecasts were above the peer consensus subsequently revise downward, while managers whose forecasts were below the consensus revise upward.

Table 6 presents the results. The coefficient on the lagged deviation from the leave-one-out consensus is negative and statistically significant for all five forecast components: equity return, real growth, dividend yield, valuation change, and inflation. The estimated magnitudes are economically large, ranging from approximately -0.49 for real growth to -0.93 for valuation change. These estimates imply that forecast outliers do not remain fixed in their relative positions. Instead, conditional on contemporaneous changes in peer forecasts and lagged realized fundamentals, managers systematically revise toward the center of the cross-sectional belief distribution.

The contemporaneous change in the leave-one-out consensus is positive and significant for equity return, dividend yield, valuation change, and inflation. This suggests that individual forecast changes move with the information incorporated by other managers during the same period. Importantly, the convergence result survives after controlling for this common contemporaneous component. Thus, the negative estimates of β^b are not simply capturing common news or aggregate shifts in return expectations. They indicate that a manager's prior distance from the peer consensus contains incremental predictive power for subsequent forecast changes.

[Table 6 here]

We note that specification (30) is not a pure test of anchoring. In the spirit of the anchoring literature (Tversky & Kahneman 1974, Campbell & Sharpe 2009, Bastianello 2025), we therefore test whether managers' deviations from an objective benchmark are related to the distance between a salient anchor and that same benchmark. We use the rolling 365-day CMA consensus as the anchor and our out-of-sample machine-learning forecasts as the objective benchmark. Specifically, we estimate:

$$G_{t,h} - F^{ER}_{i,t,h} \quad = \quad \alpha_i + \beta^{ER} \left(G_{t,h} - \bar{F}^{ER,365d}_{t-1,h} \right) + \varepsilon_{i,t,h} \tag{31}$$

where $F^{ER}_{i,t,h}$ is manager i's headline U.S. equity return forecast at report date t and horizon h, $G_{t,h}$ is the machine-learning benchmark forecast for the same year and horizon, and $\bar{F}^{ER,365d}_{t-1,h}$ is the rolling consensus forecast constructed from horizon-matched headline U.S. equity CMA forecasts issued during the previous 365 days. The key parameter is β. If managers anchor on the recent CMA consensus, then individual forecast deviations from the machine-learning benchmark should be positively related to the distance between the rolling consensus and that benchmark. Hence, a positive and statistically significant β is consistent with anchoring: when the recent consensus lies above the machine-learning benchmark, individual forecasts also tend to lie above the benchmark, and vice versa.

Table 7 presents the results. Column (1) shows a large and statistically significant coefficient of 0.748, with a within-R^2 of 0.232. This indicates that the rolling benchmark–consensus wedge explains a substantial share of within-manager variation in individual deviations from the machine-learning benchmark. The magnitude is also economically meaningful: a one percentage point increase in the distance between the benchmark and the recent CMA consensus is associated with a roughly 0.75 percentage point increase in the individual manager's benchmark-minus-forecast deviation. Column (2) provides a robustness check using the lagged machine-learning benchmark in the consensus deviation, $G_{t-1,h} - \bar{F}^{ER,365d}_{t-1,h}$, while the dependent variable remains the manager's current deviation from the current benchmark, $G_{t,h} - F^{ER}_{i,t,h}$. The coefficient remains positive and statistically significant, though smaller at 0.400. This shows that past deviations of the CMA consensus from the machine-learning benchmark persist in current individual forecast biases: when the recent consensus was low relative to last year's benchmark, individual managers continue to issue forecasts that are low relative to the current benchmark.

[Table 7 here]

The results point to a coherent view of forecast dynamics. Individual forecasts converge toward the cross-sectional consensus: managers above consensus revise downward, while managers below consensus revise upward. At the same time, relative to the machine-learning benchmark, managers deviate systematically in the direction of the recent consensus. Thus, long-horizon CMA forecasts appear to be shaped by a consensus anchor rather than solely by mechanical responses to realized fundamentals. The consensus is therefore an economically meaningful belief component, not just a statistical average of noisy individual forecasts.

Forecasts and News. Building on our analysis of institutional return expectations and their building-block decomposition, we now examine how forecast changes relates to the flow of earnings news. To implement this analysis, we construct an aggregate S&P 500 earnings-news sentiment index, following the methodology described in Section 1.2, as a proxy for fundamental news shocks.

We study the relationship between annual forecast changes and earnings-news sentiment. For the headline return expectation and each forecast component, $b \in \{ER, g, DY, VC\}$, we estimate:

$$F_{i,t}^b - F_{i,t-1}^b \quad = \quad \mu_i^b + \beta^b S_t + \varepsilon_{i,t}^b \tag{32}$$

where S_t denotes the annual value-weighted S&P 500 RavenPack earnings-news sentiment measure.

Table 8 reports whether changes in institutional long-horizon forecasts respond systematically to earnings-news sentiment. The results show a clear effect on return forecasts and valuation-change forecasts: more positive earnings news leads to upward changes in return and valuation-change expectations.

[Table 8 here]

2.4 Relevance: Expectations and Allocations

Having established how institutional return expectations decompose into fundamental building blocks, we now examine whether these stated beliefs relate to actual investment decisions. To investigate this relationship, we link Capital Market Assumptions (CMA) forecasts to fund allocation data from Morningstar Direct (described in Section 1.2), testing whether managers' published expectations align with their portfolio positioning.

We estimate the following panel specification:

$$w_{f,t} = \mu_f + \tau_t + \boldsymbol{\beta}'\mathbf{X}_{i(f),t} + \varepsilon_{f,t} \tag{33}$$

where $w_{f,t}$ denotes fund f's U.S. equity net allocation measured between the previous and next quarter of the CMA reporting date, μ_f are fund fixed effects, τ_t are year fixed effects, and $\mathbf{X}_{i(f),t}$ denotes the CMA return forecast, excess-return forecast, or vector of building-block components of the asset manager $i(f)$ associated with fund f. The notation $i(f)$ makes explicit that the dependent variable is measured at the fund level, while CMA expectations are measured at the asset-manager level.

Table 9 presents three specifications. The first examines CMA-implied subjective excess returns across equity asset classes, the second focuses on subjective U.S. equity returns, and the third decomposes these returns into their constituent building blocks: nominal growth (NG), valuation change (VC), and dividend yield (DY). Following Dahlquist & Ibert (2024), column 1 shows that institutional investors exhibit systematic relationships between their stated return expectations and portfolio allocations. Our results replicate their key finding: higher subjective U.S. equity return expectations are associated with increased domestic equity weights, while more optimistic projections for foreign markets correspond to reduced U.S. allocations, as shown in Column 1. This indicates that CMA expectations are economically relevant for portfolios.

Using our building-block decomposition data, Column 3 examines whether valuation expectations, growth assumptions, or income components are differentially associated with portfolio weights. The results show that while the dividend-yield component has limited explanatory power, both nominal growth and valuation change exhibit positive and statistically significant coefficients—0.46 and 0.23, respectively. This component-level evidence is new: prior CMA papers document overall pass-through or identity-based mechanisms but do not identify which specific published components are most closely associated with allocations. We show that allocations load on g and VC, rather than on DY, thereby highlighting the specific channels through which beliefs are reflected in portfolio weights. The estimates indicate that both the procyclical growth component and the countercyclical valuation component are closely associated with asset allocation decisions.

These findings underscore the importance of understanding how the individual components of CMA return expectations relate to portfolio choices. If one interprets valuation change as capturing perceived mispricing in equity markets, then the positive coefficient on this term aligns with Bastianello & Peng (2024), who document that global fund managers' asset allocation intentions are positively correlated with their perceived mispricings.

[Table 9 here]

2.5 Volatility and Correlation Forecasts

Volatility. We next examine volatility forecasts. We decompose the variation in subjective volatility forecasts to assess whether heterogeneity stems from persistent manager characteristics or time-varying market conditions. We then investigate how managers form volatility forecasts—specifically, whether they anchor to long-term historical averages or incorporate recent market information—and assess whether these forecasts contain information about future realized volatility.

Let $F^{\sigma}_{i,t,h}$ denote asset manager i's subjective U.S. equity volatility forecast at CMA reference date t and horizon h. We decompose the variation in volatility forecasts $F^{\sigma}_{i,t,h}$ to assess whether heterogeneity stems from persistent manager characteristics or time-varying market conditions. We estimate a saturated fixed-effects model:

$$F^{\sigma}_{i,t,h} = \alpha + \mu_i + \tau_t + \eta_h + \varepsilon^{\sigma}_{i,t,h} \tag{34}$$

where μ_i captures manager-specific effects, τ_t represents common time variation, and η_h accounts for forecast-horizon effects.

To quantify each component's contribution, we apply a Shapley-value decomposition of the R^2, which provides order-invariant variance contributions by averaging marginal effects across all possible variable orderings. Figure 14 shows that the model explains 88.1% of total variation in volatility forecasts. Manager fixed effects account for the largest share: 72.1% of explained variance, while time effects contribute 25.3% and horizon effects only 2.6%.

[Figure 14 here]

This decomposition shows that volatility forecasts are predominantly shaped by persistent manager-specific characteristics rather than shared responses to market conditions. The dominance of manager effects suggests that institutional investors maintain stable, heterogeneous frameworks for assessing market risk over time.

To understand how asset managers form volatility forecasts, we examine how these forecasts relate to historical realized volatility. We employ exponentially weighted moving average (EWMA) models with endogenously estimated decay parameters to quantify the extent to which volatility expectations are backward-looking. Our analysis shows that managers place substantial weight on long-term historical averages while exhibiting limited responsiveness to recent volatility innovations.

Figure 15 shows that managers' volatility forecasts cluster tightly within a narrow band of 15–17% annually, displaying minimal variation despite substantial fluctuations in realized volatility over the sample period. The subjective forecasts remain highly stable across time, even as EWMA realized-volatility measures with different decay parameters exhibit considerable variability. This compression of forecasts suggests that managers anchor their volatility forecasts on long-term historical norms rather than responding strongly to current market conditions.

[Figure 15 here]

To investigate how asset managers weight past volatility when forming forecasts, we estimate an optimal exponentially weighted moving average model following the approach in Greenwood & Shleifer (2014). Suppressing the forecast horizon for this exercise, we fit the nonlinear regression

$$F^{\sigma}_{i,t} = \alpha + \beta \cdot \text{EWMA}_t(\lambda) + \varepsilon^{\sigma}_{i,t} \tag{35}$$

where $F^{\sigma}_{i,t}$ is asset manager i's subjective volatility forecast and

$$\text{EWMA}_t(\lambda) = \sum_{\ell=1}^{L} \omega_\ell(\lambda)\, RV_{t-\ell}, \qquad \omega_\ell(\lambda) = \frac{\lambda^{\ell-1}}{\sum_{m=1}^{L} \lambda^{m-1}}$$

Here, $RV_{t-\ell}$ denotes realized volatility in lagged quarter $t-\ell$, and the decay parameter $\lambda \in [0,1]$ governs the persistence of memory in the averaging process: values near unity indicate that distant observations retain substantial influence, while lower values imply rapid decay in the relevance of past information. We estimate this model using quarterly realized volatility computed from daily S&P 500 returns over a 10-year lookback window.

Table 10 presents the optimal EWMA estimates using quarterly data. The estimated decay parameter is $\hat{\lambda} = 0.9814$ (s.e. $= 0.0103$), corresponding to a half-life of 36.82 quarters—approximately 9.2 years. This high persistence parameter indicates that managers place substantial weight on long-term historical volatility when forming forecasts, consistent with our earlier finding that subjective volatility is compressed relative to realized measures.

[Table 10 here]

Correlations. We study managers' correlation forecasts for equity and bond asset classes. We benchmark managers' subjective correlation assumptions against two references. The first is the cross-manager consensus, defined for each asset-class pair and survey vintage as the leave-one-out average of the other managers' assumed correlations, which measures disagreement without contaminating a manager's benchmark with its own value. The second is the realized correlation, computed from trailing ten-year monthly total returns on the Bloomberg indices mapped to the two asset classes in each pair, which measures deviation from recent history.

Figure 16 contrasts the full cross-manager distribution of subjective correlations with U.S. large-cap equity against the realized benchmark, class by class—U.S. Small-Cap Equity, Developed ex-US Equity, Emerging-Market Equity, U.S. Aggregate Bonds, U.S. Government Bonds, and U.S. High-Yield Corporate Bonds. Subjective correlations among the major equity classes are high and close to their realized counterparts: U.S. large-cap equity is assumed to correlate near 0.88 with U.S. small-cap equity and around 0.81 with developed ex-US equity, with a lower and more dispersed assumption for emerging-market equity, near 0.71. Similarly, equity-bond correlations strongly comove to their realized historical counterparts.

[Figure 16 here]

Figure 17 summarizes the mean absolute deviation of subjective correlations from the leave-one-out consensus—Panel (a)—and from realized history—Panel (b)—at the asset-class level. Panel (a) shows that disagreement in each asset class's correlations with all other classes is modest, with mean absolute deviations of about 0.05 for U.S. cash and 0.13 for Global Bonds ex-US. Comparing Panels (a) and (b) shows that disagreement and deviation from history are related but distinct: deviations from realized history are slightly larger, with the mean absolute deviation largest for Global Bonds ex-US (0.20) and smallest for U.S. cash (0.07).

[Figure 17 here]

Figure 18 provides a more granular decomposition of the mean absolute deviation of subjective correlations from the leave-one-out consensus—Panel (a)—and from realized history—Panel (b)—at both the manager and asset-class-pair level. The heatmaps show that deviations from both consensus and history are present but limited: across managers, average deviations range from 0.04 to 0.21.

[Figure 18 here]

Figure 19 reports the magnitude of typical year-over-year revisions in return, volatility, and correlation beliefs. We document that return expectations are revised more from year to year than volatility expectations: the average absolute year-over-year revision is about 0.65 percentage points for return expectations, compared with about 0.25 percentage points for volatility expectations. Correlation beliefs move by about 0.02 correlation units per year on average. Because these quantities are measured in different units, we also scale revisions by the corresponding annual consensus average. On this basis, the average annual revision is roughly 11% for return expectations, compared with 4% for correlations and 2% for volatilities.

[Figure 19 here]

Taken together, the evidence suggests that first moments vary substantially more over time and are based on economic and fundamental intuition, while second moments appear to be more backward-looking and anchored in historical estimates.

3 Text-Based Evidence

Building on the framework-based decomposition of return expectations, we next examine the mechanisms through which institutional investors process information and form beliefs. We proceed in four steps, followed by a validation exercise using shareholder-letter narratives. First, we document who produces CMA reports and examine heterogeneity in the assumptions and methods used to construct the inputs to asset managers' forecasting frameworks. We then relate this heterogeneity to disagreement across managers. Second, we study the broader narrative environment in which these assumptions are embedded: using text analysis and topic modeling, we identify recurring themes in how managers describe economic and financial conditions. Third, we examine the structure of managers' reasoning by using large language models to extract the causal networks embedded in these reports, revealing how managers explicitly link economic phenomena to return expectations. Finally, we connect these qualitative mechanisms to quantitative belief formation by studying how institutional return expectations respond to earnings-news sentiment, and how this response varies with managers' causal-network characteristics, particularly their attention to specific economic channels and the complexity of their causal reasoning. This allows us to capture not only how managers incorporate information flows from earnings signals, but also how their narrative structures condition the translation of those signals into return expectations. We then validate the economic content of CMA narratives by comparing them with the shareholder-letter outlooks that affiliated fund managers provide in N-CSR filings.

3.1 CMA Assumptions, Models, and Authors

Having shown in Section 2 that asset managers' equity return expectations can be mapped into a common set of building blocks, we now examine the construction of these forecasts. We begin by studying the assumptions and methods described in the reports, with the goal of understanding how modeling choices map into the numerical forecasts that managers publish. We then turn to the individuals who produce these forecasts, documenting heterogeneity in CMA authors' professional profiles and examining its relationship with disagreement across asset managers.

This exercise is motivated by the idea that professional forecasts are shaped not only by the formal decomposition being used, but also by the assumptions, anchors, and methods through which each component is constructed. Because these choices are made within investment teams, the professional backgrounds of the individuals involved may also matter for how forecasts are formed. CMA reports often describe whether assumptions are based on historical averages, explicit mean-reversion arguments, trend assumptions, simulations, or regression-based approaches. They also provide textual explanations of the assumptions underlying individual building blocks, such as growth, inflation, dividend yield, and valuation change. Together, these methodological choices and team characteristics provide a window into the forecasting process behind the published numbers.

Declared Forecasting Techniques. We use LLM-based text classification to extract and construct methodology indicators from the CMA reports. Specifically, we first classify the report text and then define manager-year methodology indicators equal to one if at least one report by asset manager i in year t is classified as disclosing the corresponding technique.[13] Let:

$$\mathcal{K} = \{Reg, Sim, Avg, MR, Trend, Oth\}$$

denote the set of disclosed forecasting-technique indicators. For each $k \in \mathcal{K}$, we write

$$\mathbf{1}\{\text{technique}_{k,i,t}\}$$

for the corresponding manager-year indicator. The first five indicators capture whether manager i discloses, in year t, the use of explicit regressions (Reg), simulations (Sim), historical or long-run averages (Avg), explicit mean-reversion assumptions (MR), or trend assumptions for fundamentals ($Trend$), such as trend GDP, trend earnings, or trend productivity. The residual indicator (Oth) equals one when none of the five named techniques is disclosed. Because the indicators are extracted from the report text using LLM-based classification and are non-exclusive, a single manager-year can be associated with more than one technique. When none of the five named techniques is disclosed, we treat the corresponding assumptions as "exogenously specified" inputs in a disclosure-based sense: the report provides the numerical assumptions but does not disclose the method used to construct them.

Figure 20 reports the prevalence of forecasting techniques across CMA reports. Histor-

[13] We extract the technique indicators with a simple LLM prompt that asks the model to fill a fixed JSON schema with five binary indicators for the techniques used in the U.S. equity forecast: regressions, simulations, historical averages, mean reversion, and a trend hypothesis. We provide a definition for each indicator in the prompt. We do not supply labeled examples, we set the temperature to zero and we ask the model to base each indicator solely on text contained in the report.

ical averages, trend hypotheses, and mean-reversion arguments are the most widely used methods, whereas explicit regressions and simulations appear in a smaller share of reports. This pattern suggests that CMA forecasts are often grounded in historical benchmarks and trend-based narratives, including both extrapolative and contrarian views, while econometric or simulation-based approaches remain less common.

[Figure 20 here]

We then ask whether these technique choices are associated with where managers stand relative to the peer consensus in their U.S. equity forecasts and building-block assumptions. We consider the headline U.S. equity return forecast and its main reported components: ER denotes the headline U.S. equity return forecast, i.e., the total return expectation for U.S. equity reported by the asset manager before decomposing it into building blocks; g denotes real growth, π inflation, DY dividend yield, and VC valuation change. For the headline return expectation and each forecast component, $b \in \{ER, g, \pi, DY, VC\}$, we estimate:

$$Dev^b_{i,t} = \alpha^b + \tau^b_t + \sum_{k \in \mathcal{K}} \beta^b_k \mathbf{1}\{\text{technique}_{k,i,t}\} + \varepsilon^b_{i,t} \tag{36}$$

The year fixed effects τ^b_t absorb common macro-financial conditions, so the coefficients compare managers within the same year. We use this as the baseline specification because forecasting techniques are often persistent within asset managers, so manager fixed effects may absorb much of the relevant between-manager variation.

Table 11 reports the results. The coefficients indicate whether a disclosed technique is associated with a manager's forecast standing above or below the same-year peer consensus. The strongest patterns are concentrated in the valuation-change component. Mean-reversion assumptions are associated with headline U.S. equity forecasts below the peer consensus in the full sample and with significantly more negative valuation-change assumptions. This pattern has a natural economic interpretation within the building-block decomposition. To the extent that U.S. equity valuation ratios are elevated relative to historical benchmarks during our sample period, mean-reversion arguments imply a negative contribution from valuation change, since part of future returns is expected to be reduced by the normalization of valuation multiples.

[Table 11 here]

Overall, the table shows that disclosed forecasting techniques are informative about where managers stand relative to their peers. The clearest evidence comes from valuation-related assumptions: the mean reversion hypothesis is associated with return forecasts and valuation-change forecasts below the peer consensus.

Authors. Appendix A.5 shows that CMA forecasts are not anonymous institutional outputs, but are produced by identifiable teams with systematic professional characteristics. The typical CMA author is a senior and experienced investment professional, and most authors hold either graduate or MBA degrees. At the same time, there is meaningful heterogeneity in educational backgrounds across authoring teams and asset managers. This variation allows us to ask whether the human capital of the forecasting team is systematically related to the assumptions embedded in published CMA forecasts.

We focus on educational composition as one observable dimension of team human capital. For each manager i, year t, and U.S. equity component $b \in \{ER, g, \pi, g, DY, VC\}$, we look at how the manager's deviation from the contemporaneous leave-one-manager-out consensus relate to the number of authors on the report whose highest degree is a Bachelor, Master/MBA, or PhD. Figure 21 reports the Pearson correlation between each education count and deviations from consensus across the six equity components: equity return, real growth, inflation, nominal growth, dividend yield, and valuation change. Because the education variables are counts, the figure captures both the type of educational background represented on the team and the intensity with which that background appears among credited authors. The main pattern is that PhD-trained authors are associated with a distinct set of deviations from consensus. Bachelor and Master/MBA counts are positively correlated with real- and nominal-growth deviations but negatively correlated with dividend-yield deviations. By contrast, PhD counts are positively correlated with deviations in real growth, nominal growth, inflation, and dividend yield, and negatively correlated with valuation-change deviations. The contrast is especially pronounced for dividend yield: reports with more Bachelor or Master/MBA authors tend to deviate below consensus, whereas reports with more PhD authors tend to deviate above consensus. This pattern suggests that educational composition, and in particular the presence of PhD-trained authors, is related to the way asset managers allocate return expectations across underlying building blocks.

[Figure 21 here]

Modeling-Technique Clusters and US Equity Return Expectations. The five methodology indicators above capture broad categories of forecasting techniques. We now move to a more granular analysis of the actual modeling descriptions used in CMA reports. Rather than coding only whether a report mentions regressions, simulations, averages, mean reversion, or trend hypotheses, we use the full set of modeling-technique phrases extracted from the equity sections of CMA reports. We convert each phrase into a numerical text embedding, so that phrases with similar meanings are located close to one another in the embedding space, and then cluster these phrase-level embeddings into seven modeling-technique groups. [14]

Figure 22 visualizes this embedding space by projecting the high-dimensional phrase embeddings into two dimensions. Each point is one modeling-technique phrase, and nearby points correspond to phrases with similar economic content. The labels summarize the main technique represented in each region of the plot. Table 12 defines the seven clusters and reports their prevalence across reports. Let:

$$\mathcal{C} = \{Dec, Hist, Scen, PV, MR, Macro, YS\}$$

[14] The seven clusters above are obtained in two stages. First, a separate LLM prompt asks the model to fill a fixed JSON schema with the list of analytical techniques used in the U.S. equity forecast, each entry consisting of a free-text technique name and a 1–10 complexity score whose levels are anchored in the prompt (1–2 simple historical averages; 3–4 trend extrapolation and basic regression; 5–6 multi-factor regression, mean-reversion and building-block approaches; 7–8 Monte Carlo, VAR, regime-switching and Bayesian estimation; 9–10 DSGE, machine learning, Kalman filtering). The prompt provides only these anchors—no closed list of admissible techniques and no labeled examples—temperature is set to zero, and the model is instructed to base each entry solely on text contained in the report. This yields 5,090 technique entries across 589 reports, which after normalisation (lowercase, collapsed whitespace, stripped trailing punctuation, complexity prefix removed) reduce to 4,145 distinct phrases. Second, the phrases are embedded with the sentence-transformer `all-mpnet-base-v2`, L_2-normalised, and partitioned with k-means; we set $k = 7$ on the basis of interpretability: smaller values of k merged methodologically distinct techniques into single clusters, while larger values produced clusters that no longer admitted a coherent methodological label when we inspected them. We retain the model's raw geometry (no per-dimension standardisation). We then label each cluster by inspecting all phrases assigned to it and choosing a concrete methodological label that captures the common analytical approach and distinguishes the cluster from the others.

denote the set of modeling-technique clusters, corresponding respectively to Decomposition, Historical calibration, Scenario adjustment, Present value framework, Mean-reversion, Macro, and Yield-spread. The most common cluster is "Decomposition", which captures general return decompositions and appears in 83.2% of reports. "Historical calibration" is also widespread, appearing in 71.5% of reports, and captures calibration on historical averages, volatilities, and correlations.

[Figure 22 and Table 12 here]

We then code each manager-year using seven non-mutually-exclusive indicators, one for each cluster. For each $k \in \mathcal{C}$, we write:

$$\mathbf{1}\{\text{cluster}_{k,i,t}\}$$

for the corresponding manager-year cluster indicator. The indicator $\mathbf{1}\{\text{cluster}_{k,i,t}\}$ equals one if at least one report by asset manager i in year t uses a modeling phrase assigned to cluster k, and zero otherwise. The analysis is conducted on manager-year observations that publish a building-block decomposition of the forecast, so that both the numerical forecast and the disclosed modeling approach are observed. Because reports can use multiple technique clusters, the coefficients should be read as conditional associations: they ask whether a given cluster is associated with a manager's forecast or building-block assumption, holding fixed the other disclosed clusters and year effects.

Table 13 relates the technique clusters to managers' positions relative to their peers. For the headline return expectation and each forecast component, $b \in \{ER, g, DY, VC, \pi\}$, we estimate:

$$Dev^b_{i,t} = \alpha^b + \tau^b_t + \sum_{k \in \mathcal{C}} \beta^b_k \mathbf{1}\{\text{cluster}_{k,i,t}\} + \varepsilon^b_{i,t}$$

The year fixed effects τ^b_t absorb common macro-financial conditions, so the coefficients compare managers within the same year. The main pattern is concentrated in the "Mean-reversion" cluster, which captures mean reversion of valuation ratios such as P/E or CAPE toward historical norms. Managers whose reports disclose "Mean-reversion" techniques report headline U.S. equity forecasts and valuation-change assumptions below the peer consensus. "Historical calibration" points in the opposite direction: managers using historically calibrated inputs report forecasts above the peer consensus, with higher dividend-yield and valuation-change components. Thus, the phrase-level taxonomy confirms that disagreement across managers is linked to specific modeling choices, especially the contrast between mean-reversion of valuation ratios and historical calibration of inputs.

[Table 13 here]

Table 14 asks whether the same technique clusters are associated with deviations from the objective machine-learning benchmark. The dependent variable is the manager's headline U.S. equity return forecast minus the objective benchmark forecast, matched at the same reference date and horizon. We estimate:

$$F_{i,t,h}^{ER} - G_{t,h} = \alpha + \tau_t + \sum_{k \in \mathcal{C}} \beta_k \mathbf{1}\{\text{cluster}_{k,i,t}\} + \varepsilon_{i,t,h}$$

As in the analysis of deviations from the peer consensus, the same contrast appears relative to the objective benchmark: Mean-reversion is associated with lower forecasts, while Historical calibration is associated with higher forecasts. This shows that the technique clusters are informative not only about peer disagreement, but also about systematic deviations from the objective benchmark.

[Table 14 here]

3.2 Topic Text Analysis

To understand the narrative structures underlying institutional return expectations, we analyze the thematic content of CMA reports using topic modeling techniques. Following the methodology detailed in Appendix A.1, we segment CMA reports to isolate discussions of equity returns and related macroeconomic drivers, then apply Latent Dirichlet Allocation (LDA) to extract latent topics from these equity-focused narratives. This approach builds on the narrative asset pricing framework of Bybee et al. (2023), which shows that textual narratives contain systematic risk factors beyond traditional quantitative measures. The choice of 55 topics, selected within a coherence plateau, balances model complexity with interpretability while capturing the primary narrative dimensions in our corpus (Appendix A.1.2).

Figure 23 visualizes the resulting narrative landscape across institutional investors. Each CMA report is represented as a 55-dimensional vector of topic proportions, which we project into two dimensions using t-SNE to preserve local similarities in the high-dimensional topic space. The application of k-means clustering to these topic distributions reveals 14 distinct narrative groups, suggesting that asset managers organize their economic reasoning around coherent thematic frameworks. The separation between clusters is consistent with systematic differences in how institutions frame market dynamics and structure their investment narratives.

[Figure 23 here]

3.3 Causal Networks

Asset managers explicitly articulate causal mechanisms in their CMA reports, documenting how they believe economic variables transmit to equity returns. Using the LLM methodology from Section 1.3, we extract these stated causal relationships to construct directed networks that summarize each institution's economic reasoning. Unlike correlation analysis, these networks capture managers' stated beliefs about transmission channels and causal ordering—distinguishing between primary drivers, intermediate mechanisms, and final outcomes in their mental models of market dynamics.

Visualization. Figure 24 illustrates this approach through DWS's 2015–2025 causal network, where node size reflects a topic's interconnectedness and edge colors encode directional effects. The network shows that growth and valuation change emerge as the most central topics, with multiple causal pathways connecting policy and regulations, interest rates, and dividend yield. Figure 25 isolates first-order drivers of U.S. equity return expectations—concepts directly linked through a single causal step—highlighting how DWS emphasizes valuation adjustments, growth, dividend yield, and inflation. This structural mapping enables systematic comparison across institutions to identify where sophisticated investors differ in their views of market mechanics. In both networks, growth, inflation, dividends, and valuation adjustments—the four building blocks that our decomposition analysis (Section 2.1) identified as primary sources of forecast disagreement—appear with distinct coloring to facilitate the identification of their causal pathways. Appendix A.10.1 presents additional network visualizations for four asset managers and reports the number of topics and causal links per report for all institutions in our sample.

[Figures 24 and 25 here]

Figure 26 presents two complementary perspectives on the causal networks. Panel (a) displays how frequently economic topics appear as causal agents across all relationships, with growth, valuation change, and policy/regulatory factors dominating managers' reasoning frameworks. Panel (b) isolates topics directly linked to U.S. equity return expectations, revealing a concentrated set of drivers: valuation change and economic growth appear in 59% and 44% of asset manager-years, respectively.

[Figure 26 here]

Figure 27 documents substantial heterogeneity in asset managers' causal frameworks. The stacked bars show average causal links per year by topic, with marked differences across asset managers. JP Morgan and Robeco construct dense networks with more than 200 links per year, while others maintain smaller frameworks. The topics receiving the most links also vary: some managers emphasize growth and valuation, whereas others prioritize policy and inflation.

[Figure 27 here]

Attention, Bias, and Complexity. Having extracted causal networks from CMA narratives, we now examine whether the structure of these networks helps explain cross-sectional heterogeneity in institutional beliefs. Our analysis tests whether managers who assign different levels of complexity or attention to specific economic mechanisms—as revealed by their causal reasoning—produce systematically different return-forecast levels, forecast changes, and ex-ante forecast errors. We examine how institutional return expectations relate to the structure and content of managers' causal reasoning networks using panel regressions at the asset-manager-year level.

We begin with forecast levels. To investigate whether managers with more complex causal reasoning structures produce different return forecasts and whether this relationship varies with earnings sentiment, we employ a within-between decomposition of causal-network measures (Mundlak 1978, Chamberlain 1982). This approach separates temporary deviations from a manager's typical complexity level, or within-manager variation, from persistent differences in complexity across managers, or between-manager variation. For each complexity measure m, let $C_{i,t}^m$ denote the value of that measure for asset manager i in year t, and let:

$$\overline{C}_i^m = \frac{1}{T_i} \sum_{t \in \mathcal{T}_i} C_{i,t}^m$$

denote manager i's time-series average complexity. We estimate the following specification:

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_W^m \left(C_{i,t}^m - \overline{C}_i^m \right) + \gamma_W^m \left(C_{i,t}^m - \overline{C}_i^m \right) \times S_t + \gamma_B^m \overline{C}_i^m \times S_t + \varepsilon_{i,t}^m \tag{37}$$

where $F_{i,t}^{ER}$ denotes asset manager i's headline U.S. equity return forecast in year t, with the horizon suppressed because the empirical specification is estimated on the matched asset-manager-year panel. The terms μ_i and τ_t represent asset manager and year fixed effects. The term $\left(C_{i,t}^m - \overline{C}_i^m \right)$ captures within-manager variation in complexity, while $\overline{C}_i^m$ captures the between-manager component. The variable S_t is the aggregate sentiment measure, constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. The coefficient β_W^m captures the association between temporary increases in causal complexity and return forecasts at average sentiment levels. The coefficient γ_W^m captures whether a manager's news sensitivity changes when its network becomes temporarily more complex, while γ_B^m captures whether managers with persistently more complex networks react differently to news. Note that the between-manager component $\overline{C}_i^m$ is absorbed by the asset manager fixed effects μ_i and therefore appears only in interaction with sentiment. This specification allows us to estimate the role of temporary complexity changes within managers while controlling for time-invariant manager characteristics and common time trends. We estimate the specification separately for three complexity measures: average shortest path length, indirect connection ratio, and network transitivity (see Appendix A.10.2 for definitions of all network measures).

Table 15 shows that causal-network complexity is not directly associated with the level of return expectations: the coefficients on the within component $C^m_{i,t} - \overline{C}^m_i$ are small and statistically insignificant across all three complexity measures. Instead, complexity matters through its interaction with news sentiment. The interaction between temporary complexity and sentiment, $(C^m_{i,t} - \overline{C}^m_i) \times S_t$, is weakly significant only when complexity is measured by network transitivity. Conversely, the interaction between structural, between-manager complexity and sentiment, $\overline{C}^m_i \times S_t$, is negative and statistically significant across all three measures. This suggests that managers with persistently more complex causal networks issue lower return forecasts in response to positive earnings-news sentiment. This heterogeneity supports the view that differences in narrative structure help explain cross-manager variation in long-horizon return expectations.

[Table 15 here]

The distinction between chain and interconnected complexity helps interpret this pattern. Average path length and the indirect connection ratio capture chain complexity, or the extent to which managers reason through multi-step causal mechanisms. Network transitivity captures interconnected complexity, or the extent to which narratives contain locally connected causal structures. The results suggest that neither dimension directly shifts return expectations in normal times. Instead, both dimensions appear to become relevant when managers process new information, consistent with narrative structure shaping the mapping from news into long-horizon forecasts.

We also examine whether managers' attention to specific topics is associated with return expectations. Topic attention is measured by the edge count—the normalized number of causal connections to each topic node—thus capturing the role of that topic as a transmission mechanism in the manager's economic reasoning. For each topic q, let $A^q_{i,t}$ denote asset manager i's topic-attention measure in year t, and let:

$$\overline{A}^q_i = \frac{1}{T_i} \sum_{t \in \mathcal{T}_i} A^q_{i,t}$$

denote manager i's time-series average attention to topic q. For this analysis, we estimate:

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) + \gamma_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) \times S_t + \gamma_B^q \overline{A}_i^q \times S_t + \varepsilon_{i,t}^q \tag{38}$$

The term $\left(A_{i,t}^q - \overline{A}_i^q\right)$ captures within-manager variation in attention to topic q, while $\overline{A}_i^q$ captures persistent between-manager differences in attention to that topic. We examine five topics: valuation change, dividend yield, economic growth, inflation, and economic downturn. The coefficient β_W^q captures the association between topic attention and forecasts at average sentiment levels. The coefficient γ_W^q captures how within-manager variation in topic attention moderates sensitivity to aggregate sentiment, while γ_B^q captures the between-manager interaction with sentiment. Note that the between-manager component $\overline{A}_i^q$ is absorbed by the asset manager fixed effects μ_i and therefore appears only in interaction with sentiment.

Table 16 shows that topic-specific attention in managers' causal networks is associated with return forecasts primarily through its interaction with news sentiment rather than through direct level effects. The within component $A_{i,t}^q - \overline{A}_i^q$ is small and statistically insignificant for most topics, with the exception of economic downturn attention, which is negative and weakly significant. Thus, temporary increases in attention to a given mechanism generally do not directly shift long-horizon expectations.

In contrast, the interaction terms $(A_{i,t}^q - \overline{A}_i^q) \times S_t$ reveal economically meaningful heterogeneity. When managers temporarily devote more attention to valuation adjustments, their return forecasts respond more negatively to positive sentiment. By contrast, temporary attention to economic downturn narratives is associated with a more positive response to sentiment. Between-manager components $\overline{A}_i^q \times S_t$ display similarly heterogeneous signs: managers whose narratives structurally emphasize valuation change react more negatively to positive sentiment, whereas those who place more persistent emphasis on dividend yield or economic downturn narratives react more positively.

Overall, Table 16 shows that the content of a manager's causal narrative—not only its complexity—shapes how sentiment is incorporated into long-horizon return forecasts, with different economic mechanisms amplifying or dampening news sensitivity in systematic ways.

[Table 16 here]

We next examine whether the structural complexity of managers' causal networks is associated with how beliefs change over time. Complex mental models may either amplify reactions to news, as belief updates propagate through interconnected concepts, or dampen them if complexity reflects more stable and well-anchored reasoning. We test this by analyzing how network structure predicts year-over-year forecast changes.

We employ three complementary complexity measures: average shortest path length, indirect connection ratio, and network transitivity. For each complexity measure m, we estimate

$$F_{i,t}^{ER} - F_{i,t-1}^{ER} = \mu_i + \tau_t + \beta_W^m \left(C_{i,t}^m - \overline{C}_i^m\right) + \gamma_W^m \left(C_{i,t}^m - \overline{C}_i^m\right) \times S_t + \gamma_B^m \overline{C}_i^m \times S_t + \varepsilon_{i,t}^m \quad (39)$$

where $F_{i,t}^{ER} - F_{i,t-1}^{ER}$ is the year-over-year change in asset manager i's headline U.S. equity return forecast. The coefficient β_W^m captures the association between temporary complexity changes and forecast changes, while γ_W^m and γ_B^m capture differential news sensitivity.

Table 17 examines whether managers' causal-network complexity predicts year-over-year forecast changes. Across all three complexity measures, the within component $C_{i,t}^m - \overline{C}_i^m$ is small and statistically insignificant, indicating that temporary increases in network complexity do not mechanically generate larger forecast adjustments. By contrast, the between-sentiment interaction $\overline{C}_i^m \times S_t$ is negative, economically large, and statistically significant across all three measures. Managers with persistently more complex causal networks revise their long-horizon forecasts more negatively in response to positive earnings-news sentiment. Overall, the results indicate that persistent differences in narrative complexity, rather than temporary changes in the annual narrative, are associated with how managers update beliefs in response to news.

[Table 17 here]

The distinction between chain and interconnected complexity helps interpret this result. Higher chain complexity means that news must travel through a larger number of intermediate mechanisms before reaching return expectations, while higher interconnected complexity means that the same news is embedded in a denser set of mutually related causal channels. The negative interaction between sentiment and the persistent components of all three complexity measures therefore suggests that both longer causal chains and more interconnected narratives dampen managers' upward revisions after positive earnings news.

We separately examine whether attention to specific economic topics is associated with forecast changes. For each topic q, we estimate:

$$F_{i,t}^{ER} - F_{i,t-1}^{ER} = \mu_i + \tau_t + \beta_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) + \gamma_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) \times S_t + \gamma_B^q \overline{A}_i^q \times S_t + \varepsilon_{i,t}^q \quad (40)$$

We examine five topics: valuation change, dividend yield, economic growth, inflation, and economic downturn.

Table 18 studies whether managers' attention to specific economic mechanisms—as captured by topic-specific causal-network edge counts—helps explain year-over-year changes in long-horizon return expectations. Across all topics, the within component $A_{i,t}^q - \overline{A}_i^q$ is small and statistically insignificant, indicating that temporary shifts in attention do not mechanically induce forecast changes.

Forecast changes instead respond to news conditional on topic attention. The interaction terms $(A_{i,t}^q - \overline{A}_i^q) \times S_t$ are significant for valuation change, dividend yield, and growth, whereas the between-manager interactions $\overline{A}_i^q \times S_t$ are positive and statistically significant for dividend yield and economic downturn. Managers who temporarily devote more attention to valuation change exhibit a negative change in response to positive news, whereas those who place greater-than-usual attention on dividend yield and growth show a positive change to the same news. In addition, managers who are persistently more attentive than others to dividend yield and economic downturn narratives react more positively to earnings-sentiment news. The results suggest that forecast changes are associated not only with how a manager's attention to different topics varies over time, but also with how their attention compares with that of other managers.

Overall, the table shows that the content of causal narratives helps explain how managers update long-horizon expectations. Topic emphasis—both temporary and structural—modulates the responsiveness of forecast changes to fundamental news, creating systematic and persistent heterogeneity in belief updating across institutions.

[Table 18 here]

To investigate whether network complexity predicts systematic deviations from objective forecasts, we employ a within-between decomposition. For each complexity measure m, we estimate:

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) + \gamma^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) \times S_t + \gamma^m_B \overline{C}^m_i \times S_t + \varepsilon^m_{i,t,h} \quad (41)$$

The coefficient β^m_W captures whether temporary changes in complexity are associated with more optimistic or pessimistic forecasts relative to the benchmark. The interaction coefficients γ^m_W and γ^m_B capture whether complexity moderates how news sentiment affects forecast bias, with γ^m_B specifically capturing whether managers with structurally complex networks exhibit different responses to news. We examine three complexity measures: average shortest path length, indirect connection ratio, and network transitivity.

Table 19 examines whether the structural complexity of managers' causal networks predicts ex-ante forecast errors relative to the objective benchmark. The within-manager component $C^m_{i,t} - \overline{C}^m_i$ is statistically insignificant across all three complexity measures, indicating that temporary changes in narrative complexity do not predict forecast errors. Instead, persistent differences in complexity across managers are associated with how they process information.

The between-manager interaction with sentiment, $\overline{C}^m_i \times S_t$, is positive and highly significant across all measures. Since the dependent variable is $G_{t,h} - F^{ER}_{i,t,h}$, this implies that, following positive earnings-news sentiment, managers with structurally more complex causal networks issue forecasts that are lower relative to the objective benchmark. In other words, structurally complex managers underreact more to positive information flows, generating systematically more pessimistic forecasts relative to the benchmark. During periods of favorable sentiment, these managers produce larger benchmark-minus-forecast errors than managers with simpler causal structures.

[Table 19 here]

The distinction between chain and interconnected complexity clarifies the nature of these forecast errors. For chain complexity, the results suggest that multi-step reasoning may attenuate the pass-through from favorable fundamentals to return forecasts. For interconnected complexity, the results suggest that densely connected narratives may introduce additional countervailing channels, causing positive news to be filtered through a broader system of mechanisms before affecting expectations. Thus, both longer causal chains and greater local interconnection are associated with more conservative belief updating after positive news.

We separately examine whether attention to specific economic topics predicts systematic deviations from the objective benchmark. For each topic q, we estimate:

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) + \gamma^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) \times S_t + \gamma^q_B \overline{A}^q_i \times S_t + \varepsilon^q_{i,t,h} \qquad (42)$$

The coefficient β^q_W captures whether temporary increases in topic-specific attention are associated with systematic optimism or pessimism relative to the objective benchmark, while γ^q_W and γ^q_B capture how topic attention moderates the influence of news sentiment on forecast bias through within- and between-manager variation, respectively. We examine five economic topics: valuation change, dividend yield, economic growth, inflation, and economic downturn.

Table 20 examines whether topic-specific attention in managers' causal networks predicts ex-ante forecast errors relative to the objective benchmark. The results show that attention to specific topics systematically predicts forecast errors by shaping how asset managers incorporate information flows.

Asset managers with greater attention to valuation change in their causal networks produce larger benchmark-minus-forecast errors following positive fundamental news, generating systematically more pessimistic forecasts relative to the objective benchmark. Managers who, on average, emphasize valuation change—the between-manager component $\overline{A}^q_i \times S_t$—exhibit this pattern persistently and significantly. When individual managers temporarily increase their attention to valuation change—the within-manager component $(A^q_{i,t} - \overline{A}^q_i) \times S_t$—the effect operates in the same direction. Since the dependent variable is $G_{t,h} - F^{ER}_{i,t,h}$, the positive valuation-change interactions imply that higher valuation attention is associated with underreaction to positive news relative to the objective benchmark.

Asset managers with greater attention to dividend yield and economic downturn narratives produce the opposite pattern. The between-manager interactions with sentiment are negative and highly significant for dividend yield and economic downturn, while the within-manager interaction is also negative and significant for economic downturn. The results show that topic-specific attention modulates news sensitivity: managers emphasizing valuation adjustments underreact to positive news, while those emphasizing dividend yield and downturn risks issue forecasts that move more strongly with positive news relative to the benchmark, introducing systematic and predictable deviations from the objective benchmark.

[Table 20 here]

3.4 Relevance: N-CSR Letter Outlook

To assess whether asset managers' CMA reports reflect professional investment views, we compare the language of each manager's CMAs with the language used in its N-CSR shareholder letters. We encode every N-CSR letter outlook and every CMA paragraph discussing equity with a pretrained sentence-transformer. For each letter, similarity to a given CMA report is computed as the mean cosine similarity across the five best-matching equity-related paragraphs of that report, restricted to CMA reports filed by the relevant manager within a six month window. We then partition the resulting pairs into two populations: same-manager pairs, where the letter and CMA share the asset manager, and cross-manager pairs, where they do not. The analysis is restricted to asset managers contributing at least five unique letters.

Figure 28 shows that the same-manager distribution stochastically dominates the cross-manager distribution across the support of cosine similarity. A Mann–Whitney test rejects equality of the two distributions, and a paired test at the manager level—comparing, for each manager, its median same-manager similarity to its median cross-manager similarity—confirms that the majority of managers exhibit a positive gap.

[Figure 28 here]

These results are consistent with CMAs capturing asset managers' professional investment views. The same-manager similarity premium indicates that the narratives embedded in CMAs also appear in N-CSR outlook letters, suggesting that CMAs reflect a persistent manager-specific house view.

4 Conclusion

We study how large asset managers form and justify long-horizon beliefs by combining the quantitative and narrative content of their Capital Market Assumptions (CMAs). Using hand-collected data, we map equity return expectations to a common set of building blocks: real growth, valuation change, income, and inflation. Expected valuation change explains the largest share of cross-sectional variation in return expectations, but growth and the other components jointly account for the majority of the remaining variation. Return expectations are countercyclical mainly because valuation-change expectations decline with valuation ratios, while growth expectations are procyclical. At the same time, the subjective relationships embedded in managers' decompositions differ from their realized counterparts. Real growth and valuation change are negatively related in managers' expectations but positively related in historical data, while inflation and valuation change display the opposite sign pattern. These differences suggest that managers' subjective models may depart from the empirical relationships observed in realized data, motivating our benchmark-based analysis of systematic deviations in CMA forecasts. The same component-level beliefs also matter for investment decisions: growth and valuation-change expectations are the components most strongly linked to equity allocations.

We show that disagreement across asset managers reflects not only different numerical inputs, but also different modeling anchors. Although managers often rely on similar return-decomposition frameworks, they differ in the assumptions used to populate them. Historical-calibration approaches are associated with more optimistic forecasts, while mean-reversion approaches are associated with lower forecasts and more negative valuation-change assumptions, both relative to peers and to the objective benchmark.

We then compare CMA forecasts with an objective machine-learning benchmark and study how managers update beliefs over time. The benchmark systematically outperforms CMA forecasts, and the comparison suggests that managers place greater weight on macro-financial variables, especially inflation, and less weight on valuation ratios than the benchmark does. Forecast dynamics point to a consensus anchor: managers revise toward the cross-sectional consensus and deviate from the benchmark in the direction of recent peer forecasts. Managers also respond to earnings-news sentiment by increasing return and valuation-change expectations. By contrast, volatility expectations are highly conservative, dominated by manager-specific frameworks, and adjust only slowly to realized volatility. Correlation beliefs display a similar second-moment pattern. Managers' pairwise correlation assumptions exhibit limited disagreement and small year-to-year revisions, and they remain closely related to realized historical correlations for equity-equity and equity-bond pairs. These results highlight a sharp contrast between first moments, which vary substantially and respond to economic information, and second moments, which appear more backward-looking and history-anchored.

Finally, we show that narratives help explain these patterns by revealing how managers organize the causal structure of their beliefs. We develop a new empirical framework that uses an LLM-based pipeline to extract directed and signed causal links from CMA text and map the connections between macro variables and the building-block drivers of return expectations. These networks reveal substantial heterogeneity in the complexity and topic attention of managers' reasoning, allowing us to study heterogeneity in the structure of belief formation rather than only heterogeneity in topics, sentiment, or textual similarity. Managers with structurally more complex narratives exhibit weaker upward forecast changes after positive earnings-news sentiment and more pessimistic forecasts relative to objective benchmarks. Topic attention provides an additional source of heterogeneity: attention to valuation-change mechanisms is associated with underreaction to positive news, while attention to dividend-yield and downturn narratives is associated with forecasts that move more strongly with positive news. Comparisons with N-CSR shareholder letters further show that same-institution letter–CMA pairs are more similar than cross-institution pairs, indicating that CMA narratives capture persistent institution-specific investment views.

Taken together, our results show that institutional beliefs are shaped by the interaction between the numerical building blocks of return expectations, the assumptions used to construct those blocks, and the narratives through which managers organize and communicate their views. Reading CMAs in both numbers and words helps explain why sophisticated institutions disagree, why their forecasts depart predictably from objective benchmarks, how they update their beliefs, and why those beliefs matter for portfolio allocations. It also shows that different objects of belief behave differently: return expectations embed forward-looking economic reasoning, whereas volatility and correlation assumptions are more stable, less dispersed, and more closely tied to historical realizations.

References

Adam, K. & Nagel, S. (2023), Expectations data in asset pricing, *in* 'Handbook of Economic Expectations', Elsevier, pp. 477–506.

Andonov, A., Couts, S. J., S Gonçalves, A., Loudis, J. & Rossi, A. (2026), 'Subjective beliefs and the portfolio allocations of institutional investors', *Working Paper* .

Andre, P., Pizzinelli, C., Roth, C. & Wohlfart, J. (2022), 'Subjective models of the macroeconomy: Evidence from experts and representative samples', *The Review of Economic Studies* **89**(6), 2958–2991.

Asquith, P., Mikhail, M. B. & Au, A. S. (2005), 'Information content of equity analyst reports', *Journal of financial economics* **75**(2), 245–282.

Baker, S., Bloom, N., Davis, S. J. & Sammon, M. C. (2021), What triggers stock market jumps?, Technical report, National Bureau of Economic Research Cambridge, MA.

Barberis, N. (2018), Psychology-based models of asset prices and trading volume, *in* 'Handbook of behavioral economics: applications and foundations 1', Vol. 1, Elsevier, pp. 79–175.

Bastianello, F. (2024), 'Time-series and cross-section of risk premia expectations: Evidence from financial analysts', *Available at SSRN 4204968* .

Bastianello, F. (2025), 'Bias and predictability in analysts' beliefs', *Available at SSRN 5733302* .

Bastianello, F., H Décaire, P. & Guenzel, M. (2024), 'Mental models and financial forecasts', *Marius, Mental Models and Financial Forecasts (October 30, 2024)* .

Bastianello, F. & Imas, A. (2025), 'Biases in belief updating within and across domains', *Available at SSRN 5413920* .

Bastianello, F. & Peng, C. (2024), 'Global fund managers' beliefs, perceived mispricing, and asset allocation', *Perceived Mispricing, and Asset Allocation (December 15, 2024)* .

Bastianello, F. & Peng, C. (2025), 'The term structure of return expectations', *Available at SSRN 5390270* .

Begenau, J., Liang, P. & Siriwardane, E. (2025), 'The rise of alternatives', *Available at SSRN 4940886* .

Bianchi, F., Gómez-Cram, R. & Kung, H. (2024), 'Using social media to identify the effects of congressional viewpoints on asset prices', *The Review of Financial Studies* **37**(7), 2244–2272.

Bianchi, F., Lee, D. Q., Ludvigson, S. C. & Ma, S. (2025), The prestakes of stock market investing, Technical report, National Bureau of Economic Research.

Bianchi, F., Ludvigson, S. C. & Ma, S. (2022), 'Belief distortions and macroeconomic fluctuations', *American Economic Review* **112**(7), 2269–2315.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent dirichlet allocation', *Journal of Machine Learning Research* **3**, 993–1022.

Bordalo, P., Gennaioli, N., La Porta, R. & Shleifer, A. (2024*a*), 'Belief overreaction and stock market puzzles', *Journal of Political Economy* **132**(5), 1450–1484.

Bordalo, P., Gennaioli, N., La Porta, R. & Shleifer, A. (2024*b*), Finance without exotic risk, Technical report, National Bureau of Economic Research.

Bordalo, P., Gennaioli, N., Porta, R. L. & Shleifer, A. (2019), 'Diagnostic expectations and stock returns', *The Journal of Finance* **74**(6), 2839–2874.

Bordalo, P., Gennaioli, N. & Shleifer, A. (2022), 'Salience', *Annual Review of Economics* **14**(1), 521–544.

Bouchaud, J.-P., Krueger, P., Landier, A. & Thesmar, D. (2019), 'Sticky expectations and the profitability anomaly', *The Journal of Finance* **74**(2), 639–674.

Boudoukh, J., Michaely, R., Richardson, M. & Roberts, M. (2007), 'On the importance of measuring payout yield: Implications for empirical asset pricing', *Journal of Finance* **62**(2), 877–915.

Bretscher, L., Malkhozov, A., Tamoni, A. & Yang, H. (2024), 'Distorted beliefs and asset prices', *Available at SSRN 4947547* .

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. et al. (2020), 'Language models are few-shot learners', *Advances in Neural Information Processing Systems (NeurIPS)* **33**, 1877–1901.

Bybee, J. L. (2023), 'The ghost in the machine: Generating beliefs with large language models', *arXiv preprint arXiv:2305.02823* .

Bybee, L., Kelly, B., Manela, A. & Xiu, D. (2024), 'Business news and business cycles', *The Journal of Finance* **79**(5), 3105–3123.

Bybee, L., Kelly, B. & Su, Y. (2023), 'Narrative asset pricing: Interpretable systematic risk factors from news text', *The Review of Financial Studies* **36**(12), 4759–4787.

Campbell, J. Y. & Ammer, J. (1993), 'What moves the stock and bond markets? a variance decomposition for long-term asset returns', *The Journal of Finance* **48**(1), 3–37.

Campbell, J. Y. & Cochrane, J. H. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**(2), 205–251.

Campbell, J. Y. & Shiller, R. J. (1987), 'Cointegration and tests of present value models', *Journal of political economy* **95**(5), 1062–1088.

Campbell, J. Y. & Shiller, R. J. (1988), 'The dividend-price ratio and expectations of future dividends and discount factors', *Review of Financial Studies* **1**(3), 195–228.

Campbell, J. Y. & Thompson, S. B. (2008), 'Predicting excess stock returns out of sample: Can anything beat the historical average?', *Review of Financial Studies* **21**(4), 1509–1531.

Campbell, S. D. & Sharpe, S. A. (2009), 'Anchoring bias in consensus forecasts and its effect on market prices', *Journal of Financial and Quantitative Analysis* **44**(2), 369–390.

Chamberlain, G. (1982), 'Multivariate regression models for panel data', *Journal of Econometrics* **18**(1), 5–46.

Chen, Y., Kelly, B. & Xiu, D. (2024), 'Expected returns and large language models', *Working Paper* . University of Chicago and Yale University.

Cochrane, J. H. (2011), 'Presidential address: Discount rates', *The Journal of Finance* **66**(4), 1047–1108.

Couts, S. J., Gonçalves, A. S. & Loudis, J. (2023), 'The subjective risk and return expectations of institutional investors', *Fisher College of Business Working Paper* **14**.

Couts, S. J., S Gonçalves, A., Liu, Y. & Loudis, J. (2024), 'Institutional investors' subjective risk premia: Time variation and disagreement'.

Dahlquist, M. & Ibert, M. (2024), 'Equity return expectations and portfolios: Evidence from large asset managers', *The Review of Financial Studies* **37**(6), 1887–1928.

Dahlquist, M. & Ibert, M. (2025), 'Institutions' return expectations across assets and time'.

Dang, T. L., Moshirian, F. & Zhang, B. (2015), 'Commonality in news around the world', *Journal of Financial Economics* **116**(1), 82–110.

De la O, R. & Myers, S. (2021), 'Subjective cash flow and discount rate expectations', *The Journal of Finance* **76**(3), 1339–1387.

De la O, R. & Myers, S. (2024), 'Which subjective expectations explain asset prices?', *The Review of Financial Studies* **37**(6), 1929–1978.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**(3), 253–263.

Décaire, P. H. & Graham, J. R. (2024), 'Valuation fundamentals', *Working Paper* . Available at SSRN.

Fama, E. F. & French, K. R. (2002), 'The equity premium', *The Journal of Finance* **57**(2), 637–659.

Fu, Y., Peng, H., Sabharwal, A., Clark, P. & Khot, T. (2023), 'Complexity-based prompting for multi-step reasoning', *arXiv preprint arXiv:2210.00720* .

Gabaix, X. (2019), Behavioral inattention, *in* 'Handbook of behavioral economics: Applications and foundations 1', Vol. 2, Elsevier, pp. 261–343.

Gabaix, X. & Graeber, T. (2024), The complexity of economic decisions, Technical report, National Bureau of Economic Research.

Gao, A. K. (2023), 'Prompt engineering for large language models', *SSRN Electronic Journal* .

Gao, P., Parsons, C. A. & Shen, J. (2017), 'Global relation between financial distress and equity returns', *The Review of Financial Studies* **31**(1), 239–277.

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–574.

Giannone, D., Lenza, M. & Primiceri, G. E. (2021), 'Economic predictions with big data: The illusion of sparsity', *Econometrica* **89**(5), 2409–2437.

Giglio, S., Maggiori, M., Stroebel, J. & Utkus, S. (2021), 'Five facts about beliefs and portfolios', *American Economic Review* **111**(5), 1481–1522.

Goetzmann, W. N., Kim, D. & Shiller, R. J. (2022), Crash narratives, Technical report, National Bureau of Economic Research.

Gonçalves, D., Libgober, J. & Willis, J. (2025), 'Retractions: Updating from complex information', *Review of Economic Studies* p. rdaf032.

Gormsen, N. J. & Huber, K. (2025), 'Corporate discount rates', *American Economic Review* **115**(6), 2001–2049.

Goyal, A., Welch, I. & Zafirov, A. (2024), 'A comprehensive 2022 look at the empirical performance of equity premium prediction', *The Review of Financial Studies* .

Greenwood, R. & Shleifer, A. (2014), 'Expectations of returns and expected returns', *The Review of Financial Studies* **27**(3), 714–746.

Gu, S., Kelly, B. & Xiu, D. (2020), 'Empirical asset pricing via machine learning', *Review of Financial Studies* **33**(5), 2223–2273.

Hafez, P. (2009), 'Construction of market sentiment indices using news sentiment', *RavenPackInternationalS.L.* .

Israel, R., Kelly, B. & Moskowitz, T. (2020), 'Can machines "learn" finance?', *SSRN Working Paper* .

Ke, Z. T., Kelly, B. & Xiu, D. (2021), ‘Predicting returns with text data’, *Working Paper* . University of Chicago, Yale University, and Harvard University.

Kendall, C. W. & Charles, C. (2022), Causal narratives, Technical report, National Bureau of Economic Research.

Laarits, T., Matthies, B., Vasudevan, K. & Yang, W. (2025), ‘Policy by committee’.

Li, X., Chan, S., Zhu, X., Pei, Y., Ma, Z., Liu, X. & Shah, S. (2023), ‘Are chatgpt and gpt-4 general-purpose solvers for financial text analytics? a study on several typical tasks’, *arXiv preprint arXiv:2305.05862* .

Lopez-Lira, A. (2024), ‘Can chatgpt forecast stock price movements?’.

Malmendier, U. & Nagel, S. (2011), ‘Depression babies: Do macroeconomic experiences affect risk taking?’, *The quarterly journal of economics* **126**(1), 373–416.

Mundlak, Y. (1978), ‘On the pooling of time series and cross section data’, *Econometrica* **46**(1), 69–85.

Nagel, S. & Xu, Z. (2022), ‘Asset pricing with fading memory’, *The Review of Financial Studies* **35**(5), 2190–2245.

Nagel, S. & Xu, Z. (2023), ‘Dynamics of subjective risk premia’, *Journal of Financial Economics* **150**(2), 103713.

Newey, W. K. & West, K. D. (1987), ‘A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix’, *Econometrica* **55**(3), 703–708.

Rapach, D. E., Strauss, J. K. & Zhou, G. (2010), ‘Out-of-sample equity premium prediction: Combination forecasts and links to the real economy’, *Review of Financial Studies* **23**(2), 821–862.

RavenPack (2020), *RavenPack Analytics—User Guide and Service Overview*, RavenPack.

Sarkar, S. K. (2025), ‘Economic representations’.

Sarkar, S. K. & Vafa, K. (2024), ‘Lookahead bias in pretrained language models’, *Available at SSRN* .

Schwartzstein, J. (2014), 'Selective attention and learning', *Journal of the European Economic Association* **12**(6), 1423–1452.

Shen, Z. & Xiu, D. (2025), 'Can machines learn weak signals?', *Working Paper* . University of Chicago Booth School of Business.

Shiller, R. J. (2019), *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press.

Stock, J. H. & Watson, M. W. (2004), 'Combination forecasts of output growth in a seven-country data set', *Journal of Forecasting* **23**(6), 405–430.

Tversky, A. & Kahneman, D. (1974), 'Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty', *American association for the advancement of science* **185**(4157), 1124–1131.

van Binsbergen, J. H., Bryzgalova, S., Mukhopadhyay, M. & Sharma, V. (2024), (almost) 200 years of news-based economic sentiment, Technical report, National Bureau of Economic Research.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017), 'Attention is all you need', *Advances in Neural Information Processing Systems* **30**.

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

Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E., Narang, S., Chowdhery, A., Zhou, D. et al. (2022), 'Self-consistency improves chain of thought reasoning in language models', *arXiv preprint arXiv:2203.11171* .

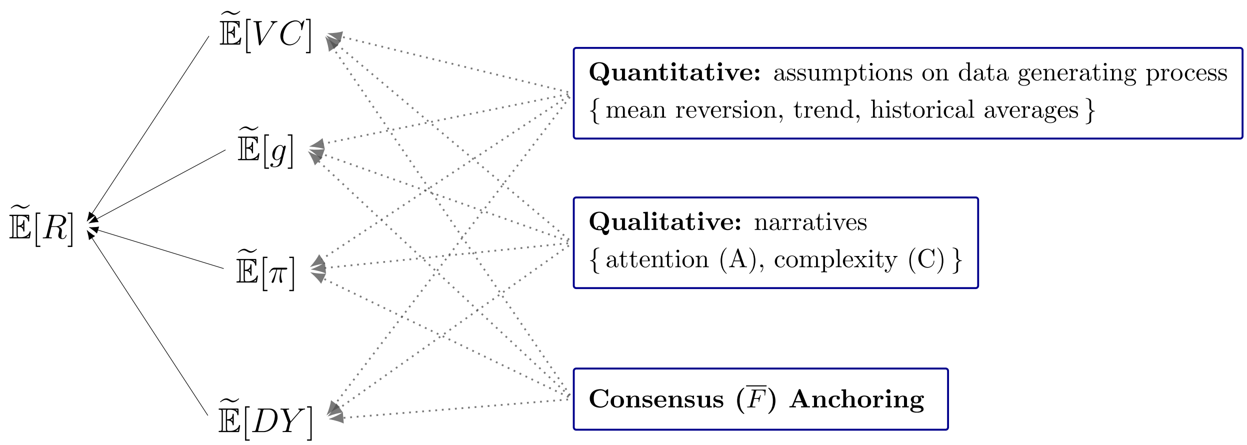
Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V. & Zhou, D. (2022), 'Chain-of-thought prompting elicits reasoning in large language models', *Advances in Neural Information Processing Systems (NeurIPS)* **35**.

Welch, I. & Goyal, A. (2008), 'A comprehensive look at the empirical performance of equity premium prediction', *Review of Financial Studies* **21**(4), 1455–1508.

Zhao, T. Z., Wallace, E., Feng, S., Klein, D. & Singh, S. (2021), 'Calibrate before use: Improving few-shot performance of language models', *arXiv preprint arXiv:2102.09690* .

Figure 1: A Unifying View of Asset Managers' Equity-Return Mental Models. This figure summarizes the empirical interpretation of institutional equity-belief formation that emerges from the paper. Most asset managers decompose headline equity-return expectations into valuation change (VC), real growth (g), inflation (π), and income or dividend yield (DY) expectations, a decomposition introduced in Section 2. The component forecasts are associated with disclosed modeling assumptions and calibration methods. Causal narratives capture the topics and transmission channels through which managers organize economic information, while narrative attention and complexity condition the response of forecasts to external earnings-news sentiment. Peer consensus provides an additional anchor for forecast revisions and benchmark-relative deviations. The arrows summarize the empirical channels studied in the paper. Volatility ($\widetilde{\text{Vol}}$) and correlation ($\widetilde{\text{Corr}}$) beliefs provide a contrasting case, as they are strongly connected to historical estimates.

Panel A: First Moment Forecasts



$$\widetilde{\mathbb{E}}[R] \approx \widetilde{\mathbb{E}}[VC] + \widetilde{\mathbb{E}}[g] + \widetilde{\mathbb{E}}[\pi] + \widetilde{\mathbb{E}}[DY] = \sum_i \widetilde{\mathbb{E}}[x_i]$$

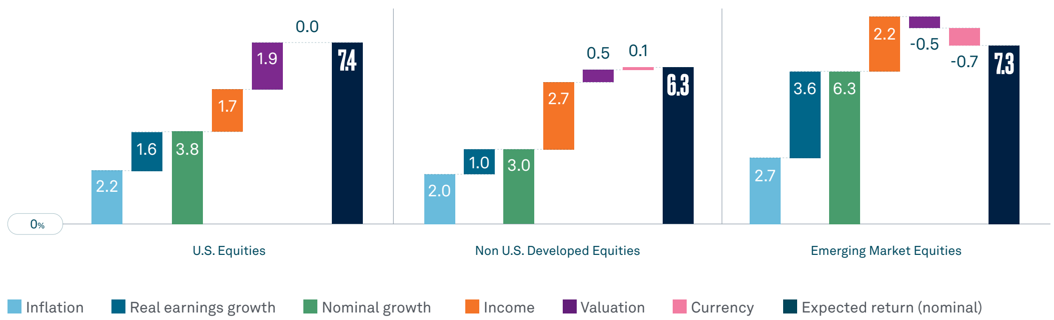
$$\widetilde{\mathbb{E}}[x_i] = f_{x_i}(\text{quantitative, qualitative, incentives}) := f_{x_i}\big(\text{assumptions, Sentiment}(\text{A}, \text{C}),\ \bar{F}\big)$$

Panel B: Second Moment Forecasts

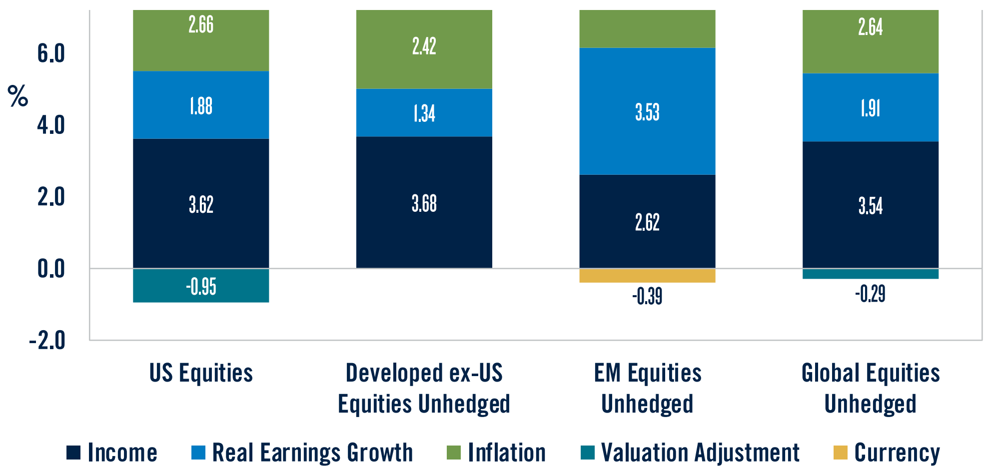
$$\widetilde{\text{Vol}},\ \widetilde{\text{Corr}} \approx g(\text{historical averages})$$

Figure 2: Building-Block Decompositions of Equity Return Expectations. This figure presents three asset managers' approaches to decomposing equity return forecasts into fundamental components for 2023. Panel (a): BNY Mellon separates returns into inflation, real earnings growth, income, valuation, and currency effects. Panel (b): PGIM uses income, real earnings growth, inflation, valuation adjustment, and currency. Panel (c): T. Rowe Price employs valuation change, expected inflation, dividend/buyback yield, and real EPS growth.

(a) BNY Mellon (30 September 2023)



(b) PGIM (30 September 2023)



(c) T. Rowe Price (31 January 2023)

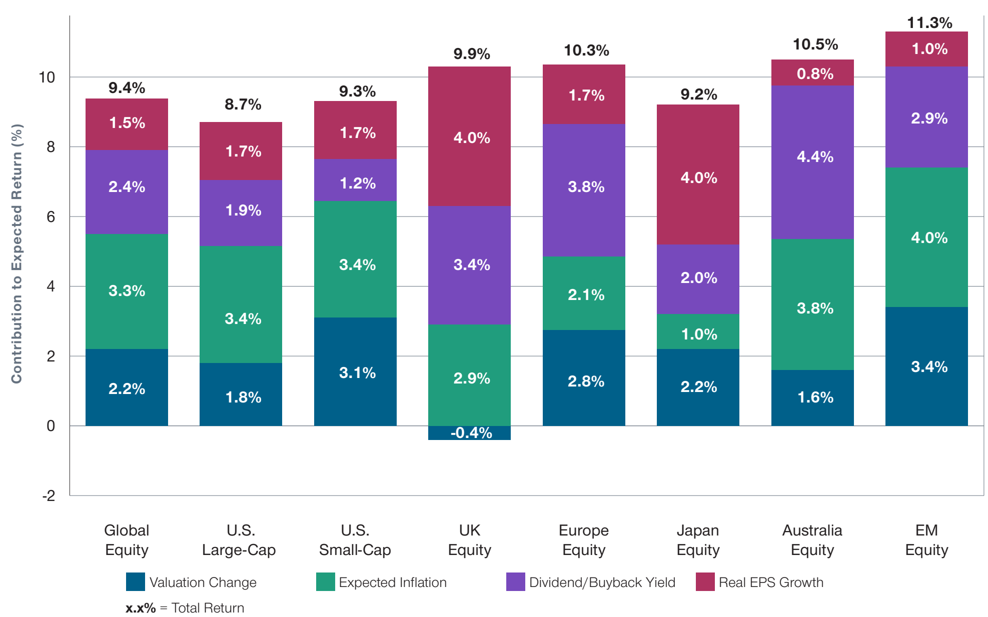


Figure 3: Subjective Return Beliefs for U.S. Equity. This figure shows forecasts for U.S. equity from CMA reports. Panel (a) shows subjective return expectations. Panel (b) shows subjective volatility forecasts. Panel (c) displays time-series of return expectations from five representative asset managers (JP Morgan, AQR, Verus, Sellwood, Northern Trust). Sample period: 2015–2026.

Figure 4: LLM Pipeline for Extracting Causal Networks. This figure illustrates the pipeline that transforms Capital Market Assumptions (CMAs) narratives into directed, signed causal networks. We segment each report into text chunks, use a large language model to extract explicit causal claims (a causing concept, an affected concept, and the directional relationship), assign the concepts to standardized topics and subtopics, and link the signed claims into a directed network that converges on U.S. equity return expectations (R). In the network, solid arrows with filled heads denote positive links and dashed arrows with open heads denote negative links. Nodes and links are illustrative.

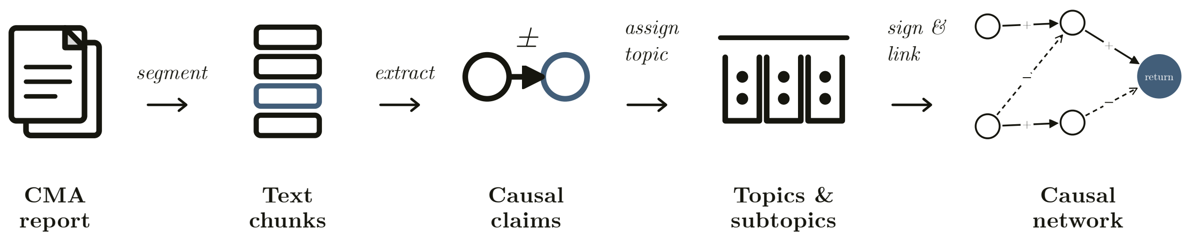


Figure 5: Schematic illustration of the main network-complexity measures. Panel (a) shows chain complexity: average path length and the indirect connection ratio are closely related measures of whether the manager explains a focal outcome through direct links or through multi-step causal chains. Panel (b) shows interconnected complexity: network transitivity measures whether the manager's causal reasoning forms closed triangles of relationships.

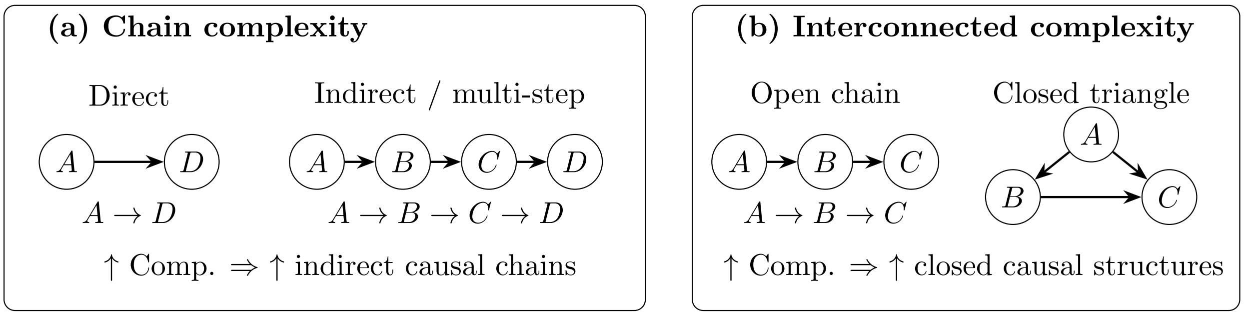


Figure 6: Building Block Structure in Institutional Equity Return Models. This figure shows the structure of return decompositions across asset managers using building blocks (described in Section 2.1). Panel (a) displays the distribution of reports by number of components used in their return decomposition. Panel (b) presents the adoption rate of each component across institutions: nominal growth, dividend yield, valuation change, buybacks, margin adjustments, and share issuance. Blue bars indicate inclusion and gold bars indicate exclusion. Sample period: 2010–2026.

(a) Distribution of building blocks per model

(b) Institutional adoption by component type

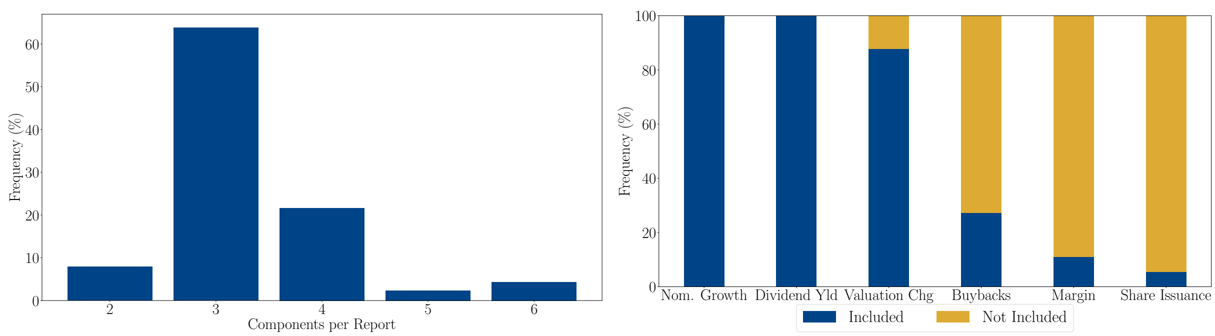


Figure 7: Sources of Variation in Subjective Equity Return Forecasts. This figure shows a variance decomposition using partial R^2 from a saturated fixed-effects model of U.S. equity return expectations. Each bar represents the percentage of total variance explained by asset manager, year, and forecast horizon fixed effects, respectively. Sample period: 2010–2026.

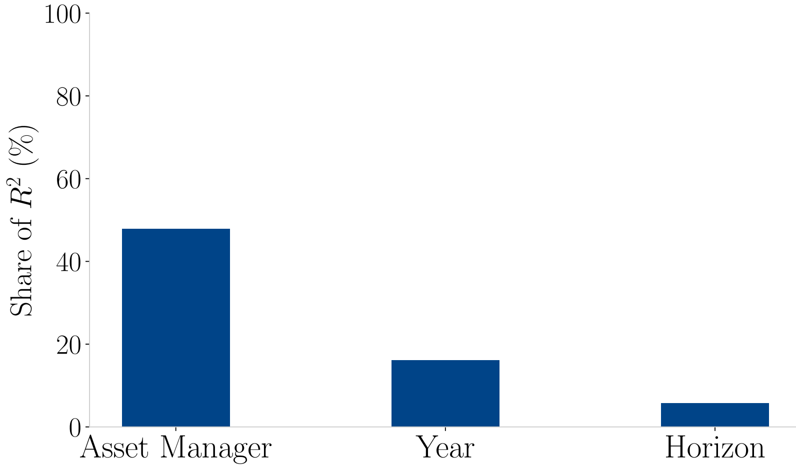


Figure 8: Time-Series of Return Expectation Components. This figure shows average subjective U.S. equity return expectations decomposed into constituent building blocks (described in Section 2.1). Each colored area represents one component's contribution: valuation change, dividend yield, real growth, and inflation. Dividend yield aggregates dividend yield, buybacks, and share issuance; real growth aggregates real earnings growth and margin adjustments. The red line indicates the average subjective equity return across all asset managers. Sample period: 2012–2025.

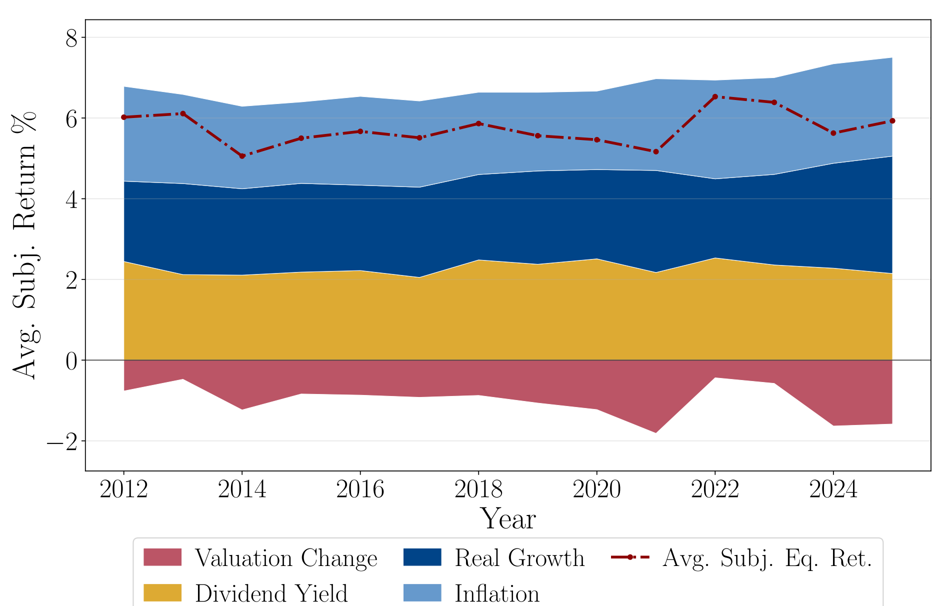


Figure 9: Time-Series of Dispersion in Return Expectation Components. This figure shows annual cross-sectional standard deviations of each building block component (described in Section 2.1), measuring disagreement among asset managers. Each colored area represents the dispersion of one component: valuation change, dividend yield, real growth, and inflation. Dividend yield aggregates dividend yield, buybacks, and share issuance; real growth aggregates real earnings growth and margin adjustments. The red dashed line shows the standard deviation of total subjective equity returns across asset managers. Sample period: 2012–2025.

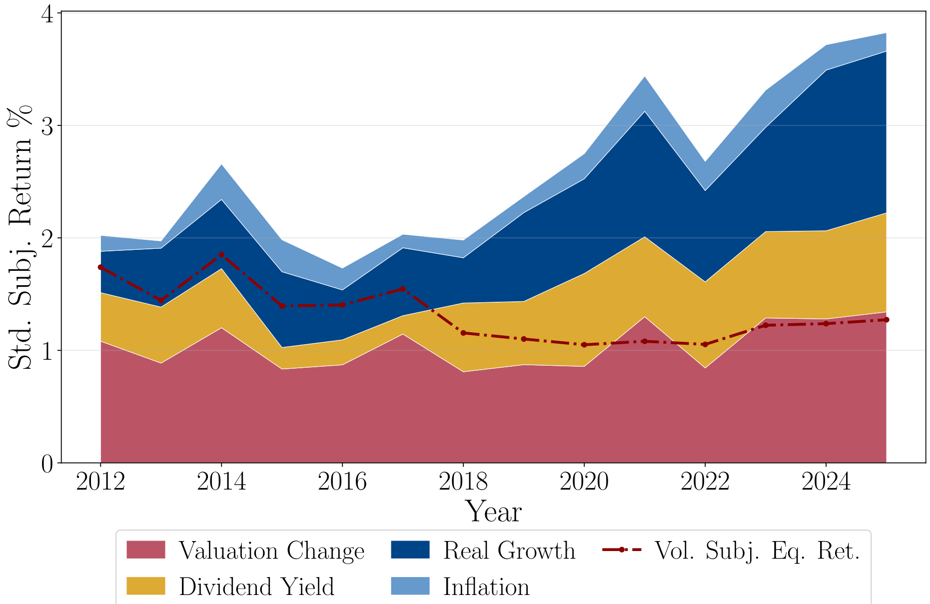


Figure 10: Variance Decomposition of Equity Return Forecasts by Building Block Components. This figure shows partial R^2 estimates from a regression of subjective U.S. equity return expectations on four building block components (described in Section 2.1) with time fixed effects. Each bar represents the percentage of cross-sectional variance in return forecasts explained by valuation change, real growth, dividend yield, and inflation. Dividend yield aggregates dividend yield, buybacks, and share issuance; real growth aggregates real earnings growth and margin adjustments. Sample period: 2010–2025.

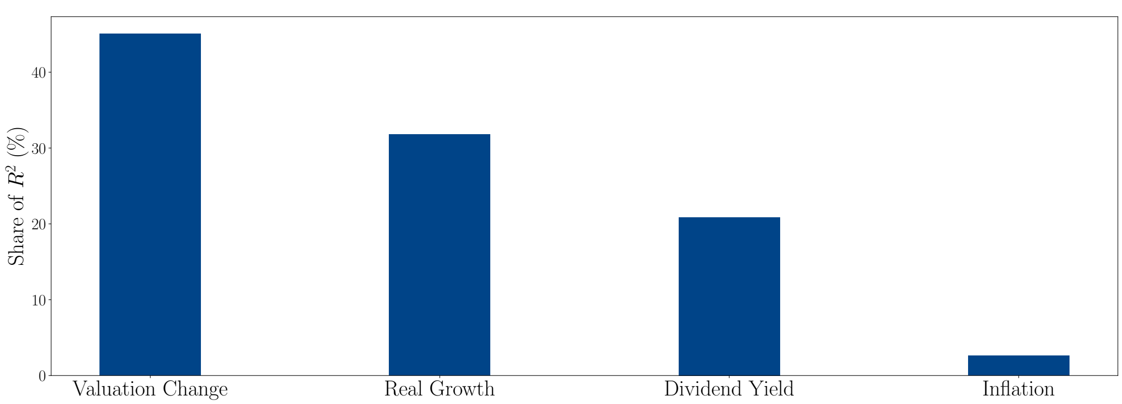
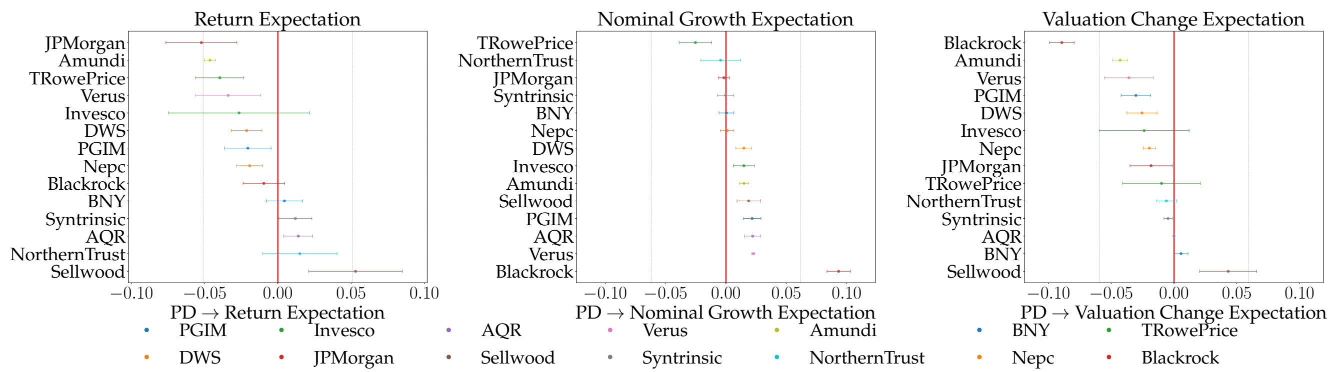


Figure 11: Campbell-Shiller Decomposition: Asset-Manager-Level Loadings on Valuation Ratios. This figure shows regression coefficients from asset-manager-level specifications in which CMA forecast components are regressed on valuation ratios. In each panel, $F^b_{i,t,h}$ denotes asset manager i's forecast for component b at CMA reference date t and horizon h, where $b \in \{ER, NG, VC\}$. The three columns report, respectively, the headline U.S. equity return forecast (ER), the nominal-growth component (NG), and the valuation-change component (VC). Panel (a) regresses forecasts on $pd_t = \ln(PD_t)$, the log price-dividend ratio. Panel (b) regresses forecasts on $px_t = \ln(CAPE_t)$, the log cyclically-adjusted price-earnings ratio. Dots denote point estimates; horizontal bars are 95% confidence intervals. Sample period: 2010–2025.

$$\text{(a)} \quad F^b_{i,t,h} = \alpha^b_i + \beta^b_{i,pd} pd_t + \varepsilon^b_{i,t,h} \qquad\qquad \text{(b)} \quad F^b_{i,t,h} = \alpha^b_i + \beta^b_{i,px} px_t + \varepsilon^b_{i,t,h}$$

(a) CMA forecast components on $pd_t = \ln(PD_t)$.



(b) CMA forecast components on $px_t = \ln(CAPE_t)$.

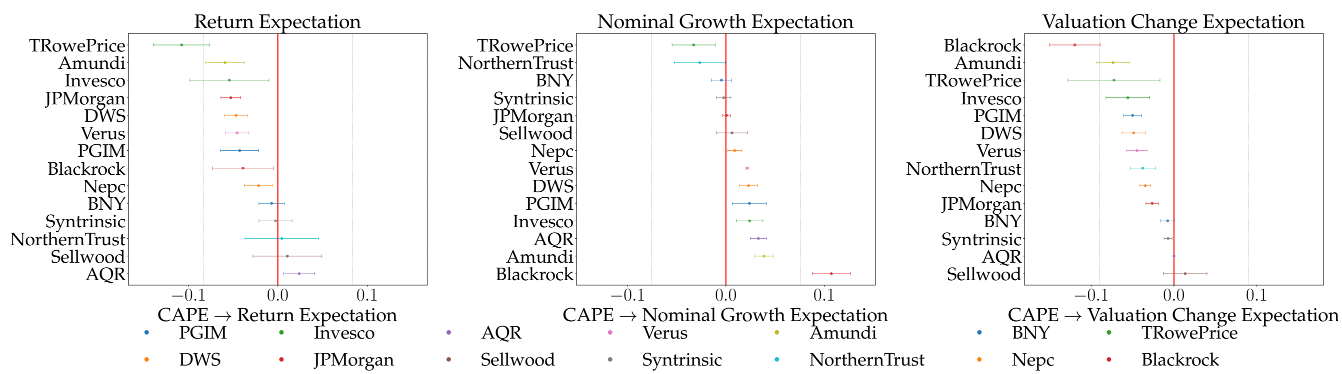


Figure 12: Causal network derived from a Campbell–Shiller (CS) approximation and a classic consumption-based asset-pricing model. Each node represents either a component of the subjective CS approximation or a state variable in the asset-pricing model of agent i, while each directed edge corresponds to a partial derivative (i.e., a "causal" link) between variables. If node A points to node B with linkage strength 1, this indicates that $\partial B/\partial A = 1$. Blue edges denote linkages among CS variables that directly relate to expected returns; black edges denote linkages among CS variables that do not directly connect to expected returns; and red nodes and edges represent additional components implied by the baseline consumption-based asset-pricing model.

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] = \kappa + \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] - (1-\rho)\, pd_t + \rho\, \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}]$$

$$\Delta c_t = (1-\phi)\mu + \phi \Delta c_{t-1} + \phi \epsilon_t$$

$$pd_t = A + B\, \Delta c_t$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] = (1-\phi)\mu + \phi \Delta c_t$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}] = B(1-\phi)(\mu - \Delta c_t)$$

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] = \kappa + (\rho - 1)A + (1-\phi)\mu(1+\rho B) + [\phi(1+\rho B) - B]\, \Delta c_t$$

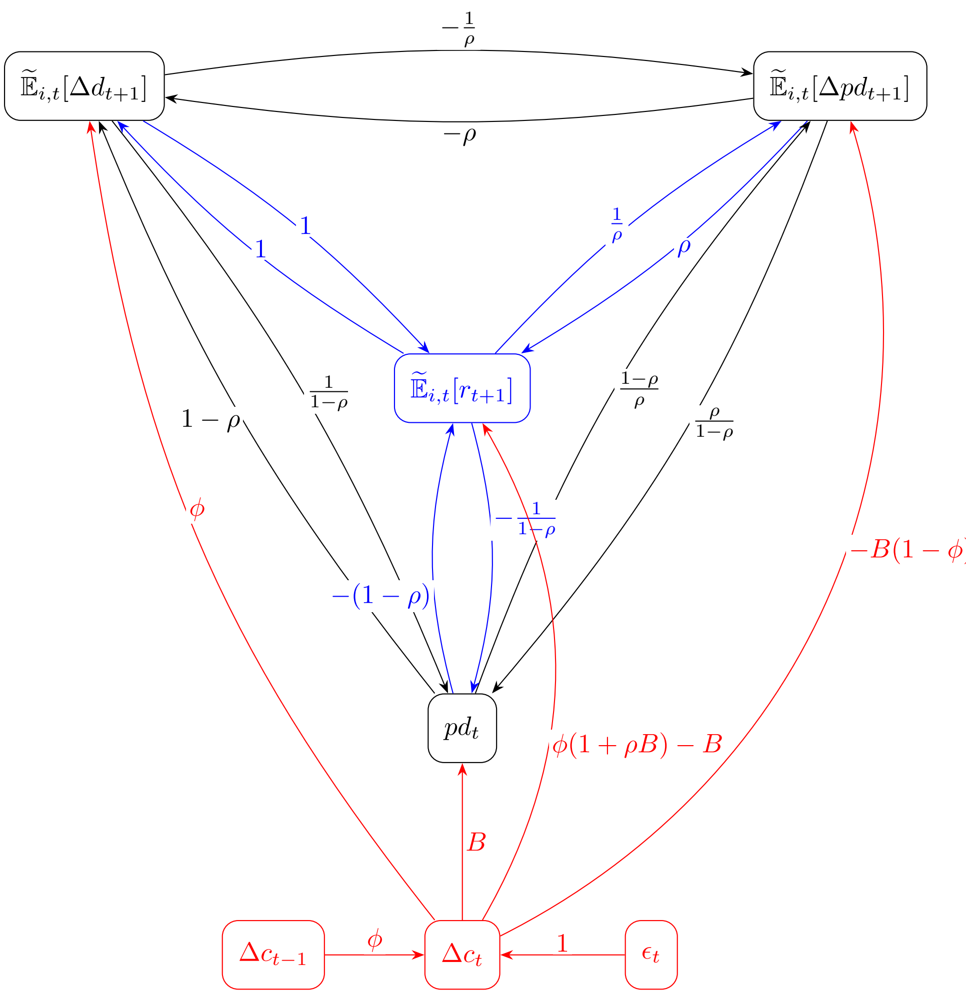


Figure 13: Model feature importance and the manager-versus-machine comparison. Panel (a) reports normalized permutation importance for the 16 Welch & Goyal (2008) predictors in models trained to predict annualized realized U.S. equity returns, $R_{t,t+h}$, the target used to construct the objective benchmark $G_{t,h}$. Columns correspond to the seven regularized learners, and rows are sorted by equal-weighted ensemble importance across learners. Panel (b) compares feature importance from machine-learning models trained on two different targets. The horizontal axis reports equal-weighted ensemble importance from models trained to predict realized U.S.-equity returns, $R_{t,t+h}$; the vertical axis reports equal-weighted ensemble importance from models trained to predict asset managers' headline CMA forecasts, $F^{ER}_{i,t,h}$. The dashed 45-degree line marks equal feature importance across the two prediction exercises. All predictors are lagged one year. Permutation importance is computed as the average reduction in in-sample R^2 after randomly shuffling each predictor, using $n = 10$ repeats, clipped at zero, and normalized to sum to one within each model column in Panel (a) and within each ensemble vector in Panel (b). Predictor definitions are provided in Table A2. The realized-return models use the Goyal et al. (2024) data over 1948–2025; the CMA-forecast models use CMA reports over 2008–2026.

(a) Model feature importance. **(b)** Manager versus machine importance.

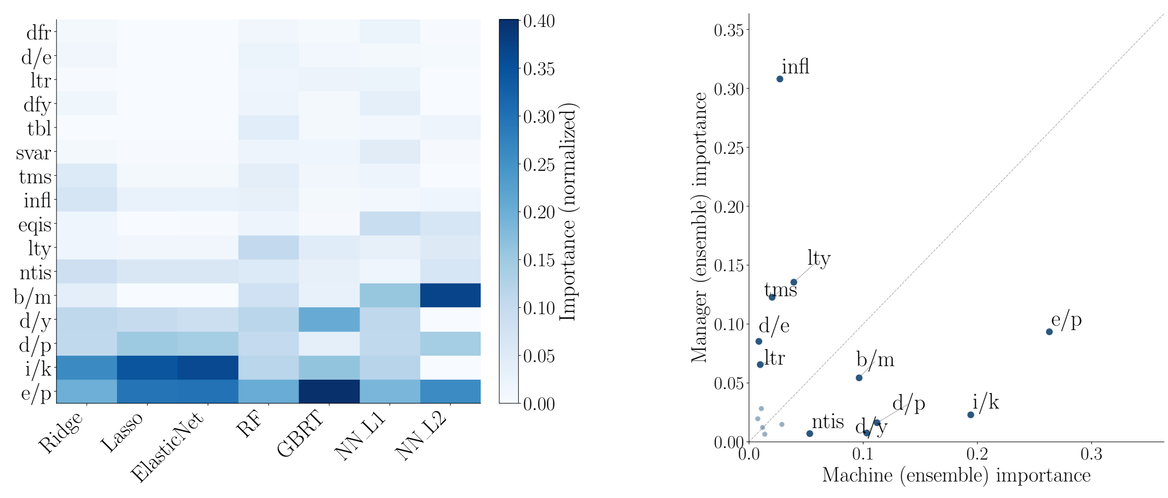


Figure 14: Sources of Variation in Volatility Forecasts. This figure shows Shapley values decomposing the R^2 from a saturated fixed-effects regression of U.S. equity volatility forecasts. Each bar represents the percentage of total variance explained by asset manager, year, and horizon fixed effects, respectively. Sample period: 2010–2025.

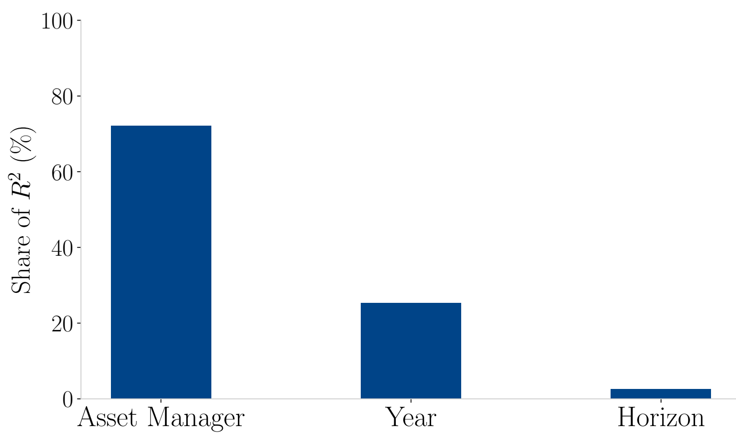


Figure 15: Subjective Volatility Forecasts versus Realized Volatility Measures. This figure compares asset managers' subjective volatility forecasts (blue box plots) with exponentially-weighted moving averages (EWMA) of trailing realized volatility for U.S. equity, quarterly frequency. Box plots show cross-sectional distribution of forecasts across managers at each point in time. Lines represent EWMA of 10-year trailing realized volatility with decay parameters: $\lambda = 0.98$, $\lambda = 0.75$, and $\lambda = 1.00$. Forecasts are winsorized at 1st and 99th percentiles. Sampl period: 2012–2025.

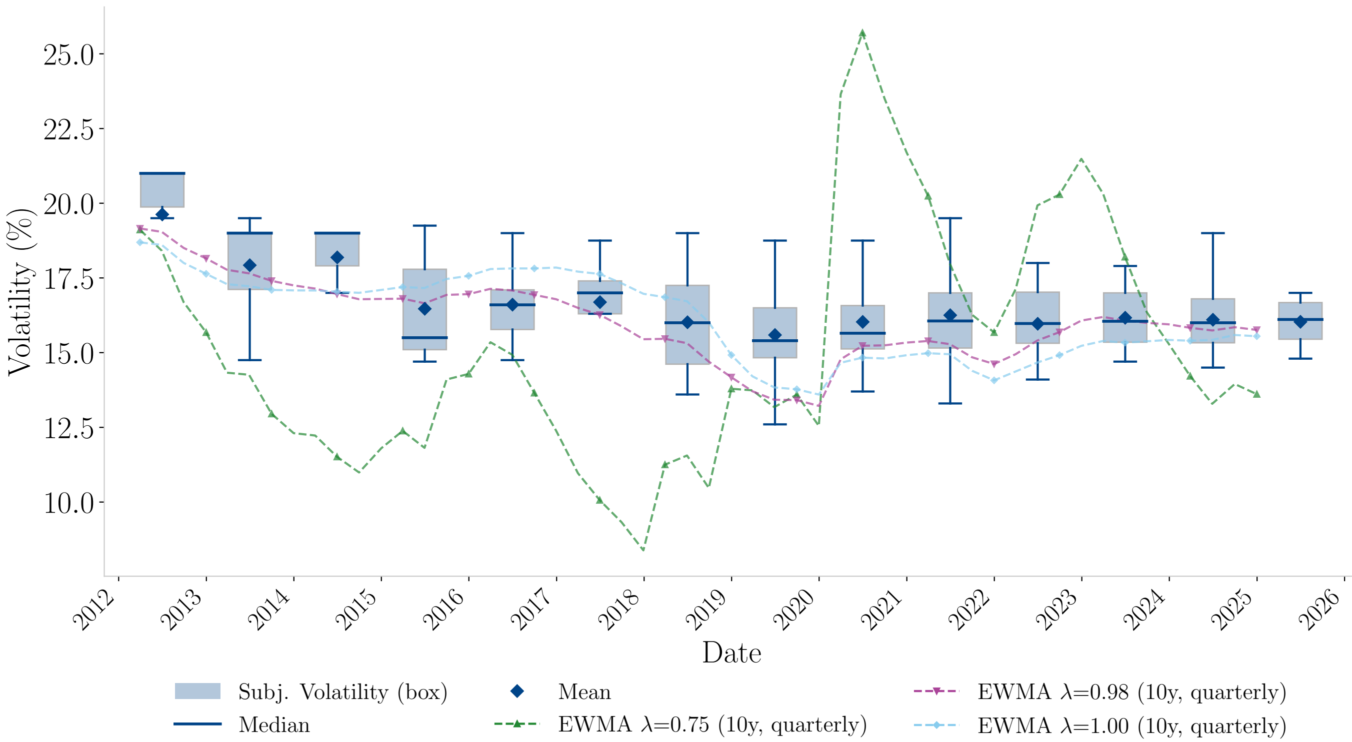
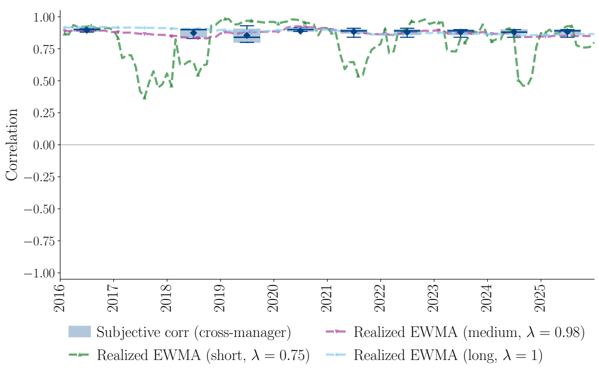
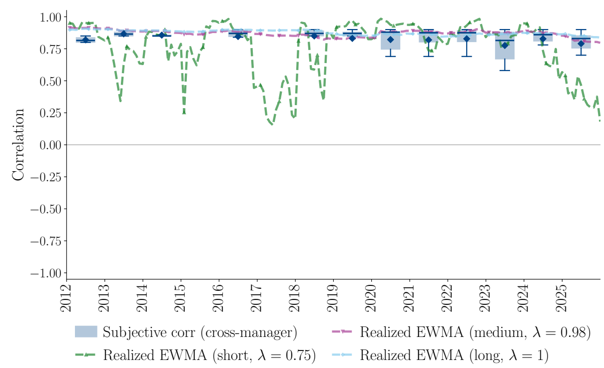


Figure 16: Subjective and Realized Correlations with U.S. Large-Cap Equity. For each counterparty class, the boxes show the cross-manager distribution of the subjective correlation between U.S. large-cap equity and that class, one box per vintage year. The dashed lines show the realized correlation between the S&P 500 total-return index and the Bloomberg total-return index mapped to the counterparty, on trailing ten-year monthly windows under three exponential-weighting schemes ($\lambda = 0.75$, 0.98, 1). The counterparty indices are the Russell 2000 (small cap), MSCI EAFE (developed ex-US), MSCI EM (emerging markets), the Bloomberg US Aggregate (aggregate bonds), US Treasury Intermediate (government bonds), and US Corporate High Yield (high-yield corporates). Sample period: 2016–2025.

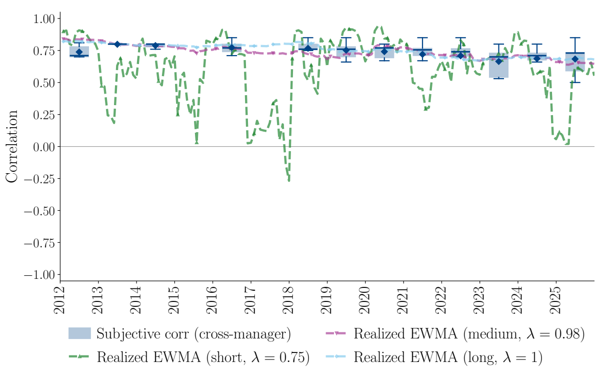
(a) U.S. Small-Cap Equity



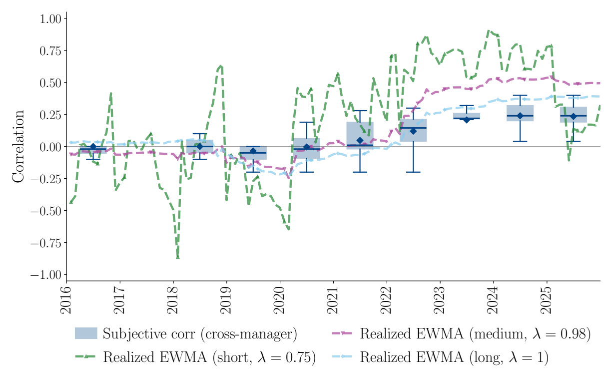
(b) Developed ex-US Equity



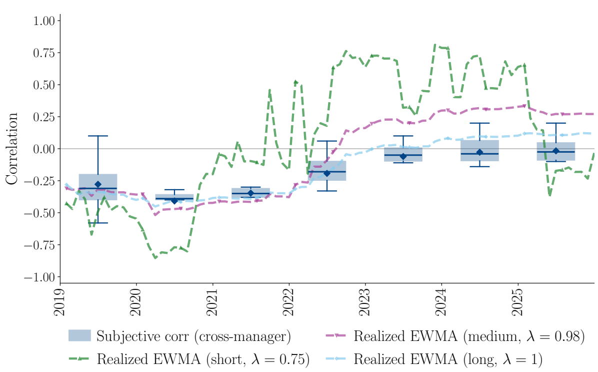
(c) Emerging-Market Equity



(d) U.S. Aggregate Bonds



(e) U.S. Government Bonds



(f) U.S. High-Yield Corporate Bonds

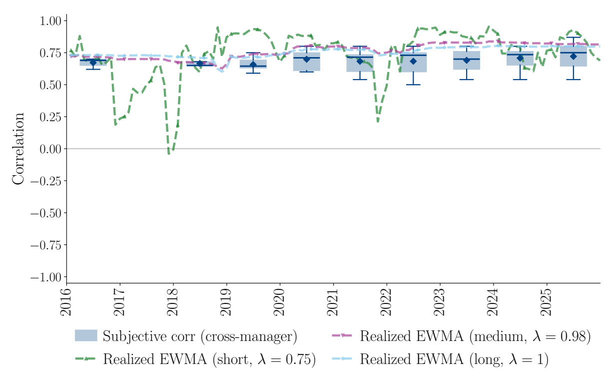


Figure 17: Disagreement and Deviation from History in Pairwise Correlation Assumptions, by Asset Class. Each heatmap reports, for an unordered pair of asset classes (m, n), the mean absolute deviation of managers' assumed correlation $\rho^{i}_{t,m,n}$ from a benchmark, averaged over managers i and vintage years t. Panel (a) uses the leave-one-out cross-manager consensus $\bar{\rho}^{-i}_{t,m,n}$ (disagreement); panel (b) uses the realized trailing ten-year monthly Bloomberg correlation $\hat{\rho}_{t,m,n}$ (deviation from history). A pair enters only when at least five managers price it in a year, with the same support on both panels; the matrix is symmetric, the diagonal blank. The right strip gives, per class, the mean of its cell values across partner classes, with the standard deviation in parentheses. Within each panel classes are ordered by mean deviation. Sample period: 2021–2025.

(a) Benchmark: leave-one-out consensus

(b) Benchmark: realized history

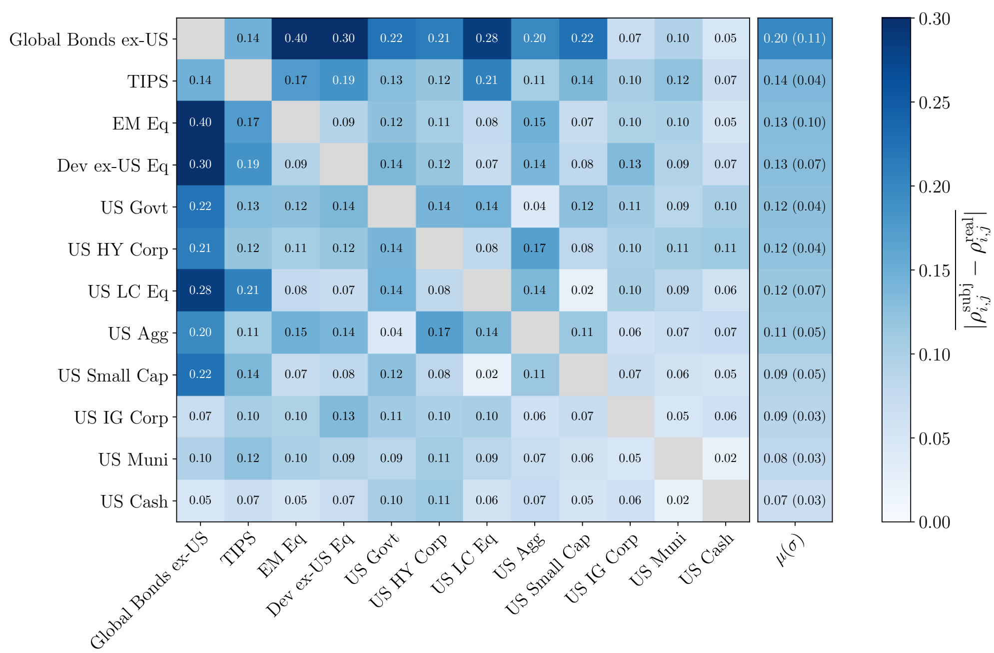
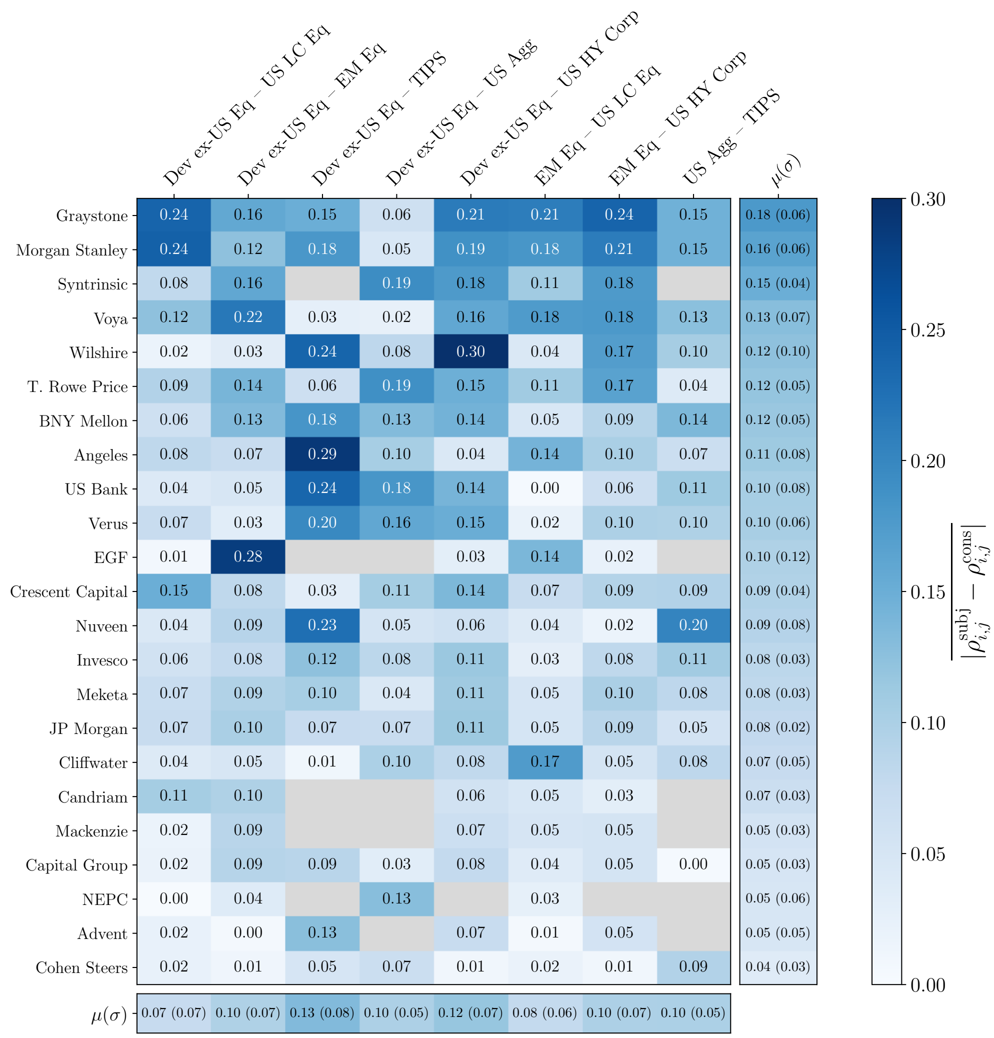


Figure 18: Disagreement and Deviation from History in Pairwise Correlation Assumptions, by Manager. For the eight most widely reported asset-class pairs, each heatmap shows, per asset manager (rows) and pair (columns), the mean absolute deviation of the manager's assumed correlation from a benchmark, averaged over vintage years: panel (a) uses the leave-one-out cross-manager consensus $\bar{\rho}^{-i}_{t,m,n}$ (disagreement); panel (b) uses the realized trailing ten-year monthly Bloomberg correlation $\hat{\rho}_{t,m,n}$ (deviation from history). A pair-year enters the sample only when at least five managers report correlation assumptions for that pair in that year. Both panels are computed using the same managers, asset-class pairs, and manager-pair-year observations. The right strip gives, per manager, the mean deviation across the eight pairs with the standard deviation in parentheses; the bottom strip gives, per pair, the mean and standard deviation across managers. Within each panel managers are ordered by mean deviation, so the two panels may list managers in a different order. Sample period: 2021–2025.

(a) Benchmark: leave-one-out consensus



(b) Benchmark: realized history

	Dev ex-US Eq – US LC Eq	Dev ex-US Eq – EM Eq	Dev ex-US Eq – TIPS	Dev ex-US Eq – US Agg	Dev ex-US Eq – US HY Corp	EM Eq – US LC Eq	EM Eq – US HY Corp	US Agg – TIPS	$\mu(\sigma)$
Graystone	0.28	0.17	0.26	0.18	0.21	0.19	0.20	0.21	0.21 (0.04)
Morgan Stanley	0.28	0.14	0.36	0.16	0.19	0.16	0.17	0.23	0.21 (0.08)
Voya	0.17	0.22	0.22	0.14	0.16	0.15	0.14	0.21	0.18 (0.03)
Wilshire	0.06	0.05	0.39	0.19	0.29	0.05	0.14	0.18	0.17 (0.12)
T. Rowe Price	0.03	0.11	0.23	0.24	0.12	0.11	0.17	0.12	0.14 (0.07)
Angeles	0.03	0.05	0.35	0.10	0.03	0.15	0.11	0.15	0.12 (0.11)
Nuveen	0.01	0.06	0.41	0.13	0.05	0.05	0.04	0.17	0.12 (0.13)
EGF	0.05	0.30			0.03	0.12	0.05		0.11 (0.11)
Crescent Capital	0.18	0.11	0.08	0.21	0.14	0.06	0.06	0.02	0.11 (0.07)
Syntrinsic	0.01	0.12		0.01	0.16	0.11	0.20		0.10 (0.08)
Cliffwater	0.02	0.01	0.24	0.26	0.08	0.17	0.02	0.01	0.10 (0.11)
Advent	0.07	0.03	0.33		0.07	0.01	0.02		0.09 (0.12)
NEPC	0.05	0.00		0.27		0.02			0.09 (0.13)
Candriam	0.15	0.12			0.06	0.04	0.01		0.07 (0.06)
JP Morgan	0.01	0.07	0.10	0.09	0.09	0.06	0.10	0.04	0.07 (0.03)
Capital Group	0.03	0.05	0.13	0.12	0.07	0.03	0.08	0.05	0.07 (0.04)
BNY Mellon	0.01	0.10	0.04	0.08	0.12	0.04	0.10	0.04	0.07 (0.04)
Mackenzie	0.07	0.11			0.08	0.05	0.03		0.07 (0.03)
Meketa	0.01	0.06	0.08	0.09	0.09	0.05	0.11	0.01	0.06 (0.04)
Cohen Steers	0.05	0.01	0.21	0.15	0.01	0.02	0.02	0.03	0.06 (0.08)
Verus	0.03	0.01	0.03	0.08	0.13	0.02	0.11	0.03	0.06 (0.04)
Invesco	0.01	0.05	0.06	0.05	0.09	0.04	0.10	0.02	0.05 (0.03)
US Bank	0.01	0.01	0.01	0.02	0.13	0.00	0.09	0.05	0.04 (0.05)
$\mu(\sigma)$	0.07 (0.08)	0.09 (0.07)	0.20 (0.14)	0.14 (0.08)	0.11 (0.07)	0.07 (0.06)	0.09 (0.06)	0.09 (0.08)	

Color scale: $\overline{|\rho^{\mathrm{subj}}_{i,j} - \rho^{\mathrm{real}}_{i,j}|}$ — 0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30

Figure 19: Year-over-Year Revision of U.S.-Equity Return, Volatility, and Correlation Beliefs. For each asset manager, the horizontal bar stacks the typical absolute year-over-year revision of three U.S. large-cap-equity beliefs: the return expectation, the return volatility, and the average correlation of U.S. large-cap equity with the other core asset classes. For each belief the revision is the mean absolute change between consecutive vintage years, in the belief's native units, with no rescaling. Because correlations are bounded in $[-1, 1]$ while returns and volatilities are expressed in percentage points, the three segments are not on a common scale, so the figure compares revision magnitudes within rather than across beliefs. Managers are ordered by the size of the return-expectation segment. Sample period: 2018–2026.

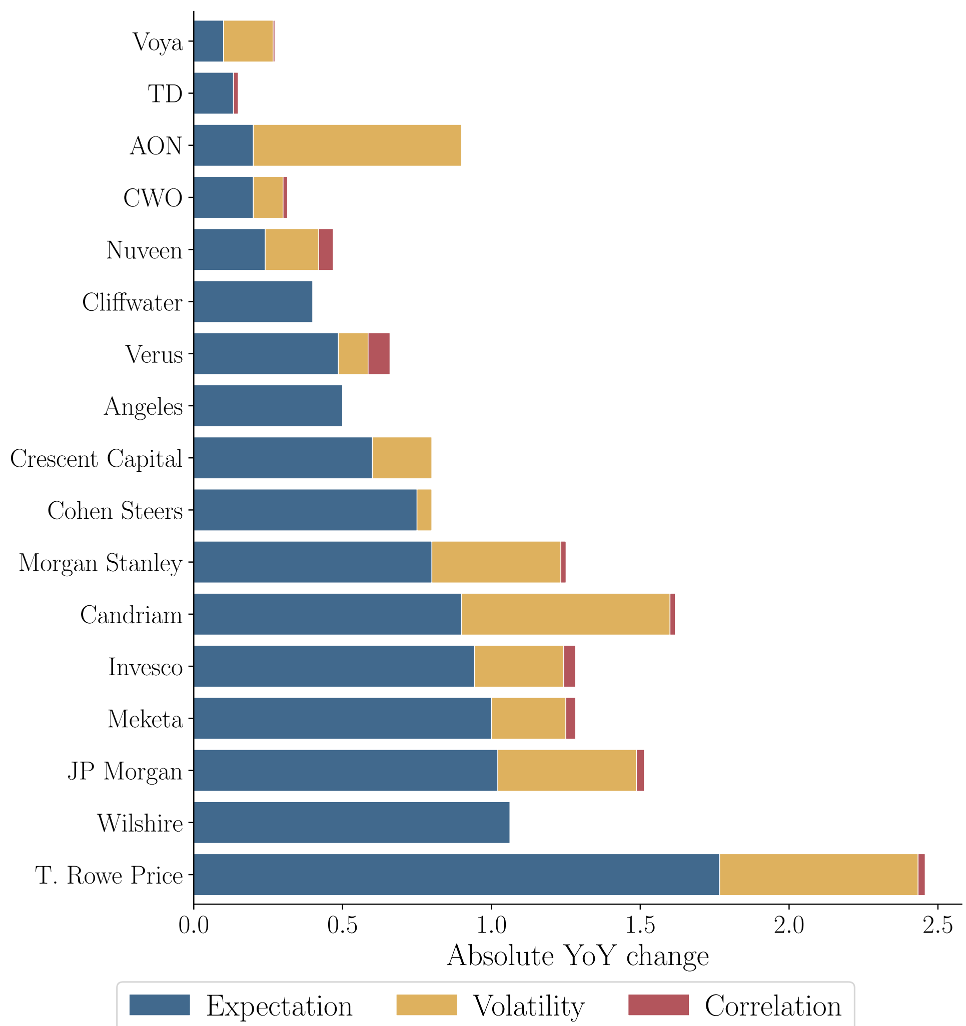


Figure 20: Prevalence of Forecasting Techniques across Reports. This figure reports the share of CMA reports in which each of six forecasting techniques is utilized: regressions, simulations, historical averages, mean reversion, trend hypotheses, and others. The indicators are non-exclusive: a single report can declare more than one technique. Sample: 2010–2025.

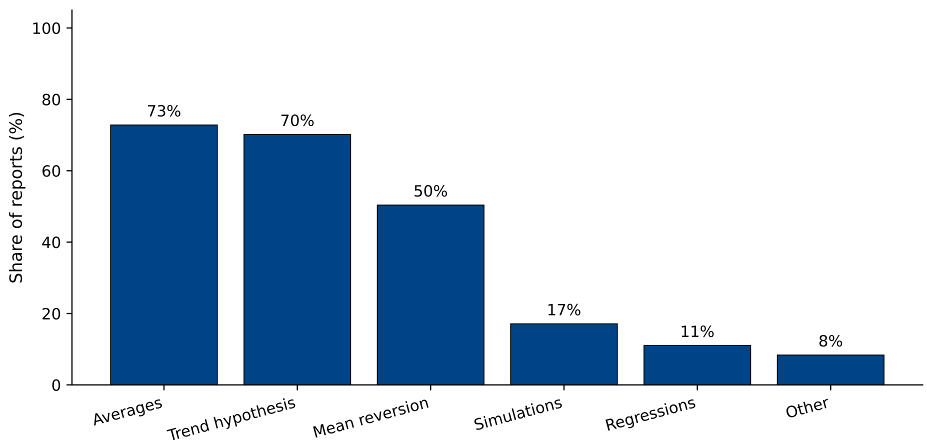


Figure 21: Author Education Counts and Building-Block Deviations from Consensus. Each cell reports the Pearson correlation between the number of authors in the report holding the indicated highest degree and the manager's deviation from the same-year leave-one-out cross-sectional consensus, $Dev_{i,t,b} = F_{i,t,b} - \bar{F}_{-i,t,b}$, on the corresponding equity building block. Authors are classified from LinkedIn profiles into Bachelor, Master / MBA, and PhD. Building blocks are equity return, real growth, inflation, nominal growth, dividend yield, and valuation change. Reports in which a given building block is not reported are excluded from both the consensus average and the correlation for that cell. Sample period: 2010 to 2024.

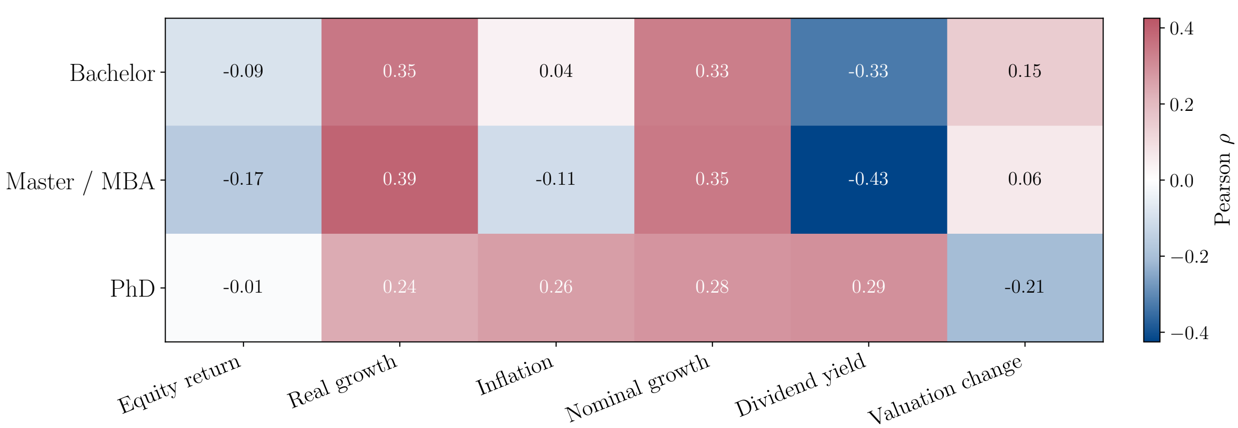


Figure 22: The Technique-Embedding Space. This figure shows a two-dimensional t-SNE projection of the 4,145 distinct modeling-technique phrases collected from the equity sections of 589 CMA reports. Each point is one technique phrase, positioned by the similarity of its meaning and coloured by the cluster to which it is assigned; labels mark the cluster centroids. Well-separated regions indicate that the seven-cluster partition captures genuine structure in the disclosed methodology. Sample period: 2008–2026.

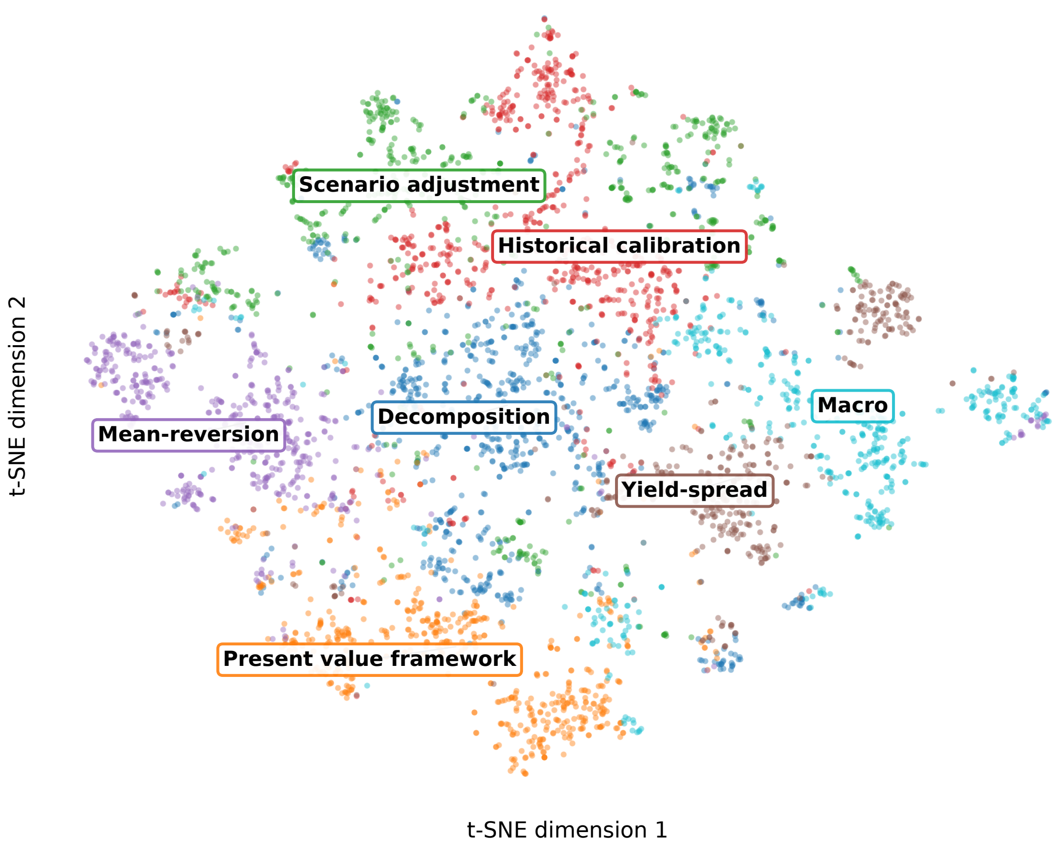


Figure 23: Topic-Based Clustering of CMA Reports. This figure shows a two-dimensional t-SNE projection of document-topic distributions from CMA reports. Each point represents one institution-report characterized by its LDA topic distribution. We first estimate a 55-topic LDA model on the reports, then apply K-Means to group them into 14 distinct clusters (shown by colors). Sample period: 2012–2025.

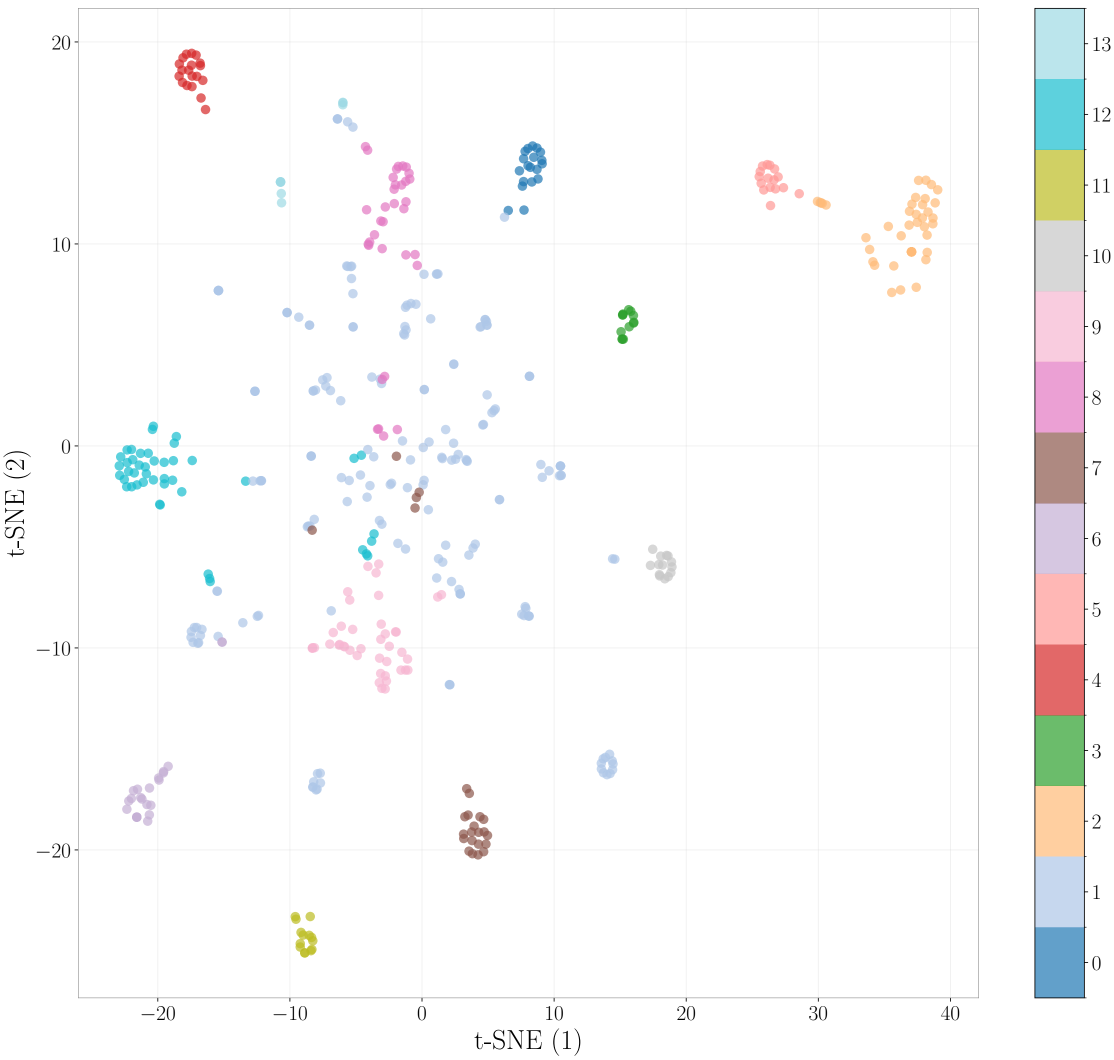


Figure 24: Asset Manager Causal Network of U.S. Equity Return Expectations. This figure shows the directed network of causal relationships (described in Section 1.3) extracted from DWS Capital Market Assumptions. Nodes represent topics identified in the text, with size reflecting centrality (number of connections). Edges indicate the direction of causality, with colors showing positive (green), negative (red), or neutral (gray) causal effects as stated in the reports. U.S. Equity Return and the four building blocks from the return decomposition (described in Section 2.1)—growth, valuation change, dividend yield, and inflation—are highlighted with distinct colors. Sample period: 2015–2025.

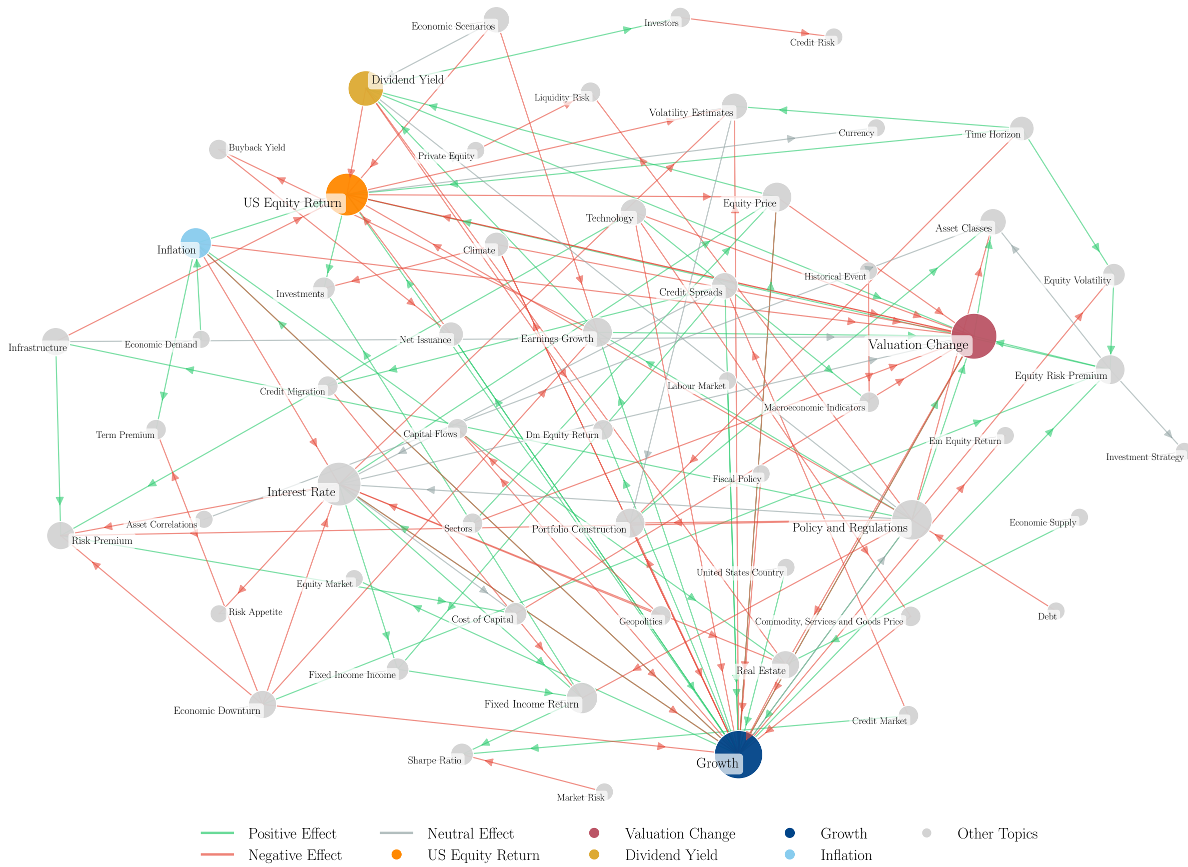


Figure 25: Asset Manager Direct Causal Drivers of U.S. Equity Return Expectations. This figure shows all topics directly linked to U.S. Equity Return in DWS Capital Market Assumptions. Node size reflects the number of causal connections (described in Section 1.3). Edges indicate the direction of causality, with colors showing positive (green), negative (red), or neutral (gray) effects as stated in the text. U.S. Equity Return and the four building blocks from the return decomposition (described in Section 2.1)—growth, valuation change, dividend yield, and inflation—are highlighted with distinct colors. Sample period: 2015–2025.

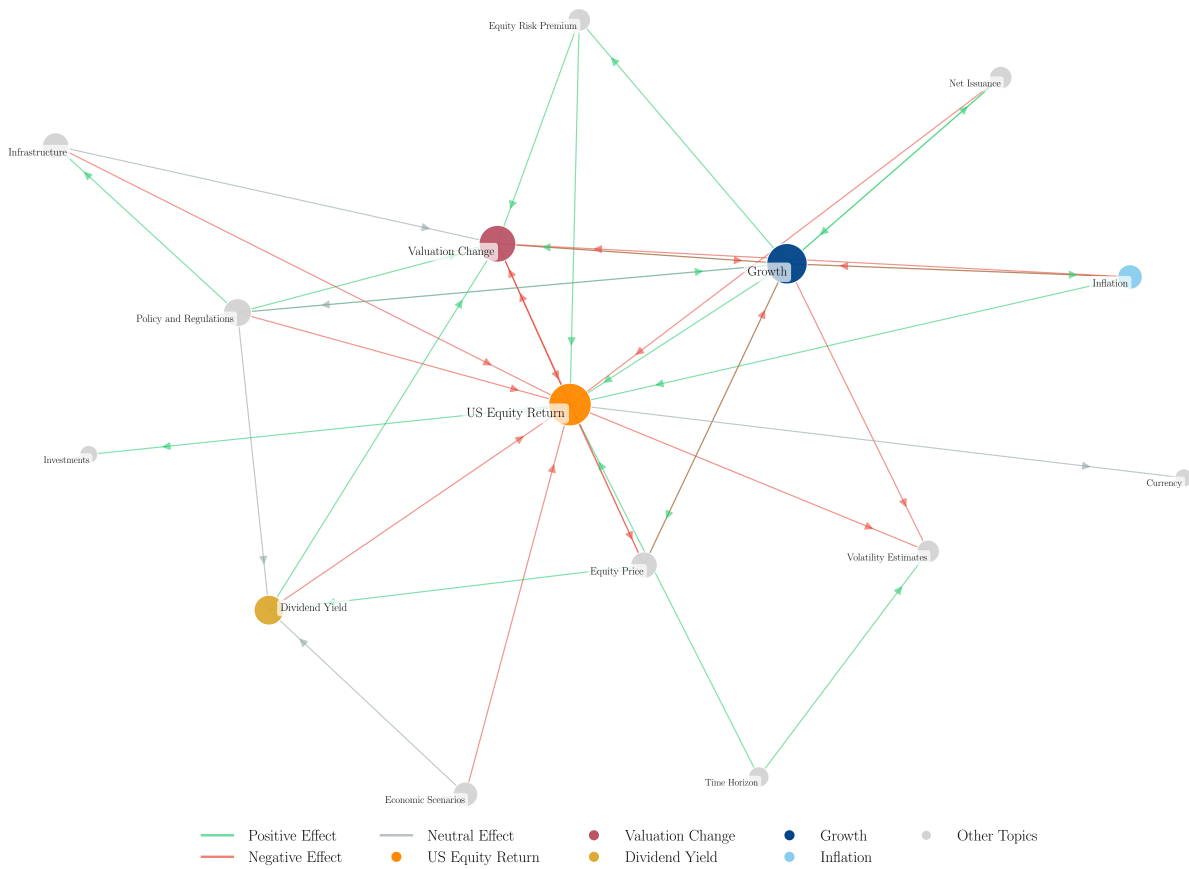
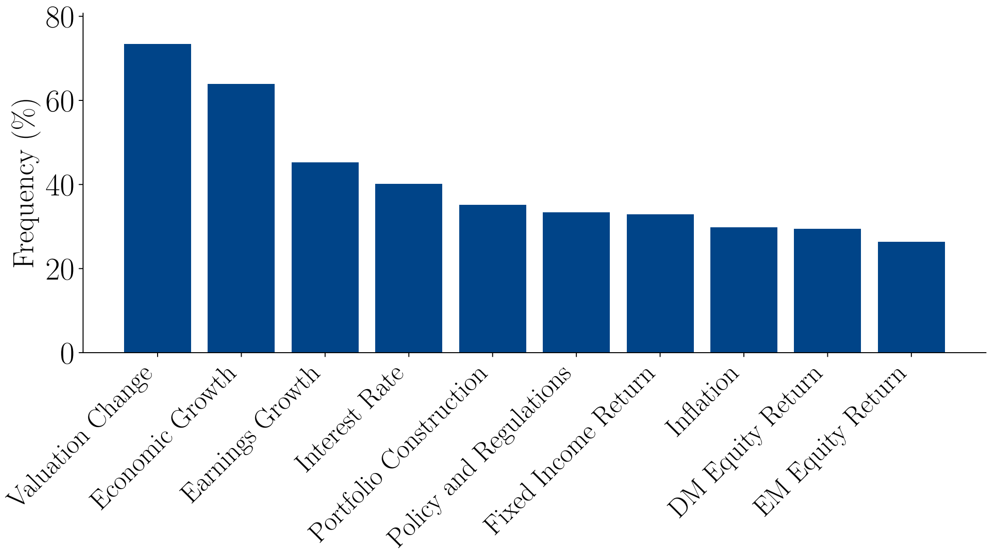


Figure 26: Causal Network Topic Composition. This figure shows the prevalence of topics in asset managers' causal networks (described in Section 1.3). Panel (a) shows the percentage of asset manager-year observations where each topic appears in any causal relationship (as cause or effect). Panel (b) shows the percentage of asset manager-year observations where each topic directly links to U.S. equity return expectations (as driver or consequence). Sample period: 2015–2025.

(a) Topic frequency



(b) Topics directly linked to U.S. equity return expectations

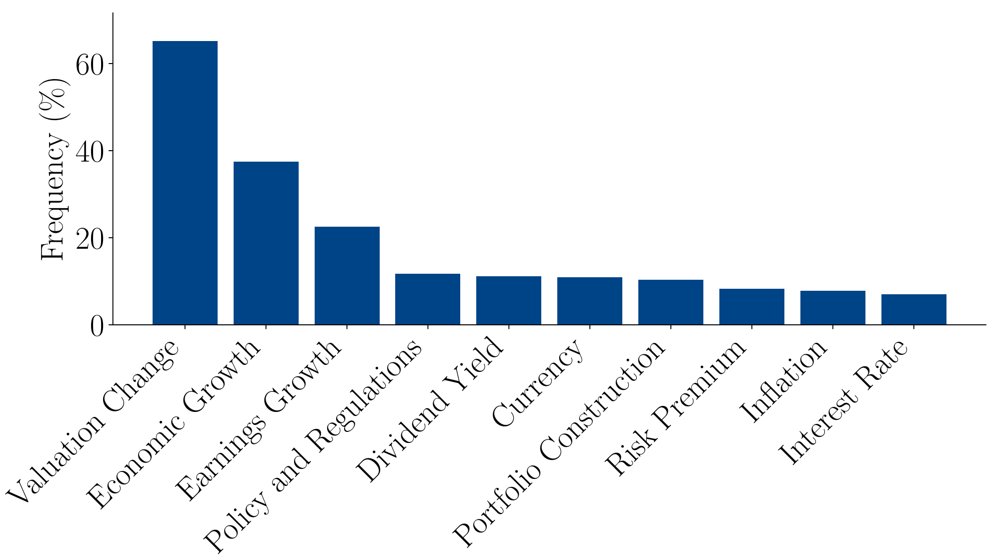


Figure 27: Heterogeneity in Causal Network Composition Across Asset Managers. This figure shows the average number of causal links (described in Section 1.3) per asset manager-year, decomposed by topic. Each horizontal bar represents one asset manager, with colored segments showing the contribution of the top 10 most frequent topics and gray segments showing all remaining topics. Sample period: 2015–2025.

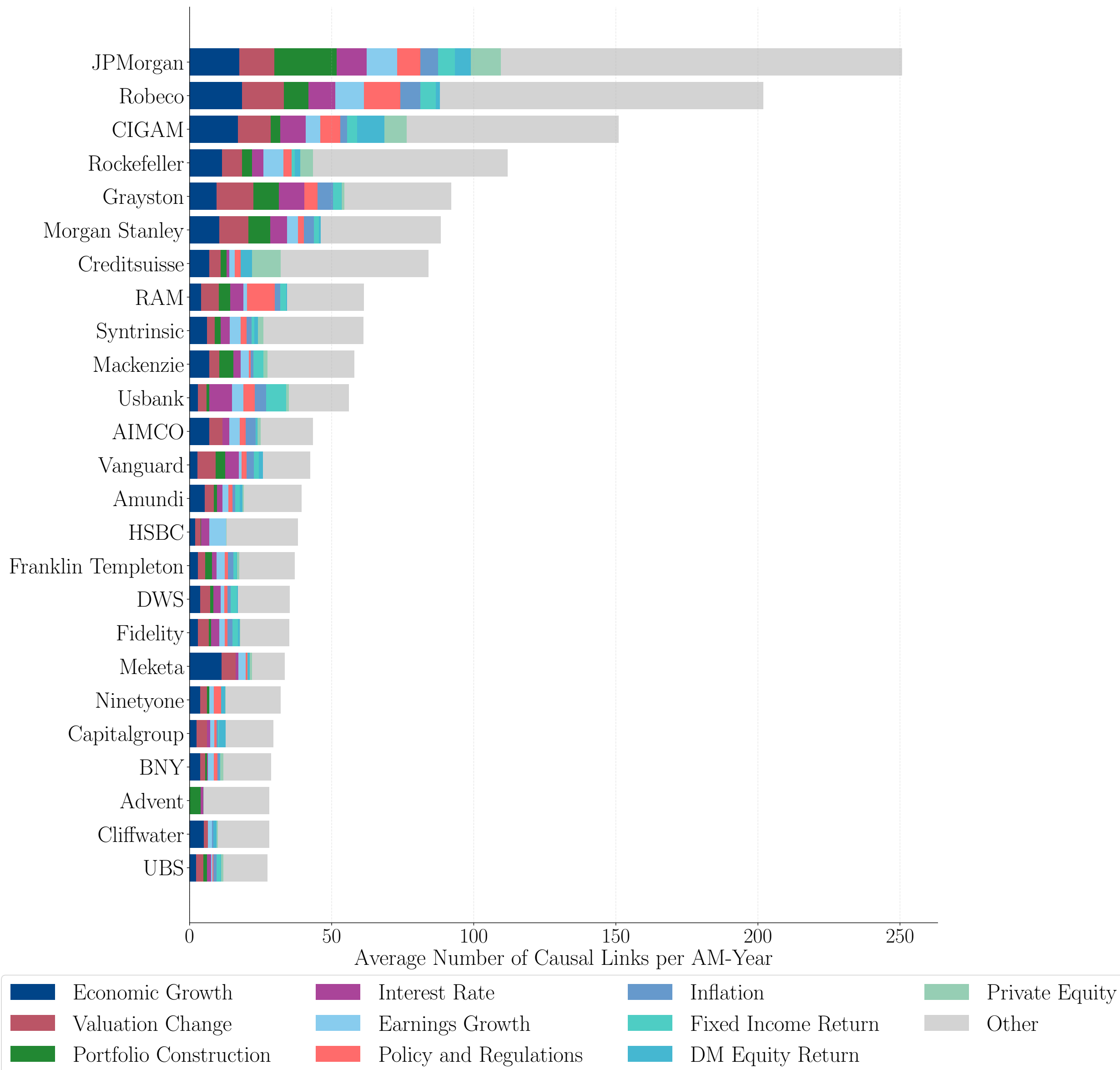


Figure 28: Cumulative Distribution of Letter–CMA Cosine Similarities. This figure shows the empirical cumulative distribution function of cosine similarities between the outlook section of each asset manager's N-CSR letter and the paragraphs of a contemporaneous capital-markets-assumption (CMA) report. Letter outlooks and CMA paragraphs are embedded with a pretrained sentence-transformer, and similarity is computed as the mean cosine similarity over the five highest-scoring paragraphs of a CMA report dated within six months of the letter. The blue curve plots the distribution over same-manager pairs, in which letter and CMA originate from the same asset manager; the grey curve plots the distribution over cross-manager pairs, in which they originate from different managers. The sample is restricted to asset managers contributing at least five unique letters, and similarities are winsorized at the 1st and 99th percentiles of the pooled distribution. Sample period: 2010–2026.

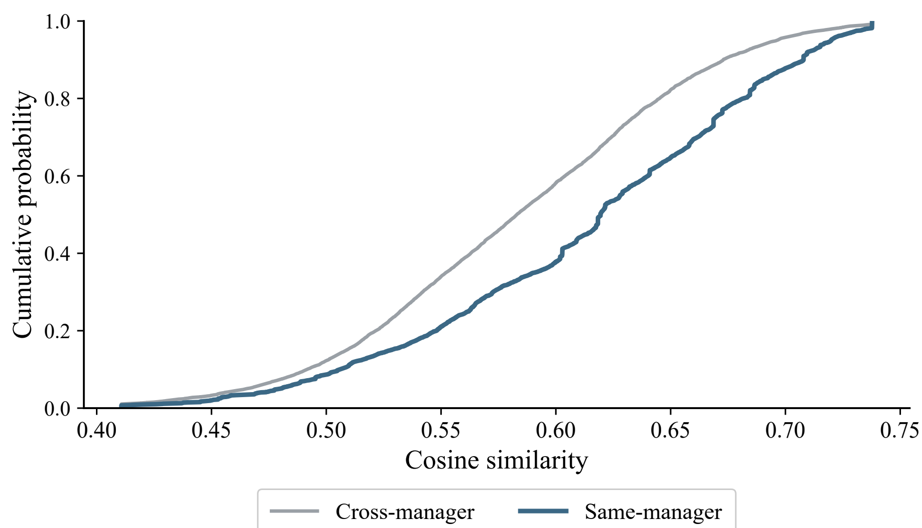


Table 1: CMA Institutional Investors. This table lists the 83 institutions in our Capital Market Assumptions dataset (59 asset managers and 24 consultants). The period column indicates the years of CMA coverage for each institution, depending on data availability and publication schedules. Sample period: 2008–2026.

Institution	Type	Years	Institution	Type	Years
ACG	Consultant	2023–2026	Mackenzie	Manager	2019–2023
Advent	Manager	2024–2024	Manulife	Manager	2020–2020
AIMCo	Manager	2021–2023	Meketa	Consultant	2020–2024
Allianz	Manager	2021–2025	Merrill Lynch	Consultant	2018–2018
American Century	Manager	2024–2025	MFS	Manager	2024–2026
Amundi	Manager	2018–2025	Morgan Stanley	Consultant	2019–2025
Angeles	Consultant	2018–2022	NEPC	Consultant	2016–2024
AON	Consultant	2012–2025	Neuberger	Manager	2022–2026
AQR	Manager	2013–2025	Ninety One	Manager	2023–2025
Baillie Gifford	Manager	2018–2018	Northern Trust	Manager	2013–2025
Barclays	Manager	2021–2025	Nuveen	Manager	2022–2025
BlackRock	Manager	2019–2024	OneAscent	Consultant	2024–2024
BNP	Manager	2020–2025	PGIM	Manager	2018–2025
BNY Mellon	Manager	2013–2025	PIMCO	Manager	2022–2024
Callan	Consultant	2016–2025	PineBridge	Manager	2024–2025
CalPERS	Manager	2021–2021	RAM	Manager	2020–2024
Candriam	Manager	2024–2025	RBC	Manager	2022–2025
Capital Group	Manager	2022–2025	Robeco	Manager	2019–2024
CIBC	Manager	2018–2025	Rockefeller	Consultant	2019–2020
Cliffwater	Consultant	2019–2026	RVK	Consultant	2021–2025
Cohen Steers	Manager	2022–2024	SEI	Consultant	2025–2025
Columbia Thread.	Manager	2017–2026	Schroders	Manager	2017–2026
Credit Suisse	Manager	2020–2020	Sellwood	Consultant	2012–2025
Crescent Capital	Manager	2020–2024	SJS	Consultant	2019–2020
CWO	Manager	2019–2020	State Street	Manager	2016–2025
Deutsche Bank	Manager	2023–2025	Syntrinsic	Consultant	2016–2025
DSA	Consultant	2011–2011	TD	Manager	2022–2025
DWS	Manager	2017–2025	T. Rowe Price	Manager	2019–2025
Edward Jones	Manager	2020–2025	UBS	Manager	2019–2023
EFG	Manager	2024–2024	Vanguard	Manager	2014–2025
Ellwood	Consultant	2017–2017	Verus	Consultant	2008–2025
Envestnet	Consultant	2013–2026	Voya	Manager	2016–2025
FI3	Manager	2022–2022	Wealthspire	Consultant	2023–2023
Fidelity	Manager	2020–2025	Wellington	Manager	2021–2025
Fiducient	Consultant	2021–2025	Wells Fargo	Manager	2023–2025
Franklin Temp.	Manager	2016–2025	Wilshire	Consultant	2021–2025
Graystone	Consultant	2021–2023	Benjamin F.E.	Manager	2024–2025
HSBC	Manager	2024–2025	BMO	Manager	2026–2026
Invesco	Manager	2018–2025	CIGAM	Manager	2024–2025
Janney	Manager	2022–2026	KKR	Manager	2025–2025
Janus Henderson	Manager	2020–2025	SIPA	Manager	2026–2026
JP Morgan	Manager	2009–2025			

Table 2: Asset Class Coverage and Return Expectations. This table shows coverage statistics for the 10 core asset classes in the CMA dataset. For each asset class: number of institutions providing forecasts, years of coverage, median forecast horizon, and average subjective return expectation and volatility. Sample period: 2008–2026.

Asset Class	No. of Inst.	Years	Horizon (median)	Avg. Subj. Exp.	Avg. Subj. Vol.
US Equity	80	2008 - 2026	10 yrs	6.28%	16.12%
Developed Equity	72	2008 - 2026	10 yrs	6.93%	17.19%
Emerging Equity	75	2008 - 2026	10 yrs	8.18%	22.24%
US Government Long Bond	25	2009 - 2025	10 yrs	2.89%	13.54%
US Government Bond	47	2008 - 2026	10 yrs	3.17%	5.70%
US IG Bond	45	2009 - 2026	10 yrs	3.93%	7.21%
US HY Bond	71	2008 - 2026	10 yrs	5.21%	9.74%
US TIPS	48	2008 - 2026	10 yrs	3.31%	6.04%
US Cash	67	2008 - 2026	10 yrs	2.52%	0.79%
US Inflation	46	2008 - 2026	10 yrs	2.31%	2.09%

Table 3: Causal Network Complexity and Topic Attention Statistics. This table reports summary statistics for causal network measures (described in Section 1.3) aggregated at the asset manager level. Panel (a) presents statistics for network complexity measures including indirect connection ratio, average shortest path length and network transitivity. Panel (b) presents statistics for topic-specific attention measures based on edge counts for valuation change/adjustment, economic growth, and dividend yield/income topics. For each measure, we report the mean, standard deviation, and coefficient of variation. Data are winsorized at the 1% and 99% levels before calculating statistics. Sample period: 2015–2025.

Panel (a): Network Complexity Measures

	Mean	Std. Dev.	Std/Mean
Indirect Ratio	0.102	0.092	0.90
Avg Shortest Path Length	2.518	0.731	0.29
Network Transitivity	0.048	0.042	0.88

Panel (b): Topic Attention Measures

	Mean	Std. Dev.	Std/Mean
Valuation Change Edge Count	6.205	6.249	1.01
Economic Growth Edge Count	6.867	8.489	1.24
Dividend Yield Edge Count	1.081	1.481	1.37

Table 4: Equity Building Blocks Pairwise Correlations. This table compares the pairwise correlations among reported CMA building-block components with the corresponding correlations in historical Shiller data. For each asset manager i, reference date t, forecast horizon h, and component $b \in \{g, \pi, DY, VC\}$, $F^b_{i,t,h}$ denotes the reported CMA forecast component: real growth (g), inflation (π), dividend yield (DY), or valuation change (VC). The subjective within-manager correlations are computed from these reported components after demeaning each component by asset manager, so that they capture within-manager co-movement across building blocks. The realised correlations are computed from the corresponding historical return-decomposition components in Shiller data. The Fisher z statistic tests each subjective correlation against its realised counterpart, while the Jennrich χ^2 statistic tests equality of the full correlation matrices. Significance in the subjective column is based on a manager-level block bootstrap; significance in the realised column is based on a circular block bootstrap of overlapping ten-year realised windows. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: subjective CMA panel, 2008–2026; realised Shiller windows, 1871–2014.

	Subj. within	Realised	Fisher z vs Real. 1871–2014	
Pair	2008-2026	1871–2014	z	p-value
DY, g	-0.255^{**}	-0.317^{**}	0.207	0.836
DY, π	-0.008	0.005	-0.038	0.970
DY, VC	0.155	-0.079	0.725	0.469
g, π	-0.178^{*}	-0.092	-0.271	0.786
g, VC	-0.411^{***}	0.308^{***}	-2.321^{**}	0.020
π, VC	0.279^{*}	-0.358^{**}	2.034^{**}	0.042
Jennrich χ^2 (6 d.o.f.) for Subj. within = Real. 1871–2014:			14.18^{**}, $p = 0.028$	

Table 5: Diebold-Mariano Test: CMA vs. Objective Model Forecasts. This table reports Diebold–Mariano test statistics comparing the predictive accuracy of asset managers' headline U.S. equity CMA forecasts, $F^{ER}_{i,t,h}$, with the objective machine-learning benchmark, $G_{t,h}$, described in Section 2.3.1. Forecast errors are computed relative to the annualized realized U.S. equity return $R_{t,t+h}$. The CMA forecast error is $e_{\text{CMA},i,t,h} = R_{t,t+h} - F^{ER}_{i,t,h}$, and the objective-model forecast error is $e_{\text{G},t,h} = R_{t,t+h} - G_{t,h}$. The Diebold–Mariano statistic is based on the squared-error loss differential $d_{i,t,h} = e^2_{\text{CMA},i,t,h} - e^2_{\text{G},t,h}$. Panel (a) reports results for our CMA reports only; Panel (b) reports results using CMA forecasts from Dahlquist & Ibert (2024). For each horizon, $h \in \{3, 5, 10\}$, and for the pooled specification, we report the number of paired observations (N), the root mean squared error (RMSE) for CMA and objective-model forecasts, the DM test statistic, and p-values. The pooled row stacks forecast pairs across horizons. A positive DM statistic indicates that the objective model has lower squared forecast errors than the CMA forecasts. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2011–2025.

Panel (a): CMA reports (Our Sample)

Horizon	N	RMSE_{CMA}	RMSE_{Obj}	DM Stat	p-value
3 years	13	12.40%	8.62%	1.99**	0.047
5 years	22	8.82%	4.90%	7.28***	<0.001
10 years	23	6.12%	2.73%	11.11***	<0.001
Pooled	66	8.79%	5.22%	4.95***	<0.001

Panel (b): CMA reports (Dahlquist & Ibert 2024)

Horizon	N	RMSE_{CMA}	RMSE_{Obj}	DM Stat	p-value
3 years	17	11.64%	7.61%	2.61***	0.009
5 years	35	9.57%	5.48%	13.26***	<0.001
10 years	29	6.92%	3.20%	7.75***	<0.001
Pooled	127	9.81%	4.92%	6.54***	<0.001

Table 6: Forecast Change Dynamics and Convergence Toward Consensus. This table reports panel regression estimates of forecast changes by asset managers. For each forecast variable $b \in \{ER, g, DY, VC, \pi\}$, where ER denotes the headline equity-return expectation and the remaining variables denote the reported CMA building-block components—real growth (g), dividend yield (DY), valuation change (VC), and inflation (π)—the dependent variable is $\Delta F^b_{i,t} = F^b_{i,t} - F^b_{i,t-1}$, the change in manager i's forecast for variable b. Columns (1)–(5) report results for ΔER, Δg, ΔDY, ΔVC, and $\Delta \pi$, respectively. The main regressor is the manager's lagged deviation from the leave-one-out consensus, $Dev^b_{i,t-1} = F^b_{i,t-1} - \bar{F}^b_{-i,t-1}$, where $\bar{F}^b_{-i,t-1}$ is the average forecast of all other managers in year $t-1$. The regressions also control for the contemporaneous change in the leave-one-out consensus, $\Delta\bar{F}^b_{-i,t} = \bar{F}^b_{-i,t} - \bar{F}^b_{-i,t-1}$, and for the lagged realized value associated with forecast variable b, X^b_{t-1}. All specifications include asset-manager fixed effects. Robust standard errors clustered by asset manager are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010–2025.

$$\Delta F^b_{i,t} = \alpha^b_i + \beta^b Dev^b_{i,t-1} + \delta^b \Delta\bar{F}^b_{-i,t} + \gamma^b X^b_{t-1} + \varepsilon^b_{i,t}$$

$\Delta F^b_{i,t}$	(1) ΔER	(2) Δg	(3) ΔDY	(4) ΔVC	(5) $\Delta \pi$
$Dev^b_{i,t-1}$	-0.692***	-0.486***	-0.580***	-0.935***	-0.876***
	(0.191)	(0.124)	(0.091)	(0.214)	(0.120)
$\Delta\bar{F}^b_{-i,t}$	0.884***	0.122	0.656***	0.893***	0.817***
	(0.138)	(0.159)	(0.141)	(0.135)	(0.162)
X^b_{t-1}	0.001	-0.001**	0.001	-0.001	0.000
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Asset Manager FE	Y	Y	Y	Y	Y
R^2 (within)	0.470	0.206	0.419	0.509	0.476
Observations	117	117	117	119	119

Table 7: CMA Forecast Anchoring. This table reports panel regression estimates testing whether individual CMA forecast deviations from the machine-learning benchmark (described in Section 2.3.1) are related to the distance between a salient consensus anchor and the same benchmark. The dependent variable is $G_{t,h} - F^{ER}_{i,t,h}$, where $F^{ER}_{i,t,h}$ is asset manager i's headline U.S. equity return forecast at report date t and horizon h, and $G_{t,h}$ is the out-of-sample machine-learning benchmark forecast for the same year and horizon. The main regressor in column (1) is $G_{t,h} - \bar{F}^{ER,365d}_{t-1,h}$, where $\bar{F}^{ER,365d}_{t-1,h}$ is the rolling consensus forecast constructed from horizon-matched headline U.S. equity CMA forecasts issued during the previous 365 days. Column (2) replaces the contemporaneous benchmark in the consensus deviation with the lagged benchmark, $G_{t-1,h} - \bar{F}^{ER,365d}_{t-1,h}$, while keeping the dependent variable $G_{t,h} - F^{ER}_{i,t,h}$. Both columns pool rounded horizons $h \in \{5, \ldots, 10\}$. All specifications include asset-manager fixed effects. Robust standard errors clustered by asset manager are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010–2025.

	Contemporaneous ML benchmark	Lagged ML benchmark
	(1)	(2)
Dependent variable	$G_{t,h} - F^{ER}_{i,t,h}$	$G_{t,h} - F^{ER}_{i,t,h}$
$G_{t,h} - \bar{F}^{ER,365d}_{t-1,h}$	0.745***	
	(0.090)	
$G_{t-1,h} - \bar{F}^{ER,365d}_{t-1,h}$		0.400***
		(0.082)
Asset Manager FE	Y	Y
R^2 (within)	0.232	0.076
Observations	450	450

Table 8: Forecast Changes and Earnings-News Sentiment. This table reports panel regression estimates of the relationship between annual changes in asset managers' forecasts and aggregate earnings-news sentiment. For the headline return expectation and each forecast component $b \in \{ER, g, DY, VC\}$, the dependent variable is $F^b_{i,t} - F^b_{i,t-1}$, the year-over-year change in asset manager i's forecast for variable b in year t. The variable ER denotes the headline U.S. equity return forecast, while g, DY, and VC denote the reported CMA building-block components: real growth, dividend yield, and valuation change, respectively. Columns (1) and (2) report results for the headline equity return forecast: column (1) uses the full available sample, and column (2) restricts the sample to observations for which all building blocks are available. Columns (3)–(5) report results for the building-block components g, DY, and VC. The key independent variable is S_t, the annual value-weighted S&P 500 RavenPack earnings-news sentiment measure, constructed as described in Section 1.2. All specifications include asset-manager fixed effects, μ^b_i. Robust standard errors clustered by asset manager are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010–2025.

$$F^b_{i,t} - F^b_{i,t-1} = \mu_i + \beta\, S_t + \varepsilon_{i,t}$$

$F^b_{i,t} - F^b_{i,t-1}$	(1) ΔER	(2) ΔER	(3) Δg	(4) ΔDY	(5) ΔVC
S_t	6.392***	7.698***	-1.351	0.265	10.152**
	(1.722)	(1.840)	(2.186)	(0.511)	(3.925)
Asset Manager FE	Y	Y	Y	Y	Y
R^2	0.225	0.229	0.303	0.206	0.220
Observations	185	113	113	113	89

Table 9: Portfolio Allocations and Return Expectations. This table reports regressions of U.S. equity allocations on CMA subjective returns and their building-block components (described in Section 2.1). The dependent variable $w_{f,t}$ is fund f's net U.S. equity allocation at time t. The explanatory variable $X_{i(f),t}$ is the CMA return forecast or building-block component of the asset manager $i(f)$ matched to fund f. In the decomposed specification, the CMA subjective return is decomposed into dividend yield ($DY_{i(f),t}$), nominal growth ($NG_{i(f),t}$), and valuation change ($VC_{i(f),t}$). Column (1) uses subjective excess returns across asset classes; column (2) uses U.S. equity return expectations only; column (3) uses the decomposed U.S. equity building blocks. The regressions include fund fixed effects (μ_f) and year fixed effects (τ_t). Standard errors, reported in parentheses, are clustered by fund and year. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010 to 2025.

$$w_{f,t} = \mu_f + \tau_t + \alpha + \boldsymbol{\beta}'\mathbf{X}_{i(f),t} + \varepsilon_{f,t}$$

	Asset Alloc U.S. Equity Net ($w_{f,t}$)		
$X_{i(f),t}$	(1)	(2)	(3)
Subj. Exc. Developed Eq.	-0.462**		
	(0.233)		
Subj. Exc. Emerging Eq.	-1.360***		
	(0.336)		
Subj. Exc. U.S. Eq.	0.598***		
	(0.225)		
Subj. U.S. Eq.		0.237***	
		(0.070)	
Dividend Yield			-0.362
			(0.286)
Nominal Growth			0.458***
			(0.145)
Valuation Change			0.237**
			(0.111)
Fund FE	Y	Y	Y
Year FE	Y	Y	Y
R^2	0.063	0.036	0.034
Observations	3251	25765	25765

Table 10: EWMA estimation of subjective volatility forecasts. This table reports nonlinear least-squares estimates of an exponentially weighted moving average model for asset managers' subjective U.S. equity volatility forecasts. Suppressing the forecast horizon for this exercise, the dependent variable is $F^{\sigma}_{i,t}$, asset manager i's subjective volatility forecast at CMA reference date t. The main regressor is $\text{EWMA}_t(\lambda)$, an exponentially weighted average of lagged quarterly realized volatility computed from daily S&P 500 returns over a 10-year lookback window. The decay parameter $\lambda \in [0, 1]$ is estimated endogenously and governs the persistence of memory in the averaging process The implied half-life reports the number of quarters required for the weight on a past realized-volatility observation to fall by one half. Robust standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010 to 2025.

$$F^{\sigma}_{i,t} = \alpha + \beta \cdot \text{EWMA}_t(\lambda) + \varepsilon^{\sigma}_{i,t}$$

$$\text{EWMA}_t(\lambda) = \sum_{\ell=1}^{L} \omega_\ell(\lambda)\, RV_{t-\ell}, \qquad \omega_\ell(\lambda) = \frac{\lambda^{\ell-1}}{\sum_{m=1}^{L} \lambda^{m-1}}$$

	$F^{\sigma}_{i,t}$
a (Intercept)	8.367***
	(1.512)
b (Slope)	0.484***
	(0.092)
γ (Decay)	0.981***
	(0.010)
Half-life	36.82 quarters
R^2	0.090
Observations	276

Table 11: Forecasting Techniques and Deviations from the Leave-One-Out Peer Consensus. This table shows how disclosed forecasting techniques relate to the signed deviation of an asset manager's forecast from the leave-one-out peer consensus. For the headline U.S. equity return forecast and each forecast component, $b \in \{ER, g, \pi, DY, VC\}$, the dependent variable is $Dev_{i,t}^b \equiv F_{i,t}^b - \overline{F}_{-i,t}^b$, where $F_{i,t}^b$ is asset manager i's forecast or building-block assumption for component b in year t, and $\overline{F}_{-i,t}^b$ is the leave-one-out peer consensus, computed as the average value of component b among all other asset managers in year t. Columns (1)–(5) report results for ER, g, π, DY, and VC, respectively. Manager-years for which no peer forecast is available are excluded. Regressions are estimated at the manager-year level. Methodology indicators are extracted and constructed from CMA report text using LLM-based classification. For each $k \in \mathcal{K} = \{Reg, Sim, Avg, MR, Trend, Oth\}$, $\mathbf{1}\{\text{technique}_{k,i,t}\}$ equals one if at least one report by asset manager i in year t is classified as disclosing technique k, and zero otherwise. The indicators are non-mutually exclusive. *Oth* equals one when none of the five named techniques is disclosed. Year fixed effects are included throughout; standard errors are clustered at the asset-manager level; deviations are winsorized at the 1st and 99th percentiles. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample: 2010 to 2025.

$$Dev_{i,t}^b \equiv F_{i,t}^b - \overline{F}_{-i,t}^b = \alpha^b + \tau_t^b + \sum_{k \in \mathcal{K}} \beta_k^b \mathbf{1}\{\text{technique}_{k,i,t}\} + \varepsilon_{i,t}^b.$$

$Dev_{i,t}^b$	(1) ER	(2) g	(3) π	(4) DY	(5) VC
$\mathbf{1}\{\text{technique}_{Reg,i,t}\}$	-0.002	0.007*	0.000	-0.004*	-0.003
	(0.003)	(0.004)	(0.001)	(0.003)	(0.003)
$\mathbf{1}\{\text{technique}_{Sim,i,t}\}$	0.001	0.004	0.000	0.002	-0.003
	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)
$\mathbf{1}\{\text{technique}_{Avg,i,t}\}$	0.003	0.001	-0.001	0.003*	0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
$\mathbf{1}\{\text{technique}_{MR,i,t}\}$	-0.004**	-0.001	-0.001	0.003	-0.005***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
$\mathbf{1}\{\text{technique}_{Trend,i,t}\}$	0.001	-0.001	-0.001	0.004	-0.005**
	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)
$\mathbf{1}\{\text{technique}_{Oth,i,t}\}$	0.001	-0.003	-0.001	0.008***	0.000
	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)
Year FE	Y	Y	Y	Y	Y
R^2	0.041	0.108	0.041	0.120	0.112
Observations	343	168	168	188	158

Table 12: The Seven Technique Clusters. This table defines the seven clusters of modeling techniques disclosed in the equities sections of CMA reports. Each of the 4,145 distinct modeling-technique phrases extracted from 589 reports is grouped, by the similarity of its meaning, into one of seven clusters; the clusters are labelled by their most representative phrases and ordered by diffusion. "% reports" is the share of the 589 reports using the cluster at least once and "Mentions" counts technique mentions (5,090 in total). Sample period: 2008–2026.

Cluster	Description	% reports	Mentions
Decomposition	General return decompositions.	83.2	1,056
Historical calibration	Calibration on historical averages, volatilities, and correlations.	71.5	770
Scenario adjustment	Probability-weighted scenario analysis under alternative macro paths, with qualitative overlays.	69.4	919
Present value framework	Equity-return framework based on income yield, real earnings or dividend growth, and valuation change.	65.0	718
Mean-reversion	Mean-reversion of valuation ratios (P/E, CAPE) toward historical norms.	64.7	584
Macro	Forward-looking forecasts of inflation and real GDP growth from consensus, breakevens, and surveys.	60.3	588
Yield-spread	Yield and credit-spread components.	39.4	455

Table 13: Technique Clusters and Deviations from the Leave-One-Out Peer Consensus, by Forecast Component. This table shows how disclosed modeling-technique clusters relate to managers' signed deviations from the leave-one-out yearly peer consensus, separately for the headline equity return forecast and each building-block component. For each component $b \in \{ER, g, DY, VC, \pi\}$, the dependent variable is $Dev^b_{i,t} \equiv F^b_{i,t} - \overline{F}^b_{-i,t}$, where $F^b_{i,t}$ is asset manager i's forecast or building-block assumption for component b in year t, and $\overline{F}^b_{-i,t}$ is the leave-one-out peer consensus, computed as the average value of component b among all other asset managers in year t. Columns (1)–(5) report results for ER, g, DY, VC, and π, respectively. Manager-years for which no leave-one-out peer consensus can be computed are excluded. Regressions are estimated at the manager-year level using manager-year observations that publish a building-block decomposition of the forecast. Technique-cluster indicators are extracted and constructed from CMA report text using LLM-based classification. Let $\mathcal{C} = \{Dec, Hist, Scen, PV, MR, Macro, YS\}$ denote the set of modeling-technique clusters, corresponding respectively to Decomposition, Historical calibration, Scenario adjustment, Present value framework, Mean-reversion, Macro, and Yield-spread. For each $k \in \mathcal{C}$, $\mathbf{1}\{\text{cluster}_{k,i,t}\}$ equals one if at least one report by asset manager i in year t uses a modeling phrase assigned to cluster k, and zero otherwise. The cluster indicators are non-mutually exclusive. Year fixed effects are included throughout; standard errors are clustered at the asset-manager level. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010–2026.

$$Dev^b_{i,t} \equiv F^b_{i,t} - \overline{F}^b_{-i,t} = \alpha^b + \tau^b_t + \sum_{k \in \mathcal{C}} \beta^b_k \mathbf{1}\{\text{cluster}_{k,i,t}\} + \varepsilon^b_{i,t}.$$

	$Dev^b_{i,t} \equiv F^b_{i,t} - \overline{F}^b_{-i,t}$				
	(1)	(2)	(3)	(4)	(5)
$Dev^b_{i,t}$	ER	g	DY	VC	π
$\mathbf{1}\{\text{cluster}_{Dec,i,t}\}$	-0.400	+0.163	-0.051	-0.381*	+0.011
	(0.285)	(0.177)	(0.131)	(0.195)	(0.071)
$\mathbf{1}\{\text{cluster}_{Hist,i,t}\}$	+0.610**	-0.276*	+0.302**	+0.456**	+0.049
	(0.247)	(0.147)	(0.153)	(0.182)	(0.039)
$\mathbf{1}\{\text{cluster}_{Scen,i,t}\}$	-0.126	+0.197	-0.217	-0.068	+0.046
	(0.214)	(0.160)	(0.146)	(0.168)	(0.047)
$\mathbf{1}\{\text{cluster}_{PV,i,t}\}$	+0.146	+0.290*	+0.109	-0.171	+0.029
	(0.248)	(0.165)	(0.210)	(0.194)	(0.053)
$\mathbf{1}\{\text{cluster}_{MR,i,t}\}$	-0.926***	+0.116	-0.060	-0.766***	-0.060
	(0.323)	(0.207)	(0.197)	(0.242)	(0.058)
$\mathbf{1}\{\text{cluster}_{Macro,i,t}\}$	+0.496	+0.226	+0.183	+0.175	+0.028
	(0.329)	(0.167)	(0.124)	(0.195)	(0.056)
$\mathbf{1}\{\text{cluster}_{YS,i,t}\}$	-0.283	-0.020	-0.122	+0.015	-0.075*
	(0.242)	(0.217)	(0.149)	(0.158)	(0.039)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.159	0.080	0.058	0.163	0.045
Observations	188	168	188	188	168

Table 14: Technique Clusters and Deviations from the Objective Model Forecast. This table shows how disclosed modeling-technique clusters relate to managers' signed deviations from the objective machine-learning benchmark. The dependent variable is $F^{ER}_{i,t,h} - G_{t,h}$, where $F^{ER}_{i,t,h}$ is asset manager i's headline U.S. equity return forecast at reference date t and horizon h, and $G_{t,h}$ is the objective machine-learning forecast (described in Section 2.3.1) matched to the same reference date and horizon. Regressions are estimated at the manager-year-horizon level. Technique-cluster indicators are extracted and constructed from CMA report text using LLM-based classification. Let $\mathcal{C} = \{Dec, Hist, Scen, PV, MR, Macro, YS\}$ denote the set of modeling-technique clusters, corresponding respectively to Decomposition, Historical calibration, Scenario adjustment, Present value framework, Mean-reversion, Macro, and Yield-spread. For each $k \in \mathcal{C}$, $\mathbf{1}\{\text{cluster}_{k,i,t}\}$ equals one if at least one report by asset manager i in year t uses a modeling phrase assigned to cluster k, and zero otherwise. The cluster indicators are non-mutually exclusive. Year fixed effects are included throughout; standard errors are clustered at the asset-manager level. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2010–2026.

$$F^{ER}_{i,t,h} - G_{t,h} = \alpha + \tau_t + \sum_{k \in \mathcal{C}} \beta_k \mathbf{1}\{\text{cluster}_{k,i,t}\} + \varepsilon_{i,t,h}$$

	$F^{ER}_{i,t,h} - G_{t,h}$
$\mathbf{1}\{\text{cluster}_{Dec,i,t}\}$	-0.539**
	(0.273)
$\mathbf{1}\{\text{cluster}_{Hist,i,t}\}$	+0.573**
	(0.223)
$\mathbf{1}\{\text{cluster}_{Scen,i,t}\}$	-0.164
	(0.201)
$\mathbf{1}\{\text{cluster}_{PV,i,t}\}$	+0.175
	(0.309)
$\mathbf{1}\{\text{cluster}_{MR,i,t}\}$	-0.630**
	(0.264)
$\mathbf{1}\{\text{cluster}_{Macro,i,t}\}$	+0.448*
	(0.257)
$\mathbf{1}\{\text{cluster}_{YS,i,t}\}$	-0.264
	(0.211)
Year fixed effects	Yes
R^2	0.369
Observations	179

Table 15: Equity Return Expectations and Network Complexity. This table presents regression results relating CMA U.S. equity return expectations to causal-network complexity using a within–between decomposition. The dependent variable is $F^{ER}_{i,t}$, asset manager i's headline U.S. equity return forecast in year t, measured at the asset-manager-year level. Each column corresponds to a different network complexity measure m: column (1) uses average shortest path length, column (2) uses the indirect connection ratio, and column (3) uses network transitivity. For each measure m, $C^m_{i,t}$ denotes the value of the causal-network complexity measure for asset manager i in year t, and $\overline{C}^m_i$ denotes manager i's time-series average complexity. The term $C^m_{i,t} - \overline{C}^m_i$ captures within-manager variation in complexity, while $\overline{C}^m_i$ captures the between-manager component. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$F^{ER}_{i,t} = \mu_i + \tau_t + \beta_w (C^m_{i,t} - \overline{C}^m_i) + \gamma_w (C^m_{i,t} - \overline{C}^m_i) \times S_t + \gamma_b \overline{C}^m_i \times S_t + \varepsilon_{i,t}$$

m	(1) APL	(2) ICR	(3) NT
$C^m_{i,t} - \overline{C}^m_i$	0.254	-0.289	-0.405
	(0.477)	(0.366)	(0.250)
$(C^m_{i,t} - \overline{C}^m_i) \times S_t$	17.884	-20.420	30.498*
	(20.137)	(16.419)	(17.043)
$\overline{C}^m_i \times S_t$	-17.387***	-19.692**	-23.476***
	(6.410)	(7.734)	(8.480)
Asset Manager FE	Y	Y	Y
Year FE	Y	Y	Y
Within R^2	0.220	0.210	0.139
R^2	0.761	0.771	0.762
Observations	173	173	173

Table 16: Equity Return Expectations and Topic Attention. This table presents regression results relating CMA U.S. equity return expectations to topic-specific attention in managers' causal networks using a within–between decomposition. The dependent variable is $F_{i,t}^{ER}$, asset manager i's headline U.S. equity return forecast in year t, measured at the asset-manager-year level. Each column corresponds to a different topic q: column (1) uses valuation change (VC), column (2) uses dividend yield (DY), column (3) uses economic growth (g), column (4) uses inflation (π), and column (5) uses economic downturn (ED). For each topic q, $A_{i,t}^{q}$ denotes asset manager i's topic-attention measure in year t, measured as the normalized number of causal connections to topic q, and $\overline{A}_i^q$ denotes manager i's time-series average attention to that topic. The term $A_{i,t}^{q} - \overline{A}_i^q$ captures within-manager variation in attention to topic q, while $\overline{A}_i^q$ captures persistent between-manager differences in attention to that topic. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) + \gamma_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) \times S_t + \gamma_B^q \overline{A}_i^q \times S_t + \varepsilon_{i,t}^q$$

q	(1) VC	(2) DY	(3) g	(4) π	(5) ED
$A_{i,t}^q - \overline{A}_i^q$	0.087	-0.117	-0.054	-0.177	-0.826*
	(0.126)	(0.384)	(0.130)	(0.306)	(0.439)
$(A_{i,t}^q - \overline{A}_i^q) \times S_t$	-15.259***	20.555	0.541	7.814	39.037*
	(5.697)	(22.060)	(5.279)	(13.519)	(20.378)
$\overline{A}_i^q \times S_t$	-12.452***	11.695**	-4.167	16.149	14.588***
	(4.145)	(5.285)	(3.444)	(10.827)	(5.497)
Asset Manager FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Within R^2	0.227	0.178	0.081	0.096	0.279
R^2	0.788	0.754	0.738	0.746	0.777
Observations	173	173	173	173	173

Table 17: Forecast Changes and Network Complexity. This table presents regression results relating year-over-year changes in CMA U.S. equity return expectations to causal-network complexity using a within–between decomposition. The dependent variable is $F^{ER}_{i,t} - F^{ER}_{i,t-1}$, the year-over-year change in asset manager i's headline U.S. equity return forecast. Each column corresponds to a different network complexity measure m: column (1) uses average shortest path length, column (2) uses the indirect connection ratio, and column (3) uses network transitivity. For each measure m, $C^m_{i,t}$ denotes the value of the causal-network complexity measure for asset manager i in year t, and $\overline{C}_i^m$ denotes manager i's time-series average complexity. The term $C^m_{i,t} - \overline{C}_i^m$ captures within-manager variation in complexity, while $\overline{C}_i^m$ captures persistent between-manager differences in complexity. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$F^{ER}_{i,t} - F^{ER}_{i,t-1} = \mu_i + \tau_t + \beta^m_W \left(C^m_{i,t} - \overline{C}_i^m\right) + \gamma^m_W \left(C^m_{i,t} - \overline{C}_i^m\right) \times S_t + \gamma^m_B \overline{C}_i^m \times S_t + \varepsilon^m_{i,t}$$

m	(1) APL	(2) ICR	(3) NT
$C^m_{i,t} - \overline{C}_i$	0.168	0.107	-0.061
	(0.312)	(0.858)	(0.201)
$(C^m_{i,t} - \overline{C}_i^m) \times S_t$	-34.238*	0.620	18.724**
	(19.414)	(39.876)	(8.916)
$\overline{C}_i^m \times S_t$	-22.240***	-24.636**	-27.306***
	(3.514)	(10.193)	(5.762)
Asset Manager FE	Y	Y	Y
Year FE	Y	Y	Y
Within R^2	0.158	0.133	0.101
R^2	0.551	0.554	0.525
Observations	122	119	122

Table 18: Forecast Changes and Topic Attention. This table presents regression results relating year-over-year changes in CMA U.S. equity return expectations to topic-specific attention in managers' causal networks using a within–between decomposition. The dependent variable is $F^{ER}_{i,t} - F^{ER}_{i,t-1}$, the year-over-year change in asset manager i's headline U.S. equity return forecast. Each column corresponds to a different topic q: column (1) uses valuation change (VC), column (2) uses dividend yield (DY), column (3) uses economic growth (g), column (4) uses inflation (π), and column (5) uses economic downturn (ED). For each topic q, $A^q_{i,t}$ denotes asset manager i's topic-attention measure in year t, measured as the normalized number of causal-network edges connected to topic q, and $\overline{A}^q_i$ denotes manager i's time-series average attention to that topic. The term $A^q_{i,t} - \overline{A}^q_i$ captures within-manager variation in attention to topic q, while $\overline{A}^q_i$ captures persistent between-manager differences in attention to that topic. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$F^{ER}_{i,t} - F^{ER}_{i,t-1} = \mu_i + \tau_t + \beta^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) + \gamma^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) \times S_t + \gamma^q_B \overline{A}^q_i \times S_t + \varepsilon^q_{i,t}$$

	(1)	(2)	(3)	(4)	(5)
q	VC	DY	g	π	ED
$A^q_{i,t} - \overline{A}^q_i$	0.189	-0.602	0.049	0.073	-0.589
	(0.190)	(0.808)	(0.186)	(0.129)	(0.895)
$(A^q_{i,t} - \overline{A}^q_i) \times S_t$	-27.413***	55.496*	12.069***	-17.518	59.476
	(8.210)	(28.861)	(1.215)	(16.514)	(47.013)
$\overline{A}^q_i \times S_t$	-10.822	15.192**	-9.460	24.976	18.144***
	(7.499)	(6.553)	(6.959)	(22.639)	(6.254)
Asset Manager FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Within R^2	0.174	0.092	0.114	0.098	0.138
R^2	0.590	0.581	0.526	0.525	0.589
Observations	122	122	122	122	122

Table 19: Ex-Ante Forecast Errors and Network Complexity. This table presents regression results examining the relationship between benchmark-relative ex-ante forecast errors and causal-network complexity using a within–between decomposition. The dependent variable is $G_{t,h} - F^{ER}_{i,t,h}$, where $G_{t,h}$ is the objective machine-learning forecast at reference date t and horizon h (described in Section 2.3.1), and $F^{ER}_{i,t,h}$ is asset manager i's headline U.S. equity return forecast matched to the same reference date and horizon. Each column corresponds to a different network complexity measure m: column (1) uses average shortest path length, column (2) uses the indirect connection ratio, and column (3) uses network transitivity. For each measure m, $C^m_{i,t}$ denotes the value of the causal-network complexity measure for asset manager i in year t, and $\overline{C}^m_i$ denotes manager i's time-series average complexity. The term $C^m_{i,t} - \overline{C}^m_i$ captures within-manager variation in complexity, while $\overline{C}^m_i$ captures persistent between-manager differences in complexity. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) + \gamma^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) \times S_t + \gamma^m_B \overline{C}^m_i \times S_t + \varepsilon^m_{i,t,h}$$

m	(1) APL	(2) ICR	(3) NT
$C^m_{i,t} - \overline{C}^m_i$	-0.428	0.460	0.252
	(0.437)	(0.315)	(0.243)
$(C^m_{i,t} - \overline{C}^m_i) \times S_t$	-14.180	18.954	-34.812*
	(20.162)	(15.955)	(17.621)
$\overline{C}^m_i \times S_t$	19.155***	21.656***	27.510***
	(6.566)	(8.002)	(9.361)
Asset Manager FE	Y	Y	Y
Year FE	Y	Y	Y
Within R^2	0.179	0.161	0.122
R^2	0.803	0.810	0.806
Observations	178	178	178

Table 20: Ex-Ante Forecast Errors and Topic Attention. This table presents regression results examining the relationship between benchmark-relative ex-ante forecast errors and topic-specific attention in managers' causal networks using a within–between decomposition. The dependent variable is $G_{t,h} - F^{ER}_{i,t,h}$, where $G_{t,h}$ is the objective machine-learning forecast at reference date t and horizon h (described in Section 2.3.1), and $F^{ER}_{i,t,h}$ is asset manager i's headline U.S. equity return forecast matched to the same reference date and horizon. Each column corresponds to a different topic q: column (1) uses valuation change (VC), column (2) uses dividend yield (DY), column (3) uses economic growth (g), column (4) uses inflation (π), and column (5) uses economic downturn (ED). For each topic q, $A^q_{i,t}$ denotes asset manager i's topic-attention measure in year t, measured as the normalized number of causal-network edges connected to topic q, and $\overline{A}^q_i$ denotes manager i's time-series average attention to that topic. The term $A^q_{i,t} - \overline{A}^q_i$ captures within-manager variation in attention to topic q, while $\overline{A}^q_i$ captures persistent between-manager differences in attention to that topic. The aggregate sentiment measure, S_t, is constructed as a one-year rolling average of the value-weighted S&P 500 RavenPack earnings-news sentiment index and centered around its time-series mean. All regressions include asset-manager fixed effects μ_i and year fixed effects τ_t. Robust standard errors clustered at the asset-manager level are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2015 to 2025.

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) + \gamma^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) \times S_t + \gamma^q_B \overline{A}^q_i \times S_t + \varepsilon^q_{i,t,h}$$

q	(1) VC	(2) DY	(3) g	(4) π	(5) ED
$A^q_{i,t} - \overline{A}^q_i$	0.023	0.032	-0.006	0.276	0.563
	(0.123)	(0.348)	(0.130)	(0.307)	(0.385)
$(A^q_{i,t} - \overline{A}^q_i) \times S_t$	14.399**	-19.143	0.599	-12.976	-35.698*
	(5.778)	(22.363)	(5.454)	(13.706)	(19.818)
$\overline{A}^q_i \times S_t$	12.955***	-13.308**	4.411	-17.570	-16.233***
	(4.527)	(5.529)	(3.493)	(10.806)	(5.830)
Asset Manager FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Within R^2	0.125	0.141	0.058	0.060	0.200
R^2	0.820	0.796	0.781	0.792	0.813
Observations	178	178	178	178	178

Appendix A

A.1 LDA Textual Data Cleaning and Document-Term Matrix Construction

Latent Dirichlet Allocation (LDA) is a widely used technique for discovering latent themes in textual data by grouping words into topics based on their co-occurrence patterns (Blei et al. 2003). This approach enables a numerical representation of the text and facilitates the extraction of narratives, aligning with the methodology of Bybee et al. (2023).

A.1.1 Text preprocessing

We prepare the CMA report text for topic modeling in six steps. First, we tokenize each report into alphabetic unigrams of at least two characters and lowercase. Second, we apply rule-based lemmatization (suffix stripping for *-sses*, *-ies*, *-s*, *-ly*, *-ed*, *-ing*). Third, we remove standard English stopwords together with a custom list of financial and document-structure terms. Fourth, we drop words that appear in fewer than 0.1% of documents. Fifth, we form bigrams and then trigrams with Gensim's *Phrases* module, using `min_count` = 5 and `threshold` = 10. Sixth, we strip asset manager identifiers from the corpus at both the unigram and bigram level. The bigram pass catches manager-name fragments that lemmatization occasionally produces and that a unigram-only strip would miss. The final corpus comprises 510 reports.

We acknowledge one limitation. Eight of the recovered topics capture regulatory and disclaimer language, fund prospectus templates, and similar standardized text. We tested aggressive disclaimer-section stripping at the PDF extraction stage and rejected it: it lowered overall coherence and surfaced tokenization artifacts in the remaining topics. We flag these topics explicitly rather than risk biasing the substantive topic estimates.

A.1.2 Latent Dirichlet Allocation for Narratives

We evaluate $K \in [10, 150]$ using three-fold cross-validation with three random seeds per fold, yielding nine repetitions per K. We report three metrics that capture different aspects of model fit: held-out log-perplexity, which measures predictive likelihood on documents not seen during estimation; c_v topic coherence, which measures whether the top-weighted words within each topic actually co-occur in the underlying documents; and UMass coherence, a log-conditional-probability variant of the same idea. The three metrics capture genuinely distinct dimensions of model quality: predictive fit on unseen documents, semantic coherence under a sliding-window co-occurrence measure, and semantic coherence under a global co-occurrence measure.

[Figure A1 here]

Figure A1 reports the three metrics across K. Topic coherence c_v exhibits an interior plateau on $K \in [30, 90]$, within which the difference between candidate values is smaller than the seed-to-seed standard deviation. Held-out log-perplexity declines monotonically with K, a pattern characteristic of small corpora where the marginal-likelihood penalty dominates as K grows. We therefore rely on the coherence criterion within the plateau rather than on perplexity in isolation, we choose $K = 55$ within the plateau on the basis of the t-SNE projection of the document-topic distribution.

A.2 LLM and Text Analysis

We develope a structured interface for using Anthropic's API to process CMA reports. Our implementation divides the analysis into distinct subtasks, each handled through a separate API call. The Online Appendix provides the full prompt-engineering design.

A.2.1 Text preprocessing

Our text preprocessing pipeline for LLM analysis follows the same general text extraction stage described in Appendix A.1 and does not require additional preprocessing steps beyond the initial filtering and cleaning.

A.2.2 LLM Pipeline Architecture

We construct causal networks from financial reports through a three-stage natural language processing pipeline. Our approach extracts explicit causal relationships, classifies agents, and determines effect sentiment to create directed networks representing causal reasoning in asset manager communications.

Causal relationship extraction. We identify explicit causal relationships within report texts using Claude Sonnet 4.5 via batch API processing. The model analyzes each text segment with surrounding context for pronoun resolution and extracts causal triplets: (i) causal agent, (ii) effect, and (iii) affected agent. Each relationship receives a confidence score based on marker clarity and relationship explicitness. The batch processing system adaptively groups text segments into API requests to efficiently handle documents of varying lengths, with prompt caching for repeated instructions.

Topic classification. We assign financial topics and subtopics to each causal and affected agent using a predefined taxonomy covering major financial domains. The classification leverages both sentence-level context (where the agent appears) and broader context to disambiguate agent meanings. Context grouping optimization ensures that agents sharing identical sentence and full text contexts are processed together, avoiding redundant transmission of duplicate contexts. Each agent occurrence is classified independently, receiving a topic assignment with a match score indicating classification confidence. Agents matching no predefined category are classified as "other".

Effect sentiment analysis. We determine the directional impact of each causal relationship by analyzing the effect description as the primary signal. Effects containing terms such as "increase," "improve," "boost," or "enhance" receive positive sentiment; those with "decrease," "reduce," "weaken," or "damage" receive negative sentiment; effects with unclear directionality (e.g., "affect," "influence") or offsetting impacts receive neutral classification. Each sentiment assignment includes a strength score based on language intensity in the effect description, relationship directness, and certainty indicators.

A.2.3 Prompt Design

We use pretrained models without fine-tuning with carefully designed prompts to improve task performance (Gao 2023).

Our prompts employ a step-by-step structure informed by chain-of-thought methodology (Wei et al. 2022, Fu et al. 2023). This approach decomposes complex analytical tasks into discrete sequential steps within the prompt, enhancing logical coherence and reducing ambiguity, though overly rigid task decomposition can limit adaptability. Each prompt includes a self-consistency verification component, instructing the model to review its outputs before finalizing responses. This technique improves accuracy with minimal computational overhead (Wang et al. 2022).

We incorporate few-shot examples within each prompt to demonstrate expected output formats and reasoning depth (Brown et al. 2020). This approach allows the model to match fine-tuned performance without parameter updates, while the natural-language format of examples enables rapid adaptation to new tasks (Zhao et al. 2021). Our prompts explicitly direct the model to rely only on the provided text, avoiding external knowledge. This approach reduces hallucination risks and ensures responses are based exclusively on the provided text content, though it may limit the model's ability to resolve ambiguous references or fill contextual gaps. We configure API parameters to encourage deterministic and precise outputs. A temperature setting of 0.2 prioritizes accuracy over response variability. These configurations align with established practices for high-precision analytical tasks (Vaswani et al. 2017).

A.3 CMA Reports Text

We examine the evolution of Capital Market Assumption (CMA) reports by analyzing their textual characteristics, using word count as a proxy for report length and depth. Figure A2 summarizes these patterns. Panel (a) shows substantial variability in CMA report lengths over time, with no clear trend in the median until a sharp increase in 2025. However, the spread of word counts widens in later years, indicating increasing heterogeneity in report length Panel (b) highlights considerable variation in typical report length across asset management firms, pointing to differences in communication strategies.

[Figure A2 here]

Figure A3 visualizes the most frequent titles, reference authors and journals, and years using word clouds extracted from the report text.

[Figure A3 here]

A.4 Sentiment Word Clouds

A.4.1 Aggregated Sentiment Keywords

We applied the LLM framework from Section 1.3 to measure sentiment toward valuation ratios, interest rates, GDP growth, and earnings. We then collect the keywords that drove these ratings, figures A4 through A7 show word clouds for each factor. Larger words appear more frequently across our dataset, highlighting the most frequently used descriptive words.

These visualizations highlight the specific terminology and phrasing used to convey optimistic or pessimistic market signals. Figure A4 shows valuation language with negative terms like ”overvalued” and ”expensive” versus positive descriptors like ”cheap” and ”undervalued.” Figure A5 captures monetary policy signals through negative keywords such as ”tightening” and ”hikes” compared to positive terms like ”accommodative” and ”supportive.” Figure A6 reflects economic conditions with negative words including ”recession” and ”contraction” against positive terms like ”robust” and ”expanding.” Finally, Figure A7 highlights corporate performance language, contrasting negative descriptors such as ”declining” and ”disappointing” with positive terms like ”growing” and ”strong.”

In future iterations, we will build on these insights to better understand the heterogeneity in asset managers' communication strategies when conveying their beliefs to financial practitioners.

[Figures A4, A5, A6, and A7 here]

A.4.2 Event Sentiments

We present three word clouds that visualize the sentiment derived from the extracted events. Figure A8 groups sentiment keywords according to three overall sentiment topics: Positive/Very Positive, Neutral/NaN, and Negative/Very Negative.

[Figure A8 here]

A.5 CMA Report Authors

We assemble a panel of biographical information on the professionals listed as authors of each CMA report in our sample. While our quantitative dataset captures the forecasts and decompositions that institutions publish, the author panel records the human inputs behind these forecasts and allows us to examine whether forecaster characteristics map into systematic differences in reported expectations.

We construct the panel as follows. For each CMA report in our document corpus, we identify the named contributors from the cover page, the methodology section, or the publication credits when available. We then locate each contributor on LinkedIn and record their professional background. For every author–report observation we capture: gender, country of origin (inferred from native language, high-school location, or biographical detail), highest academic degree, current job title at the time of report publication, and total years of professional experience measured to the report date. Job titles are harmonized into a nine-level seniority scale ranging from Analyst to CEO, and educational attainment is harmonized into the standard sequence of Bachelor, Master/MBA, and PhD.

The resulting panel contains 2,088 author–report observations spanning 260 unique CMA reports, 39 distinct asset managers, and the period 2005–2024. Each report is associated with one or more contributors, so the panel preserves both the team composition of individual publications and the cross-sectional variation in author profiles across institutions and over time.

Figure A10 summarizes the panel along four dimensions. Panel (a) reports the distribution of job seniority across all author–report observations. Authors are concentrated in senior decision-making roles: roughly 30% are Chief Investment, Risk, or Financial Officers, and a further 21% are Managing Directors, indicating that CMA reports are typically signed by individuals with senior portfolio or research responsibility. At the same time, 18% of observations correspond to Analysts, consistent with a production model in which junior researchers contribute alongside senior authors. The remaining mass is spread across Vice Presidents, Senior Managers, Managers, and a small number of Presidents and CEOs.

Panel (b) shows the distribution of educational attainment. Master and MBA degrees are the most common credential at 47%, followed by Bachelor degrees at 41% and PhDs at 12%.

Panel (c) plots the distribution of years of professional experience at the time of report publication. The distribution is approximately bell-shaped and centered between 15 and 25 years, with a clear mode near 20 years. Few authors have fewer than five years of experience or more than 35 years, so CMA report production is concentrated among mid-to-late-career professionals. Taken together, the three panels paint a coherent picture: institutional CMA documents are produced by experienced professionals with applied graduate training and senior firm responsibility, with a meaningful share of junior analytical contributors.

Panel (d) reports the distribution of authoring-team size, measured as the number of listed contributors per CMA report. The distribution is strongly right-skewed: most reports are produced by relatively small teams, with the largest mass concentrated below ten authors. However, the distribution has a long right tail, with a small number of reports listing more than 20 contributors and a few cases exceeding 50 authors. This pattern suggests that while CMA reports are typically authored by compact research or investment teams, some institutions attribute reports to much broader committees or cross-functional groups.

Taken together, the four panels paint a coherent picture: institutional CMA documents are produced by experienced professionals with applied graduate training and senior firm responsibility, typically organized in teams with meaningful variation in size across institutions.

We note that while the aggregate distributions describe the typical author profile in our sample, the educational composition of authoring teams varies markedly across institutions.

Figure A11 reports the resulting manager-level distribution. This dispersion indicates that the human-capital composition behind CMA forecasts differs systematically across institutions.

[Figure A10 and Figure A11 here]

A.6 Subjective Return Expectations Across Asset Classes

Figure A9 presents return expectations across Developed Market Equity, Emerging Market Equity, U.S. Investment Grade, U.S. High Yield, U.S. Government Bond, and U.S. Cash.

[Figure A9 here]

A.7 RavenPack Data Processing

RavenPack provides comprehensive structured metadata that includes sentiment scores, event classifications, and relevance indicators for news stories across multiple financial and economic domains. We employ two primary sentiment indicators in our analysis. The Event Sentiment Score (ESS) rates each firm-level news event on a -1 to +1 scale based on experts with extensive experience and backgrounds in linguistics, finance, and economics. The Composite Sentiment Score (CSS) indicates how the market responds to news articles. The CSS variable is estimated based on stock price reactions, which are empirically modeled using intraday data from a portfolio of approximately one hundred large-cap stocks (RavenPack 2020, Dang et al. 2015).

We categorize news into various topics using alternative classifications based on the variables *GROUP* and *TYPE*, allowing for different levels of granularity in topic analysis. To mitigate the risk of overcounting due to repeated coverage of the same events by multiple news outlets, we introduce a mandatory one-day gap to classify stories as distinct new events (i.e., setting EVENT SIMILARITY DAYS ≥ 1). To ensure the selected news has substantial implications for the companies under study, we apply two relevance thresholds simultaneously: entity relevance must exceed 75, and event relevance must surpass 90. Additionally, we exclude social media platforms and low-quality news sources to preserve data integrity and emphasize professionally curated content. These filtering criteria follow established practices documented in the literature and RavenPack's user guidelines (RavenPack 2020, Dang et al. 2015, Gao et al. 2017, Hafez 2009).

In constructing news horizons aligned with the CMA report date, we aggregate news based on their publication dates at the daily frequency, following equations (1) or (2). Subsequently, using the resulting daily time series, we calculate average news measures over specified time intervals. We focus on news about S&P 500 companies included in CRSP monthly market cap data. We first aggregate news at the parent company level, then perform fuzzy matching between company names in RavenPack and CRSP.

A.8 Additional Details on the Building-Block Decomposition

Asset managers' return forecasts employ a "building-block" decomposition that breaks expected equity returns into fundamental drivers. This approach parallels the logic of the dividend discount model and has been widely adopted by both practitioners and academics. To keep notation consistent with the main text, let $F^{ER}_{i,t,h}$ denote asset manager i's headline U.S. equity return forecast at CMA reference date t and horizon h. In its simplest form, the expected long-term equity return can be approximated as the sum of expected real growth, expected inflation, and the initial income yield. Formally, one can write:

$$F^{ER}_{i,t,h} \approx F^{g}_{i,t,h} + F^{\pi}_{i,t,h} + F^{DY}_{i,t,h},$$

where $F^{g}_{i,t,h}$ denotes the real-growth component, $F^{\pi}_{i,t,h}$ expected inflation, and $F^{DY}_{i,t,h}$ the income or dividend-yield component.

To capture additional drivers of returns, more elaborate decompositions include terms for share repurchases, net issuance, profit-margin mean reversion, and valuation changes. Accordingly, we define the net distribution yield component as

$$F^{DY}_{i,t,h} = F^{DY,div}_{i,t,h} + F^{DY,buyback}_{i,t,h} - F^{DY,issuance}_{i,t,h},$$

which augments the traditional dividend-yield component with buyback payouts and subtracts the dilution from net share issuance.

Similarly, if current profit margins are elevated relative to historical norms, one can include a margin-adjustment component, $F^{margin}_{i,t,h}$, to capture expected profit-margin reversion. When margin adjustments are reported separately, we write the real-growth block as

$$F^{g}_{i,t,h} = F^{g,adj}_{i,t,h} + F^{margin}_{i,t,h},$$

where $F^{g,adj}_{i,t,h}$ is the real-growth component net of margin effects.

A more comprehensive additive model for expected returns is therefore:

$$F^{ER}_{i,t,h} \approx \underbrace{\left(F^{g,adj}_{i,t,h} + F^{\pi}_{i,t,h}\right)}_{\text{Growth and Inflation}} + \underbrace{F^{margin}_{i,t,h}}_{\text{Margin Adjustment}} + \underbrace{F^{DY}_{i,t,h}}_{\text{Net Distribution Yield}} + \underbrace{F^{VC}_{i,t,h}}_{\text{Valuation Change}} . \tag{A1}$$

Equivalently, when reports combine real growth and inflation, we denote nominal growth by

$$F^{NG}_{i,t,h} = F^{g}_{i,t,h} + F^{\pi}_{i,t,h}.$$

Each component can be adjusted or grouped depending on the forecaster's reporting convention. For example, some forecasters subsume margin reversion into an adjusted growth term, while others combine dividends, buybacks, and issuance into a single net-yield component. In all cases, the building-block decomposition provides a structured framework that ensures all key drivers of long-run equity returns are considered. It is worth noting that one could also express the decomposition in a multiplicative form to account for compounding interactions among factors.

Institution-specific building blocks. Figure A12 summarizes the components included by each asset manager in their CMA equity-return decompositions. Each colored bar denotes one of six reported building blocks—nominal growth, dividend yield, valuation change, buybacks, margin adjustments, and share issuance. These reported blocks are subsequently harmonized into the common taxonomy used in the main analysis: real growth (g), inflation (π), dividend yield or net distribution yield (DY), and valuation change (VC), with nominal growth (NG) defined as real growth plus inflation when needed.

[Figure A12 here]

A.9 A Practical Application to CMAs: AQR 2017 CMA

As discussed in Section 2.1, many asset managers rely on building-block approaches that can be motivated by approximate Campbell–Shiller decompositions. For a generic asset manager i, these decompositions can be written in additive form as

$$\begin{aligned}\widetilde{\mathbb{E}}_{i,t}^{CMA}[R_{t+1}] &= \widetilde{\mathbb{E}}_{i,t}\left[DY_{div,t+1} + DY_{buyback,t+1} - DY_{issuance,t+1}\right] \\ &\quad + \widetilde{\mathbb{E}}_{i,t}\left[G_{px,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[G_{real\,adj.\,x,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[G_{margin,t+1}\right] + \widetilde{\mathbb{E}}_{i,t}\left[\pi_{t+1}\right]. \quad \text{(A2)}\end{aligned}$$

where

$$\widetilde{\mathbb{E}}_{i,t}[G_{real\,x,t+1}] = \widetilde{\mathbb{E}}_{i,t}[G_{real\,adj.\,x,t+1}] + \widetilde{\mathbb{E}}_{i,t}[G_{margin,t+1}].$$

In a similar spirit, the AQR (2017) CMA report describes a methodology, largely maintained in subsequent reports, that starts from a dividend-discount model approximated as

$$\widetilde{\mathbb{E}}_{t}^{AQR}[R_{t+1}] \approx DY_t^{AQR} + \widetilde{\mathbb{E}}_{t}^{AQR}[g_{t+1}] + \widetilde{\mathbb{E}}_{t}^{AQR}[VC_{t+1}], \quad \text{(A3)}$$

where DY_t^{AQR} is the dividend yield, interpreted broadly as the payout to shareholders, $\widetilde{\mathbb{E}}_{t}^{AQR}[g_{t+1}]$ is expected real dividend or earnings-per-share growth, and $\widetilde{\mathbb{E}}_{t}^{AQR}[VC_{t+1}]$ captures the expected valuation-change component. AQR assumes no mean reversion in valuations, so that

$$\widetilde{\mathbb{E}}_{t}^{AQR}[VC_{t+1}] = 0,$$

implying

$$\widetilde{\mathbb{E}}_{t}^{AQR}[R_{t+1}] = DY_t^{AQR} + \widetilde{\mathbb{E}}_{t}^{AQR}[g_{t+1}]. \quad \text{(A4)}$$

Because dividend yield alone may not capture the full payout to shareholders when repurchases are important, AQR constructs two complementary estimators of expected returns and takes their average. The first estimator, denoted $\widehat{r}_t^{(E)}$, is based on earnings yields. Using the inverse of the Shiller CAPE ratio, denoted X_t/P_t, AQR approximates the dividend yield by assuming a constant payout ratio of $D/X = 50\%$. Combining this with a constant long-run real earnings growth rate $\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}]$ yields the earnings-yield-based estimate

$$\widehat{r}_t^{(E)} = \frac{D}{X} \cdot \frac{X_t}{P_t} + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}]. \tag{A5}$$

For developed markets, AQR sets $\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}] = 1.5\%$, corresponding to the long-run historical real growth rate of earnings per share.

The second estimator, denoted $\widehat{r}_t^{(P)}$, modifies the dividend-discount model to incorporate total shareholder payouts. Instead of using dividend yield alone, AQR uses the net total payout yield, defined as

$$NTY_t = Div.Yield_t + NBY_t,$$

where NBY_t denotes the net buyback yield, i.e., buybacks net of issuance. Expected returns are then expressed as

$$\widehat{r}_t^{(P)} = NTY_t + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}]. \tag{A6}$$

where $\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}]$ denotes the expected growth rate of real aggregate total payouts. This growth term is estimated as the average of two measures. The first is a bottom-up estimate based on long-run historical real EPS growth adjusted for dilution,

$$\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GP}] = g_t^{EPS} + \delta^{dil},$$

where δ^{dil} denotes long-run equity dilution, approximately 1.6%. The second is a top-down estimate based on long-run forecasts of real aggregate GDP growth, $\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GDP}]$. The expected growth of aggregate total payouts is therefore

$$\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}] = \frac{\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GP}] + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GDP}]}{2}. \tag{A7}$$

Finally, AQR obtains its expected real equity return by averaging the earnings-yield-based and payout-based estimates:

$$\widetilde{\mathbb{E}}_t^{AQR}[R_{t+1}] = \frac{1}{2}\left(\widehat{r}_t^{(E)} + \widehat{r}_t^{(P)}\right). \tag{A8}$$

Substituting the expressions above yields the following representation of expected real equity returns in the AQR framework:

$$\begin{aligned}
\widetilde{\mathbb{E}}_t^{AQR}[R_{t+1}] &= \frac{1}{2}\left[\left(\frac{D}{X}\cdot\frac{X_t}{P_t} + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}]\right) + \left(NTY_t + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}]\right)\right] && \text{(A9)}\\
&= \underbrace{\frac{1}{2}\left[\frac{D}{X}\cdot\frac{X_t}{P_t} + NTY_t\right]}_{DY_t^{AQR}} + \underbrace{\frac{1}{2}\left[\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}] + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}]\right]}_{\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}]}. && \text{(A10)}
\end{aligned}$$

This formulation combines an earnings-yield anchor with a payout-based dividend-discount model to produce a long-run estimate of expected equity returns. We can represent the resulting set of relationships as a network, as shown in Figure A13.

[Figure A13 here]

A.10 Causal Networks

Each Capital Market Assumption report generates a directed graph $G = (V, E)$ where nodes $v \in V$ represent economic topics and directed edges $(u, v) \in E$ represent causal relationships from topic u to topic v extracted via the LLM pipeline described in Section 1.3.

A.10.1 Network Visualizations

Figure A14 presents complete causal networks for four representative asset managers, constructed by aggregating all available CMA reports for each institution across our sample period.

[Figure A14 here]

Table A1 reports the average number of edges and nodes per report for each asset manager in our dataset.

[Table A1 here]

A.10.2 Network Measures

Average Path Length: Mean shortest path length between node pairs in the largest connected component:

$$\text{Avg. Path Length} = \frac{1}{|P|} \sum_{(i,j)\in P} d(i,j) \tag{A11}$$

where $d(i,j)$ is the shortest path length from node i to node j, and P is the set of node pairs in the largest connected component. The strong component version considers directed paths, while the weak component version considers undirected paths.

Indirect Connection Ratio: Proportion of edges involving focal topics that participate in multi-step causal chains to U.S. equity return:

$$\text{Indirect Ratio} = \frac{|\mathcal{E}indirect|}{|\mathcal{E}total|} \tag{A12}$$

where $\mathcal{E}indirect$ is the set of edges involving the focal topics that are part of indirect causal paths leading to (or from) U.S. equity return, and $\mathcal{E}total$ is the total number of edges in the network. This measure captures the extent to which these topics affect U.S. equity through multi-step causal chains rather than direct connections.

Network Transitivity: Global clustering coefficient measuring the tendency for triangles to form:

$$\text{Transitivity} = \frac{3 \times \text{number of triangles}}{\text{number of connected triples}} \tag{A13}$$

where a triangle consists of three nodes with edges $(A \to B, B \to C, A \to C)$, and a connected triple has at least two edges $(A \to B, B \to C)$. Values range from 0 (no transitive closures) to 1 (complete transitivity).

A.11 Objective Model

This appendix details the construction of the objective benchmark $G_{t,h}$ introduced in Section 2.3.1. Our predictor set consists of the 16 canonical Welch & Goyal (2008) variables, drawn from the annual Goyal et al. (2024) dataset and listed in Table A2. This is the same set of predictors used by Shen & Xiu (2025) and Giannone et al. (2021). All predictors are lagged by one year relative to the forecast start date to avoid contemporaneous look-ahead bias.

[Table A2 here]

The realised target is the annualised geometric total return:

$$r_{t,t+h} = \left(\frac{P_{t+h}}{P_t}\right)^{1/h} - 1 \tag{A14}$$

where P_t is the cum-dividend total-return price index at the end of year t.

The objective forecast $G_{t,h}$ is the equal-weighted average of $M = 7$ regularised return-forecasting models:

$$G_{t,h} = \frac{1}{M}\sum_{m=1}^{M} \max\left(0, \hat{r}_{m,t,t+h}\right) \tag{A15}$$

where the Campbell & Thompson (2008) non-negativity constraint is imposed on each base forecast before averaging. The seven base models are Ridge, Lasso, Elastic Net, Random Forest, Gradient Boosting Regression Trees, and two single-hidden-layer feed-forward neural networks: one with L_2 weight decay and one with an L_1 subgradient penalty. We use equal weights rather than estimated weights because, in small samples with multiple models, optimised weight vectors tend to overfit and underperform simple averages.

Each base model is trained on an expanding window with a minimum of $T_{\min} = 30$ training observations. Hyperparameters are tuned using forward-only expanding-window 5-fold cross-validation within the training window, and features are standardised within each training window only.

Figure A1: Evaluation Metrics for LDA Topic Selection: Held-out Log-Perplexity and Coherence. The figure plots three metrics used to assess the quality and interpretability of LDA models estimated on CMA reports. Held-out log-perplexity (left) measures out-of-sample predictive fit (lower is better); topic coherence C_v (middle) and topic coherence UMass (right) measure how interpretable or semantically consistent the extracted topics are. We evaluate LDA models with the number of topics ranging from 10 to 150. Both held-out perplexity and cross-validation coherence criteria suggest that $L = 55$ yields the best overall balance among these metrics. Sample: CMA reports, 2012–2025.

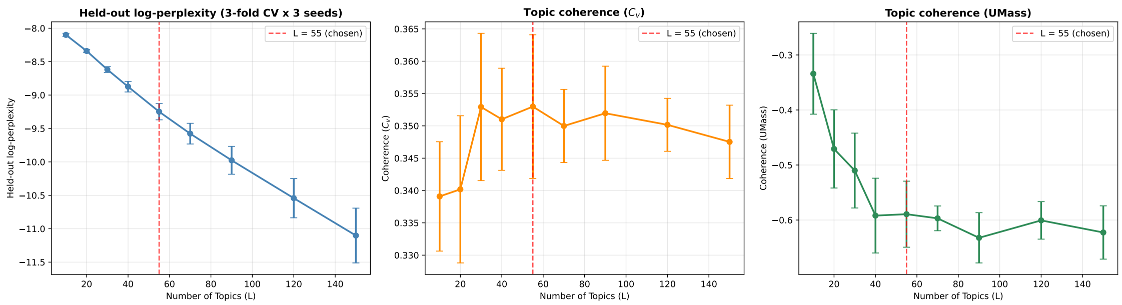


Figure A2: Evolution of CMA Report Textual Characteristics. This figure shows the evolution of CMA report lengths measured by word count. Panel (a) displays the temporal evolution of CMA report lengths measured by word count from 2015 to 2025. Panel (b) shows the histogram of word count distribution across all CMA reports in the sample, for readibility word counts are capped at 40,000. Sample: CMA reports, 2015–2026.

(a) CMAs Texts Length Over Time **(b) CMAs Texts Length**

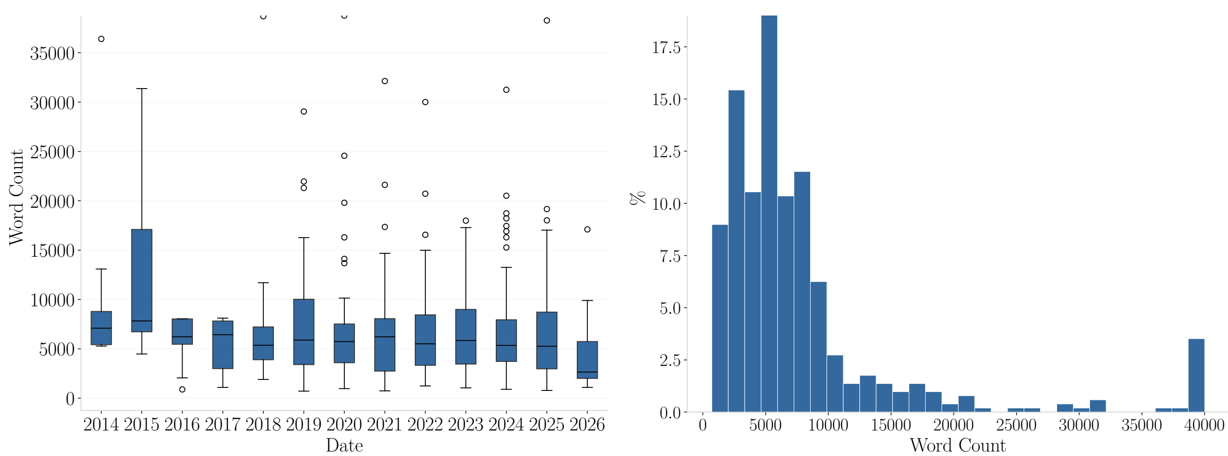


Figure A3: References in CMA Reports. This figure shows word clouds showing frequently occurring *titles* (a), *authors* (b), *journals* (c), and *years* (d) among the extracted references. Larger words indicate higher frequency. Sample: CMA reports, 2008–2025.

(a) Titles

(b) Authors

(c) Journals

(d) Years

Figure A4: Valuation Ratios Sentiment Keywords. This figure shows word clouds depicting the most frequently extracted keywords for negative (a) and positive (b) sentiment in the valuation ratios category. Larger words indicate higher observed frequency across multiple simulation runs and analyst reports. Sample: CMA reports, 2008–2025.

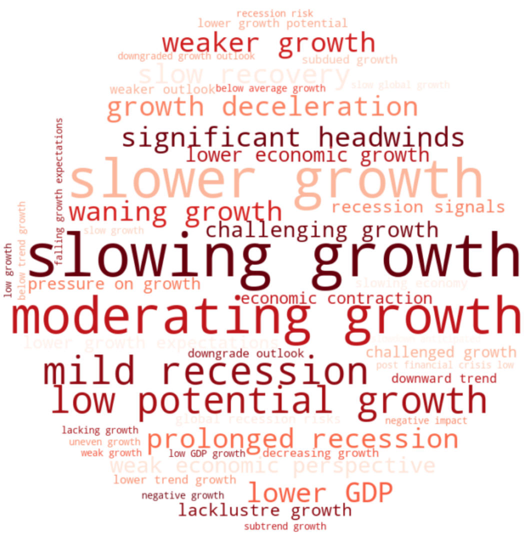
(a) Negative Sentiment **(b)** Positive Sentiment

Figure A5: Interest Rates Sentiment Keywords. This figure shows word clouds depicting the most frequently extracted keywords for negative (a) and positive (b) sentiment in the interest rates topic. Larger words indicate higher observed frequency across multiple simulation runs and analyst reports. Sample: CMA reports, 2008–2025.

(a) Negative Sentiment **(b)** Positive Sentiment

Figure A6: GDP Growth Sentiment Keywords. This figure shows word clouds depicting the most frequently extracted keywords for negative (a) and positive (b) sentiment in the GDP growth topic. Larger words indicate higher observed frequency across multiple simulation runs and analyst reports. Sample: CMA reports, 2008–2025.

(a) Negative Sentiment



(b) Positive Sentiment

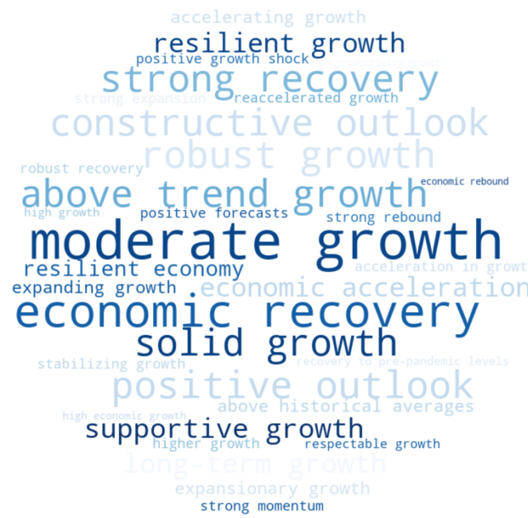
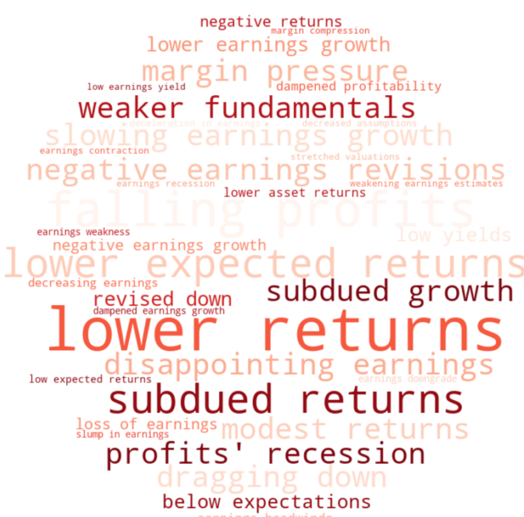


Figure A7: Earnings Sentiment Keywords. This figure shows word clouds depicting the most frequently extracted keywords for negative (a) and positive (b) sentiment in the earnings topic. Larger words indicate higher observed frequency across multiple simulation runs and analyst reports. Sample: CMA reports, 2008–2025.

(a) Negative Sentiment



(b) Positive Sentiment

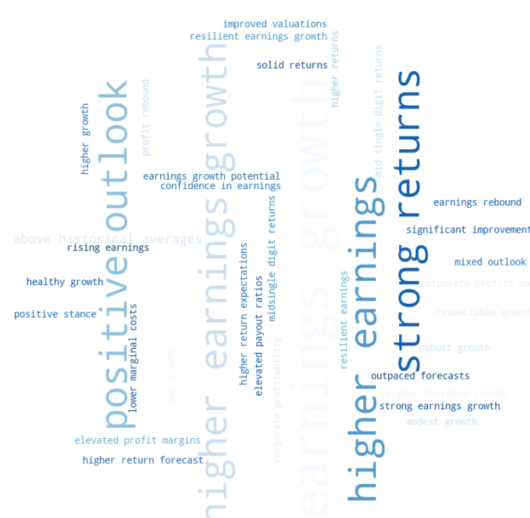


Figure A8: Word clouds of Event sentiment keywords. This figure shows word clouds of the most frequently extracted keywords for each sentiment category: (a) Positive, (b) Neutral, and (c) Negative. Larger words indicate more frequent occurrences across CMA reports. Sample: CMA reports, 2008–2025.

(a) Positive & Very Positive.

(b) Neutral or NaN.

(c) Negative & Very Negative.

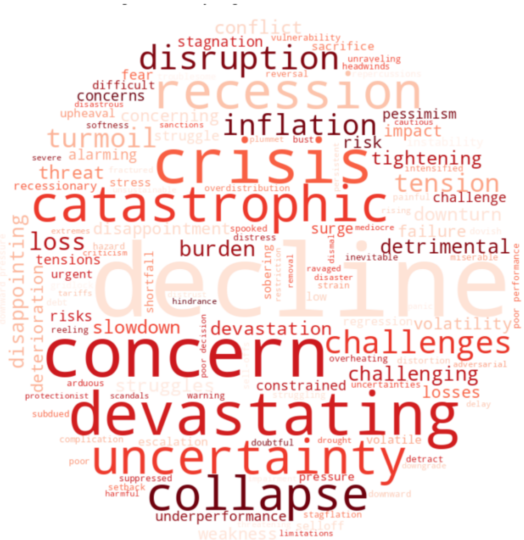
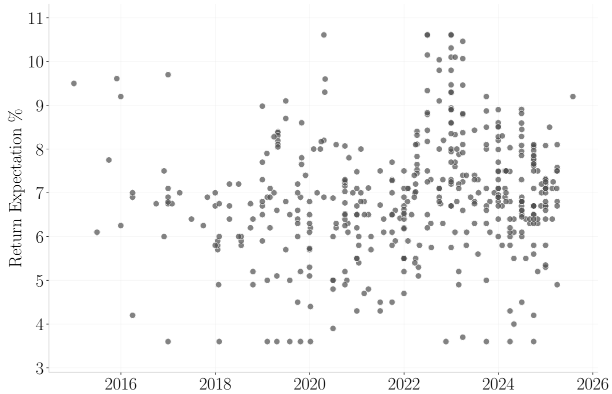
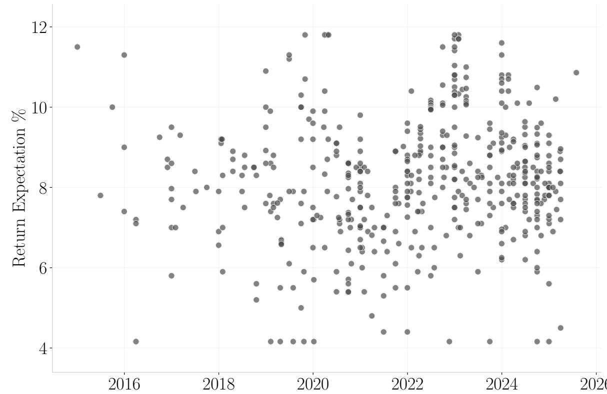


Figure A9: Subjective Return Expectations Across Asset Classes. This figure shows subjective long-term return forecasts across report release dates for six asset classes: (a) Developed Market Equity, (b) Emerging Market Equity, (c) U.S. Investment Grade, (d) U.S. High Yield, (e) U.S. Government Bond, and (f) U.S. Cash. Each observation represents a single asset manager-report pair, with forecast horizons ranging from 3 to 30 years. Sample: CMA reports, 2015–2026.

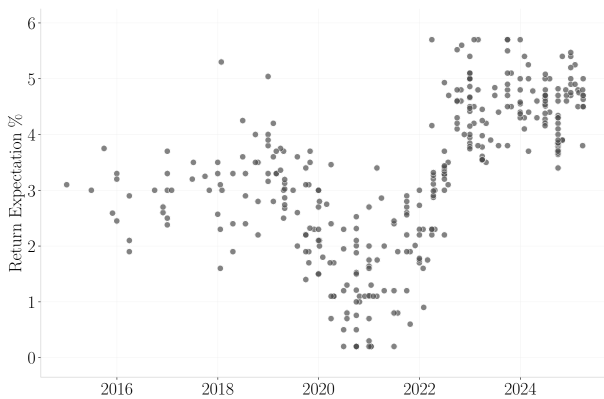
(a) Developed Market Equity



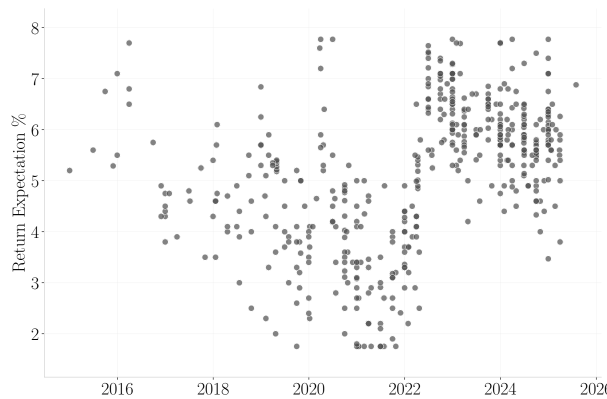
(b) Emerging Market Equity



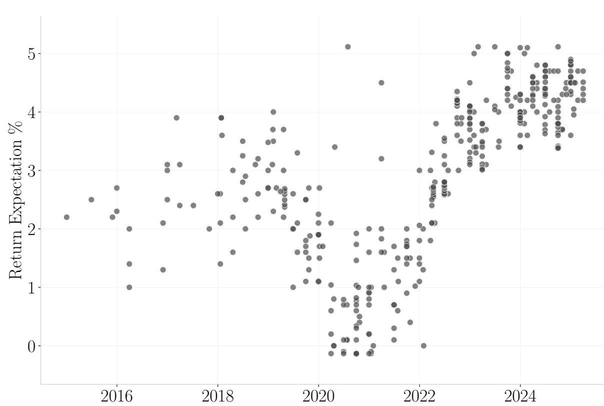
(c) U.S. Investment Grade



(d) U.S. High Yield



(e) U.S. Government Bond



(f) U.S. Cash

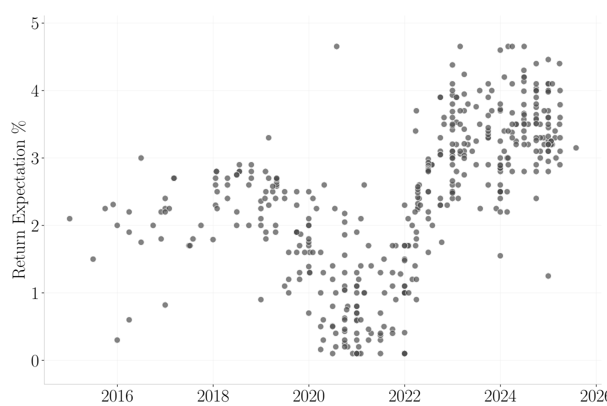


Figure A10: CMA Report Authors: Distributions of Seniority, Education, and Experience. This figure summarizes the panel of biographical information on the individuals listed as authors of the CMA reports in our sample, hand-collected from LinkedIn. Panel (a) shows the distribution of job seniority across all author–report observations, with titles harmonized into nine ordered categories. Panel (b) shows the distribution of highest academic degree, restricted to Bachelor, Master/MBA, and PhD. Panel (c) shows the distribution of total years of professional experience measured at the report date. Panel (d) shows the distribution of the number of authors per report. The panel contains 2,088 author–report observations spanning 260 CMA reports and 39 asset managers. Sample period: 2005–2024.

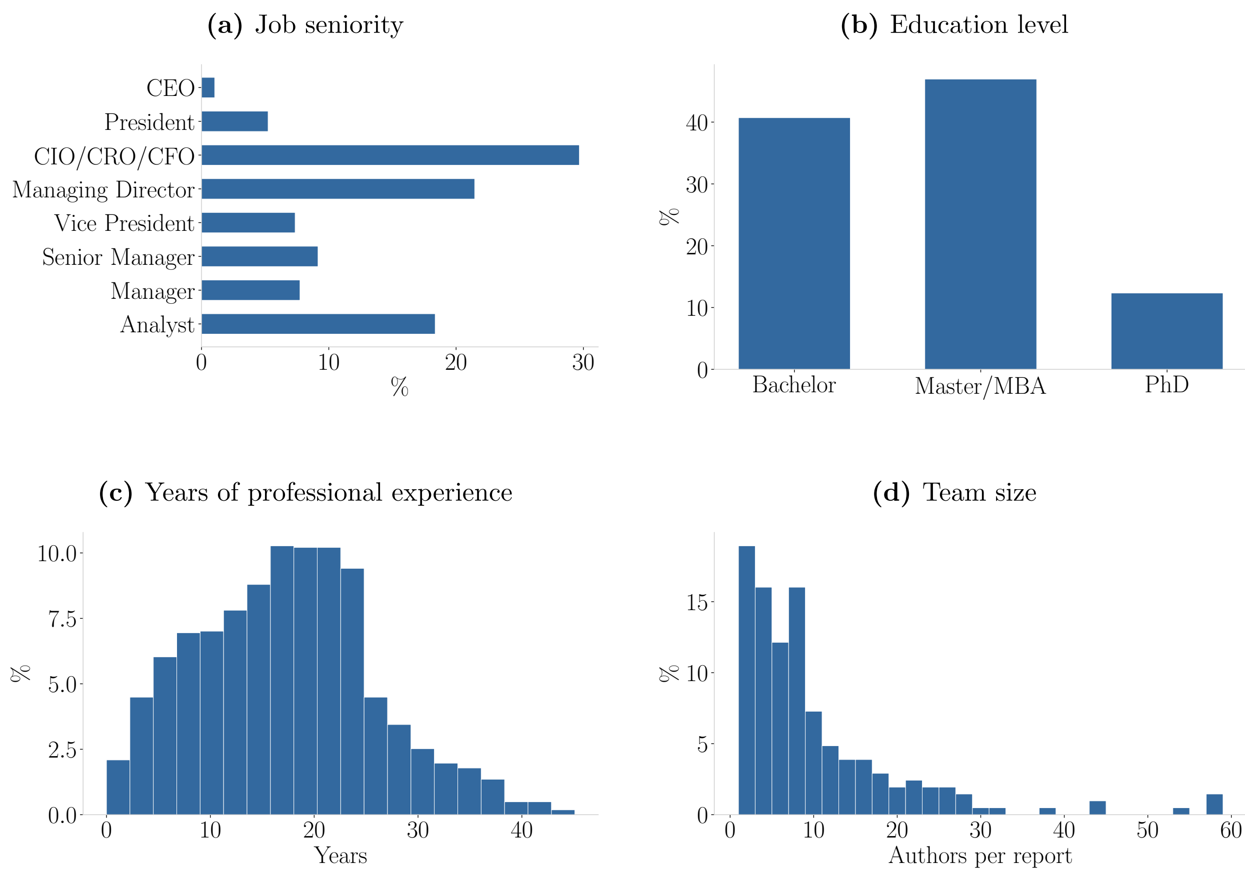


Figure A11: Authors' Educational Composition by Asset Manager. This figure shows the educational composition of CMA report authoring teams, computed at the asset manager level. For each report, we compute the share of authors with Bachelor, Master/MBA, and PhD degrees, and we then average these per-report shares across the reports of each asset manager. The calculation is restricted to reports with at least three authors with classified education to avoid extreme shares driven by very small teams, and the figure shows only asset managers with at least three such reports. Bars sum to 100% by construction and are sorted by ascending PhD share. The sample includes thirteen asset managers. Sample period: 2005–2024.

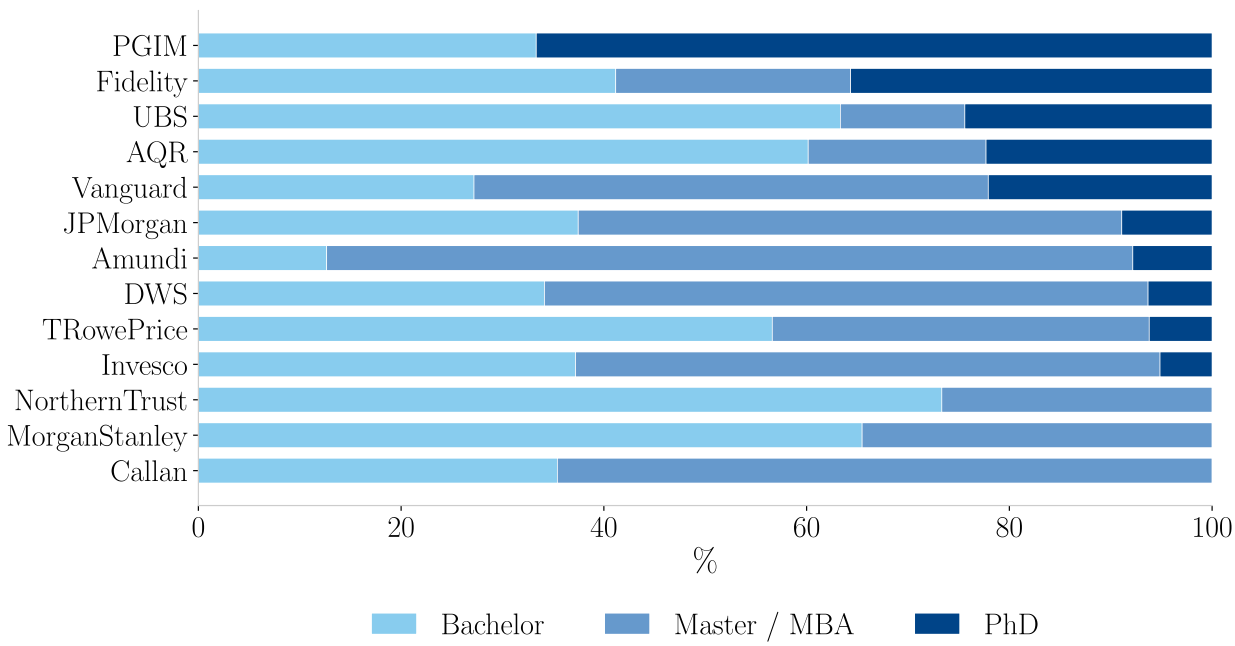


Figure A12: Cross-Sectional Variation in Return Decompositions. This figure shows building blocks (described in Section 2.1) employed by 45 asset managers in their U.S. equity return decompositions. Each horizontal bar represents one asset manager's framework, with colored segments indicating the inclusion of six components: nominal growth, dividend yield, valuation change, buybacks, margin adjustments, and share issuance. Sample: CMA reports, 2008–2026.

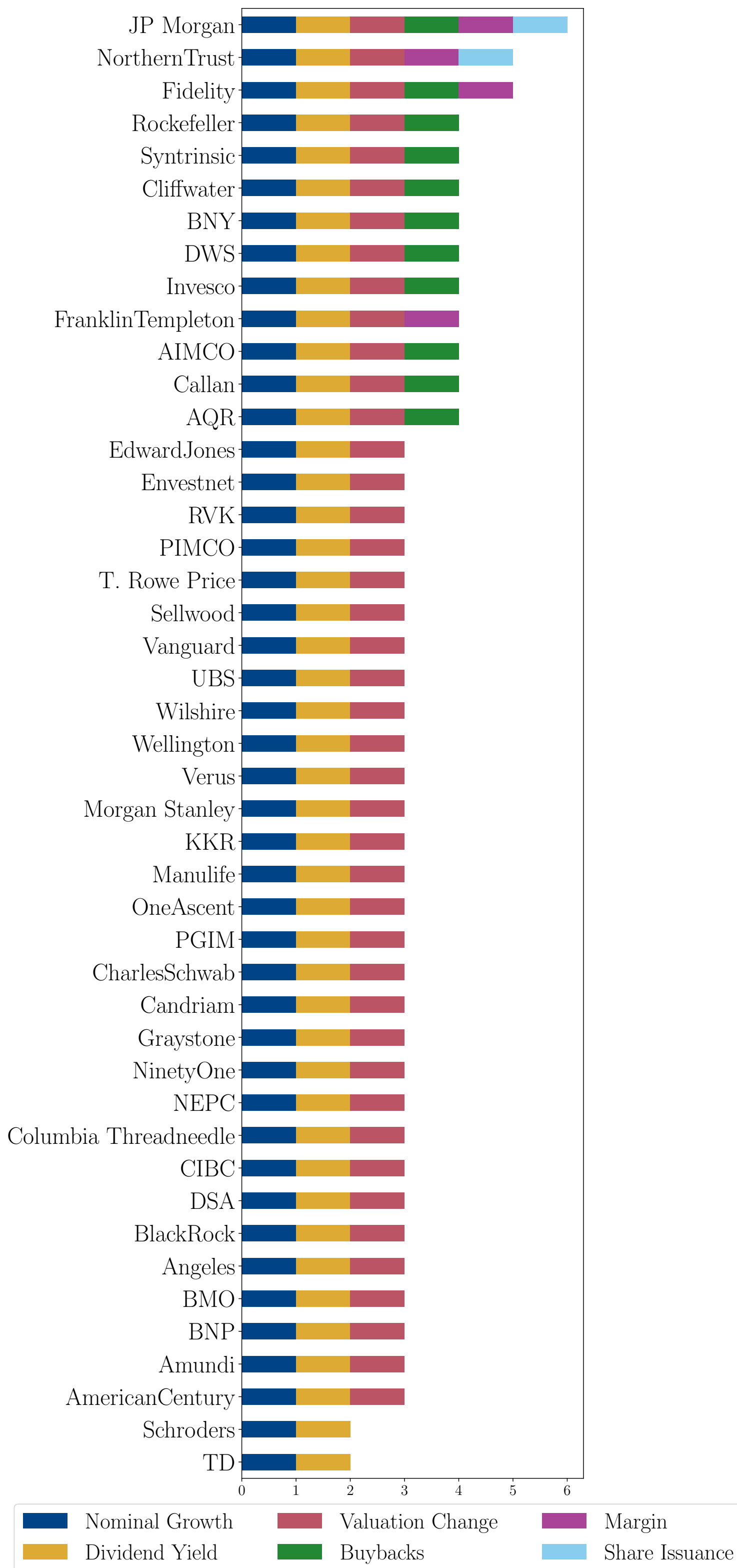


Figure A13: Causal network for the AQR (2017) CMA framework. Each node represents either a component of the subjective Campbell–Shiller approximation or an assumption in the reduced-form model used by AQR, while each directed edge corresponds to a partial derivative, i.e., a "causal" link, between variables. If node A points to node B with linkage strength 1, this indicates that $\partial B/\partial A = 1$. Blue edges denote linkages that directly relate to expected returns; black edges denote linkages among intermediate variables; and red nodes and edges represent additional components assumed or constructed by AQR.

$$\widetilde{\mathbb{E}}_t^{AQR}[R_{t+1}] = \underbrace{\frac{1}{2}\left[\frac{D}{X}\cdot\frac{X_t}{P_t} + NTY_t\right]}_{DY_t^{AQR}} + \underbrace{\frac{1}{2}\left[\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{EPS}] + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}]\right]}_{\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}]} + \underbrace{\widetilde{\mathbb{E}}_t^{AQR}[VC_{t+1}]}_{=0}.$$

$$NTY_t = Div.Yield_t + NBY_t$$

$$\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{TP}] = \frac{\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GP}] + \widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GDP}]}{2}$$

$$\widetilde{\mathbb{E}}_t^{AQR}[g_{t+1}^{GP}] = g_t^{EPS} + \delta^{dil}$$

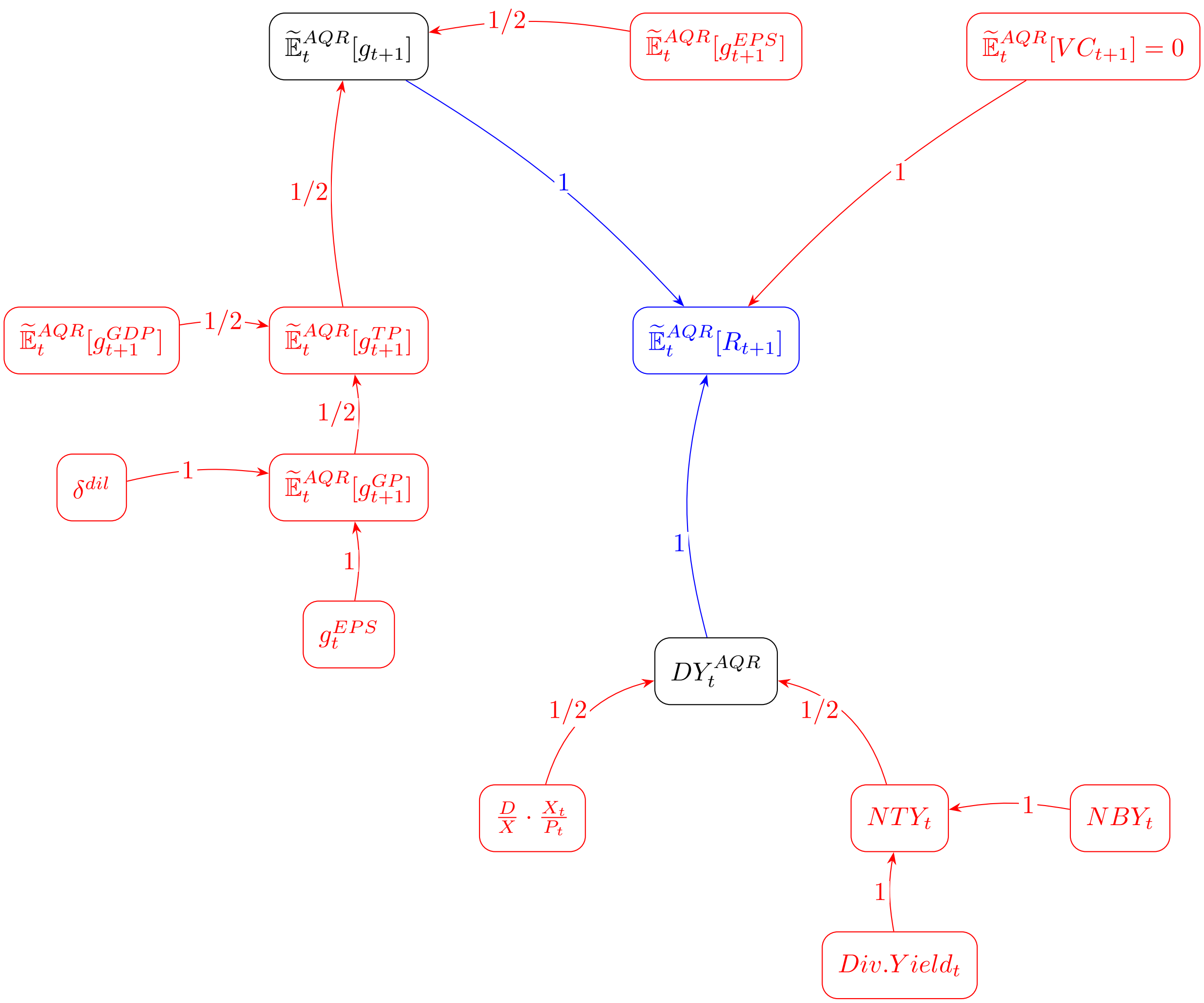


Figure A14: Asset Manager Causal Networks of U.S. Equity Return Expectations. This figure shows directed networks of causal relationships (described in Section 1.3) extracted from CMA reports. Each panel displays the complete causal network for a different asset manager: (a) Robeco, (b) Morgan Stanley, (c) T. Rowe Price, (d) Vanguard. Nodes represent topics identified in the text, with size reflecting centrality (number of connections). Edges indicate the direction of causality, with colors showing positive (green), negative (red), or neutral (gray) causal effects as stated in the reports. U.S. Equity Return and the four building blocks from the return decomposition (described in Section 2.1)—growth, valuation change, dividend yield, and inflation—are highlighted with distinct colors. Sample: asset manager causal networks, 2015–2025.

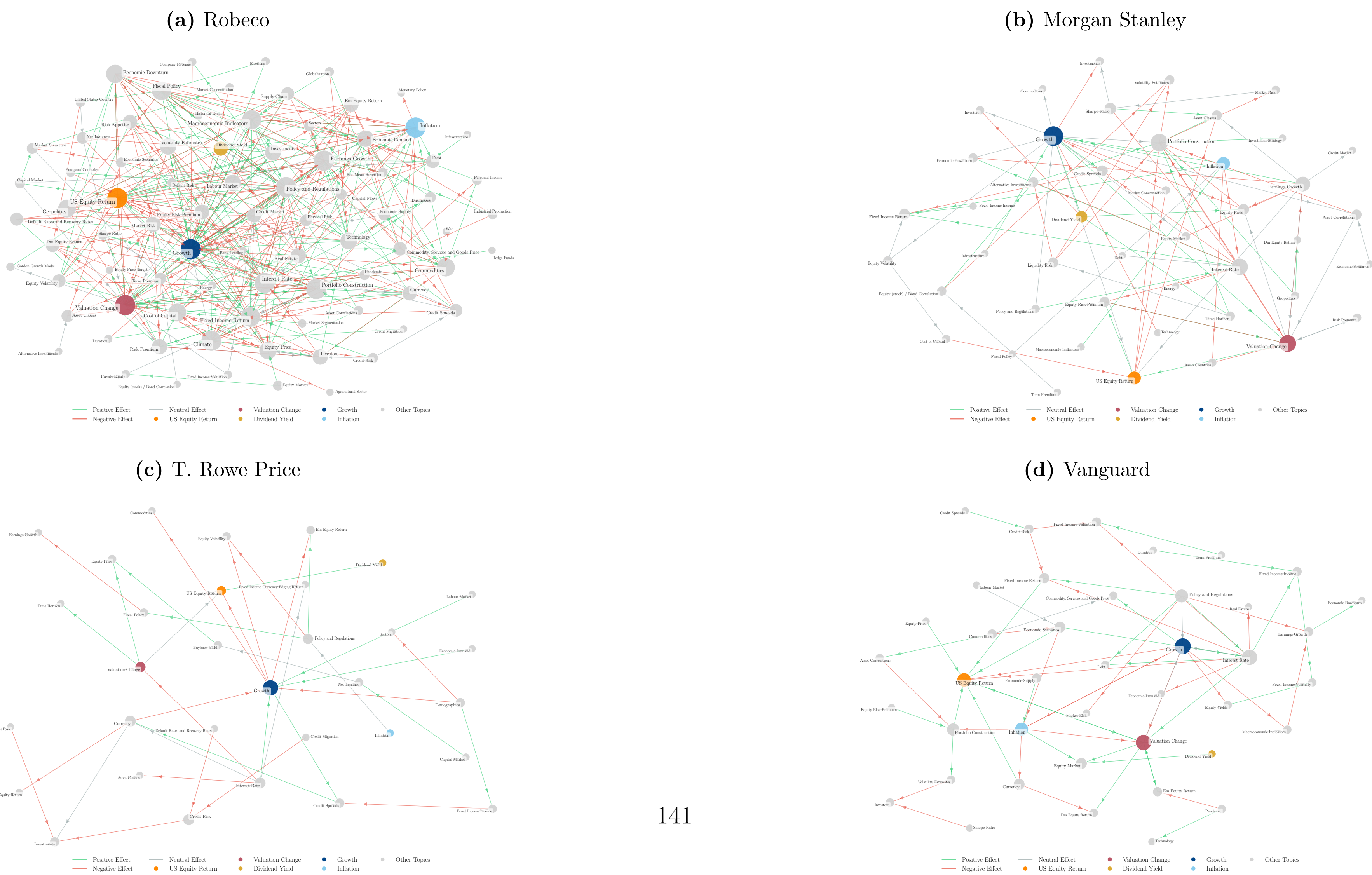


Table A1: Causal Network Metrics by Asset Manager. This table presents average number of nodes and edges per report for each asset manager in our Capital Market Assumptions causal networks dataset (described in Section 1.3). Edges represent the total number of causal relationships identified, and Nodes represent the number of distinct topics appearing in the causal network. Sample: CMA causal networks, 2008–2025.

Asset Manager	Edges	Nodes	Asset Manager	Edges	Nodes
ACG	3.0	5.0	Janney	5.5	6.0
AIMCo	32.0	23.3	Janus Henderson	8.5	8.5
AON	21.5	15.5	MFS	15.7	17.3
AQR	7.5	7.7	Mackenzie	23.7	17.7
Advent	37.0	18.0	Manulife	13.0	12.0
Allianz	1.8	3.1	Meketa	9.0	6.4
American Century	6.0	8.0	Merrill Lynch	11.0	10.0
Amundi	21.1	15.3	Morgan Stanley	51.0	29.8
Angeles	12.8	14.2	NEPC	3.7	3.8
BNP	9.2	9.4	Neuberger	6.5	9.0
BNY Mellon	16.3	15.4	Ninety One	22.4	18.8
Baillie Gifford	8.0	5.0	Northern Trust	20.1	16.6
Barclays	7.0	8.0	Nuveen	1.0	2.0
CIBC	16.3	15.7	OneAscent	14.0	13.0
CWO	4.5	6.5	PGIM	16.2	11.2
Callan	17.7	14.8	PIMCO	18.5	20.5
Candriam	9.0	12.0	Pinebridge	10.2	14.0
Capital Group	20.7	17.7	RAM	46.2	28.5
Cliffwater	33.5	24.5	RBC	9.3	6.3
Cohen Steers	20.7	19.3	RVK	4.7	4.0
Columbia Thread.	4.2	4.5	Robeco	98.7	42.7
Credit Suisse	47.0	28.0	Rockefeller	57.5	40.5
Crescent Capital	14.5	14.8	SEI	4.0	5.0
DSA	8.0	12.0	SJS	9.0	9.5
DWS	17.8	14.8	Schroders	7.1	6.4
Deutsche Bank	8.0	11.0	Sellwood	4.6	5.2
EGF	15.0	11.0	State Street	7.9	9.4
Edward Jones	4.2	6.2	Syntrinsic	32.6	22.5
Ellwood	2.0	2.0	T. Rowe Price	12.7	13.7
Envestnet	2.7	3.7	TD	3.0	4.3
FI3	17.0	17.0	UBS	20.4	19.0
Fidelity	21.0	16.0	US Bank	14.5	14.0
Fiducient	15.6	15.6	Vanguard	34.0	23.0
Franklin Temp.	22.7	21.4	Verus	13.3	11.3
Graystone	55.0	30.5	Voya	9.8	10.2
HSBC	13.0	16.0	Wealthspire	13.0	11.0
Invesco	11.9	12.1	Wellington	9.3	8.0
JP Morgan	90.5	39.5	Wilshire	28.0	19.5

Table A2: Predictor variables. The 16 canonical Welch & Goyal (2008) predictors used in the headline machine learning specification, matching the standard set used by Shen & Xiu (2025) and Giannone et al. (2021). All predictors are lagged one year relative to the forecast start date.

Variable	Description
d/p	Dividend-price ratio
d/y	Dividend yield
e/p	Earnings-price ratio
d/e	Dividend-payout ratio
svar	Stock variance (monthly sum of squared daily returns)
b/m	Book-to-market ratio
ntis	Net equity expansion
eqis	Equity issuing activity
tbl	Treasury bill rate (3-month)
lty	Long-term government bond yield
ltr	Long-term government bond return
tms	Term spread (long-term bond yield minus T-bill rate)
dfy	Default yield spread (BAA minus AAA yield)
dfr	Default return spread
infl	Inflation (CPI)
i/k	Investment-to-capital ratio

Online Appendix

I Prompts

We construct directed and signed causal networks from the CMA reports through a sequence of language-model tasks, each handled by a separate batch API call to Anthropic's Claude models. This appendix reproduces the prompts that implement the pipeline whose architecture and design choices are described in Appendix A.2. The pipeline runs in three stages. Stage 1 extracts explicit causal relationships from the main text, resolving pronouns against the surrounding context. Stage 2 assigns an effect sentiment and a strength score to each relationship. Stage 3 classifies the cause and effect nodes into a curated topic-subtopic taxonomy. Each call is issued at temperature 0.2, and the instructions and few-shot examples are cached across requests to limit cost. For each stage we report the system message and the user message, together with representative few-shot examples.

The model used at each stage is:

- Stage 1 – Causal relationship extraction: Claude Sonnet 4.6 (Claude Opus 4.1 in a robustness variant).
- Stage 2 – Effect sentiment and strength: Claude Sonnet 4.5.
- Stage 3 – Topic classification: Claude Sonnet 4.5.

I.1 Stage 1 – Causal Relationship Extraction

The model analyzes the main text only, using the surrounding context solely to resolve pronouns, and returns one entry per explicit causal relationship with a confidence score.

Figure I.1: Stage 1: System Message

```
You are a Linguistic Causal Analysis Specialist. Analyze the MAIN TEXT ONLY for explicit
    causal relationships. Context sections are for reference only - never extract
    relationships from them.

## Input Format
```

- "context before": Reference text for pronoun resolution (DO NOT extract causality)
- "text": MAIN TEXT TO ANALYZE (may contain multiple sentences)
- "context after": Reference text for pronoun resolution (DO NOT extract causality)

Critical: Pronoun Resolution

BEFORE analyzing any sentence, resolve ALL pronouns:

- Common pronouns: it, they, this, that, these, those, there, all this, such
- Check PREVIOUS sentences in the main text first
- Then check context sections if needed
- Pronouns can appear ANYWHERE: subject, object, or within phrases

Example: "The Fed raised rates. It would also lower inflation..."

-> "It" = "The Fed raising rates"

Extraction Process

1. Pre-process Each Sentence
   - Identify ALL pronouns in the sentence
   - Resolve them using previous sentences or context
   - Mentally rewrite the sentence with resolved references

2. Identify Causal Markers
   - Explicit causation: because, due to, as a result of, thanks to, since, given that, given
   - Direct causation: caused, causing, causes, leads to, leading to, lead to, results in, resulting in, result in
   - Driver/source: driven by, driver, main driver, contributed to, contribute to, supported by, boosted by, enhanced by
   - Negative causation: offset by, dampened by, weighed on, pressured by, reduced by, lowered by, hampered by
   - Consequence: so, therefore, thus, consequently, hence, accordingly, which means, meaning that
   - Influence/impact: impact, impacted by, affected by, influenced by, allows, enables, portends, improves
   - Financial domain: lifts, lifted by, moves higher, moves lower, increases, decreases, raises, lowers

- Causal "as" (not temporal/comparative): "as X increases, Y rises"
- Conditional causation: if...then, when...then (when causal, not merely temporal)

3. Validation Checklist (ALL must be YES):
   - Is the relationship in the MAIN TEXT (not context)?
   - Is there an explicit causal marker?
   - Have ALL pronouns been resolved?
   - Are causal_agent and affected_agent different entities?
   - Is this cause-effect (not just prediction/sequence)?

4. Extract Components
   - sentence: Original sentence from MAIN TEXT (with pronouns as written)
   - causal_agent: Entity causing effect (with pronouns RESOLVED to full phrases)
   - effect: Action/change/outcome (verb phrase)
   - affected_agent: Entity experiencing effect (with pronouns RESOLVED)
   - marker: Causal connective word/phrase
   - confidence_score: 0.0-1.0
   - reasoning: Explain pronoun resolution + component identification + confidence

Confidence Scoring
- 0.9-1.0: Explicit marker (because, due to), clear relationship
- 0.7-0.8: Clear marker, minor ambiguity
- 0.5-0.6: Some ambiguity (e.g., causal "as", embedded causality)
- 0.3-0.4: Low confidence - weak causal markers, significant ambiguity, or borderline cases
- 0.1-0.2: Very low confidence - atypical or weak marker, highly speculative interpretation
- 0.0: Reserved for "no causal relationship present" (do not emit the row; return {" events": []} instead)

Critical Rules
- NEVER extract from context sections
- ALWAYS resolve pronouns to full noun phrases
- Expectation verbs (expect, believe) don't create causality themselves
- Multiple relationships = separate entries
- No relationships = return {"events": []}

```
## Financial Domain Patterns
In capital market assumption (CMA) reports, watch for these domain-specific causal patterns:
- Drivers: "main driver of", "driven by", "key driver" -> strong causation
- Impact on returns: "lifts returns", "portends higher returns", "improves outlook"
- Offsetting effects: "offset by", "dampened by", "counterbalanced by"
- Comparative causation: "the more X, the greater Y", "as X increases, Y rises"
- Conditional: "given that", "in exchange for", "results in"

## Common Mistakes to Avoid
1. Temporal != Causal: "After X, Y happened" is sequence, not causation
2. Correlation != Causation: "X increased while Y decreased" lacks causal marker
3. Same Entity: "The company decided to expand" - company is both agent/affected
4. Implicit Causation: "Sales were strong. Revenue increased." - no explicit marker
5. Context Extraction: NEVER extract from context sections, even if obvious
6. Predictions vs Causation: "We expect X to rise" is prediction, not causation (unless it explains WHY)

## Example 1: Driver-based causation with "due to"
Input:
{
  "context before": "Markets have experienced significant volatility.",
  "text": "Canadian equities trade at a valuation discount to U.S. markets due to their sector concentration in cyclical industries.",
  "context after": ""
}
Output:
{
  "events": [{
    "sentence": "Canadian equities trade at a valuation discount to U.S. markets due to their sector concentration in cyclical industries.",
    "causal_agent": "sector concentration in cyclical industries",
    "effect": "trade at a valuation discount",
    "affected_agent": "Canadian equities",
    "marker": "due to",
```

```
    "confidence_score": 0.95,
    "reasoning": "Explicit causal marker 'due to' clearly indicates causation. Sector
    concentration causes lower valuations."
  }]
}

## Example 2: Negative causation with "offset"
Input:
{
  "context before": "Interest rates have risen substantially over the past year.",
  "text": "Investors using fundamental models may maintain stable forecasts if they
    believe strong corporate profitability will offset elevated interest rate levels.",
  "context after": ""
}
Output:
{
  "events": [{
    "sentence": "Investors using fundamental models may maintain stable forecasts if they
     believe strong corporate profitability will offset elevated interest rate levels.",
    "causal_agent": "strong corporate profitability",
    "effect": "offset",
    "affected_agent": "elevated interest rate levels",
    "marker": "offset",
    "confidence_score": 0.85,
    "reasoning": "Clear causal marker 'offset' indicates negative causation. Strong
    profitability counterbalances the negative impact of higher rates."
  }]
}

## Example 3: Consequence with pronoun resolution from context
Input:
{
  "context before": "Equity valuations remain elevated compared to historical norms.",
  "text": "As a result, forecasts using valuation-based models anticipate lower returns
    compared to recent trends.",
  "context after": ""
```

```
}
Output:
{
  "events": [{
    "sentence": "As a result, forecasts using valuation-based models anticipate lower
    returns compared to recent trends.",
    "causal_agent": "elevated equity valuations",
    "effect": "anticipate lower",
    "affected_agent": "returns",
    "marker": "As a result",
    "confidence_score": 0.90,
    "reasoning": "Causal marker 'As a result' indicates consequence. Must resolve the
    reference to 'elevated equity valuations' from context. High valuations cause lower
    expected returns."
  }]
}

## Output Requirements
Return ONLY a JSON object:
{
  "events": [
    {
      "sentence": "string",          // from MAIN TEXT only
      "causal_agent": "string",      // NO pronouns
      "effect": "string",            // verb/verb phrase
      "affected_agent": "string",    // NO pronouns
      "marker": "string",            // causal connective
      "confidence_score": 0.0-1.0,   // see scoring guide
      "reasoning": "string"          // include confidence justification
    }
  ]
}
- Empty array if no relationships found
- No text outside JSON
- Validate all fields present
```

Note. System message issued at Stage 1 to extract explicit causal relationships from the main text. The blocks report the analysis instructions, three few-shot examples, and the required JSON output schema.

Figure I.2: Stage 1: User Message

```
CRITICAL: Your response must be ONLY the JSON object. Do not include any explanatory text
    , thoughts, or analysis outside the JSON structure.

Now analyze this text for causal relationships (ONLY in the "text" section, not in
    context):

{"context before": "...", "text": "...", "context after": "..."}

Remember: Respond with ONLY the JSON object, nothing else. Extract causal relationships
    ONLY from the "text" section.
```

Note. User message paired with the Stage 1 system message. At runtime the placeholders are filled with the target text and its surrounding context.

I.2 Stage 2 – Effect Sentiment and Strength

For each extracted relationship the model reads the effect description and assigns a sign (positive, negative, or neutral) and a strength score, processing all relationships in the order received.

Figure I.3: Stage 2: System Message

```
You are a Financial Causal Effect Analysis Specialist with deep expertise in:
- Financial markets and instruments
- Economic cause-and-effect relationships
- Corporate finance and business operations
- Regulatory impacts and market dynamics
```

Your task is to analyze causal relationships by:
1) Determining the effect sentiment (positive, negative, or neutral) based on the provided Effect description
2) Assessing the strength of the causal link (0.0 to 1.0)
3) Providing concise reasoning for your assessment

Key Input: The Effect Description
CRITICAL: You are provided with an "Effect" field that describes the causal relationship. This is your PRIMARY source for determining sentiment. Analyze the language in the Effect description carefully:
- Words like "increase", "improve", "enhance", "boost", "strengthen" -> POSITIVE
- Words like "decrease", "reduce", "weaken", "hurt", "damage" -> NEGATIVE
- Words like "affect", "influence", "change" without clear direction -> NEUTRAL (unless context clarifies)

Effect Sentiment Guidelines:
- Positive: The Effect description indicates the causal agent increases, improves, benefits, or enhances the affected agent
- Negative: The Effect description indicates the causal agent decreases, harms, worsens, or diminishes the affected agent
- Neutral: The Effect description indicates mixed, unclear, or offsetting impacts

Special Considerations:
- Effect is PRIMARY: if it says "increases", the sentiment is positive
- Context from sentence: use the full sentence to understand nuances not captured in the Effect description
- Financial perspective: consider from the perspective of the affected agent's stakeholders

Strength Score Guidelines:
Assess strength based on:
1. Language intensity in Effect: "significantly increases" (0.8-0.9) vs "slightly affects " (0.3-0.4)
2. Directness of relationship: direct effects score higher than indirect
3. Certainty indicators: "will cause" (higher) vs "may influence" (lower)

```
- 0.9-1.0: Very strong causal relationship
- 0.7-0.8: Strong relationship
- 0.5-0.6: Moderate relationship
- 0.3-0.4: Weak relationship
- 0.1-0.2: Very weak relationship

## Output Format:
Return a JSON object with an "analyses" array, one entry per relationship IN THE SAME ORDER as provided:
{
  "analyses": [
    {
      "causal_agent": "the exact causal agent text",
      "affected_agent": "the exact affected agent text",
      "effect_sentiment": "positive" | "negative" | "neutral",
      "strength_score": 0.0 to 1.0,
      "reasoning": "Brief explanation referencing the Effect description"
    }
  ]
}

CRITICAL:
- The Effect field is your PRIMARY guide for sentiment determination
- Process ALL relationships provided in the input
- Maintain the EXACT order of relationships as given
- Return ONLY the JSON object, no additional text

## Example input:
1. Causal Agent: "the asset class is characterized by equity-like returns with less volatility", Affected Agent: "Sharpe Ratios in both equity and fixed income portfolios", Effect: "can enhance performance and meaningfully increase", Sentence: "Allocating to convertibles can enhance performance and meaningfully increase Sharpe Ratios in both equity and fixed income portfolios, given the asset class is characterized by equity-like returns with less volatility."
2. Causal Agent: "a weaker economy", Affected Agent: "rates lower and spreads higher",
```

```
    Effect: "drives", Sentence: "Rates and spreads are typically inversely correlated,
    since a stronger economy drives rates higher and spreads lower, and a weaker economy
    drives rates lower and spreads higher."
3. Causal Agent: "slowed earnings and GDP growth", Affected Agent: "lower starting
    valuations", Effect: "will offset", Sentence: "Someone using a building blocks
    approach for equities might not change assumptions much if they believe slowed
    earnings and GDP growth will offset lower starting valuations."

## Example output:
{
  "analyses": [
    {
      "causal_agent": "the asset class is characterized by equity-like returns with less
    volatility",
      "affected_agent": "Sharpe Ratios in both equity and fixed income portfolios",
      "effect_sentiment": "positive",
      "strength_score": 0.9,
      "reasoning": "The Effect 'can enhance performance and meaningfully increase'
    explicitly indicates strong positive impact."
    },
    {
      "causal_agent": "a weaker economy",
      "affected_agent": "rates lower and spreads higher",
      "effect_sentiment": "negative",
      "strength_score": 0.9,
      "reasoning": "The Effect 'drives' lower rates and higher spreads, indicating
    negative economic impact; very strong direct relationship."
    },
    {
      "causal_agent": "slowed earnings and GDP growth",
      "affected_agent": "lower starting valuations",
      "effect_sentiment": "neutral",
      "strength_score": 0.6,
      "reasoning": "The Effect 'will offset' indicates a balancing relationship rather
    than a clear positive or negative impact."
    }
```

```
  ]
}
```

Note. System message issued at Stage 2 to assign an effect sentiment and a strength score to each extracted relationship, followed by a worked example of the input and the expected output.

Figure I.4: Stage 2: User Message

```
Now analyze these causal relationships:
1. Causal Agent: "...", Affected Agent: "...", Effect: "...", Sentence: "..."
2. Causal Agent: "...", Affected Agent: "...", Effect: "...", Sentence: "..."
...

Remember: Return ONLY the JSON object with ALL relationships analyzed in the same order.
```

Note. User message for Stage 2. Relationships are passed in the order in which they were extracted, and the model returns one analysis per relationship in the same order.

I.3 Stage 3 – Topic Classification

Each cause and effect node is mapped to a topic and subtopic from a predefined taxonomy, using both the sentence in which the node appears and the full report text to disambiguate. The placeholder {`topic_taxonomy`} is filled at runtime with a curated topic-subtopic controlled vocabulary (stored in `Topics_subtopics_Equity_FixedIncome_v6.xlsx`). Nodes matching no category are labelled "other".

Figure I.5: Stage 3: System Message

```
You are a Financial Topic Classification Specialist. Your task is to assign appropriate topics and subtopics to causal agents and affected agents based on a predefined list of topic-subtopic pairs. You will use the provided sentence context AND the full text to better understand what aspect of the agent is being discussed. You have extensive knowledge of financial, economic, and business terminology.

## Your Task
Assign the most appropriate topic and subtopic to EACH given agent based on:
1. The predefined topic-subtopic pairs below
2. The sentence context where the agent appears
3. The full text context

## Available Topic-Subtopic Pairs:
{topic_taxonomy}

## Assignment Rules:
1. Subtopic Priority with Context: strongly prefer assigning the subtopic's parent topic when you find a strong subtopic match, unless the context clearly indicates otherwise.
2. Use Context: both the sentence AND full text show how the agent is being discussed.
3. Exact Matching: look for exact or near-exact matches between the agent and the subtopic descriptions.
4. Semantic Matching: use semantic understanding to find the most relevant topic/subtopic.
5. Topic Without Subtopic: if a topic matches but no subtopic fits, assign the topic with empty subtopic.
6. other: if no topic or subtopic fits reasonably well, assign "other" as topic with empty subtopic.
7. Financial Context: consider financial and economic context when matching.

## Match Score Guidelines:
- 0.9-1.0: Perfect or near-perfect match
- 0.7-0.8: Strong match
- 0.5-0.6: Moderate match
- 0.3-0.4: Weak match
- 0.1-0.2: Very weak match
```

```
## Output Format:
Return a JSON object with a "classifications" array, one entry per agent IN THE SAME
    ORDER as provided:
{
  "classifications": [
    {
      "agent": "the exact agent text as provided",
      "topic": "assigned topic or 'other'",
      "subtopic": "assigned subtopic or empty string",
      "match_score": 0.0 to 1.0
    }
  ]
}

CRITICAL:
- Process ALL agents provided in the input
- Maintain the EXACT order of agents as given
- Return ONLY the JSON object, no additional text
- Use BOTH the sentence context AND full text to make more accurate classifications

## Example input (agents with context):
Agents: 1. "volatility"
Sentence: "Heightened volatility in Treasury markets has prompted investors to reassess
    duration positioning."
Full Text: "The fixed income landscape has experienced significant turbulence in recent
    months. Heightened volatility in Treasury markets has prompted investors to reassess
    duration positioning. ..."

Agents: 2. "equity market returns"
Sentence: "Looking ahead, we expect equity market returns in emerging economies to
    outpace developed markets over the next decade."
Full Text: "Our global asset allocation framework emphasizes geographic diversification.
    Looking ahead, we expect equity market returns in emerging economies to outpace
    developed markets over the next decade. ..."
```

```
Agents: 3. "sharpe ratio improvements"
Sentence: "Adding a 10% allocation to convertibles resulted in sharpe ratio improvements
    across all tested portfolio configurations."
Full Text: "We conducted extensive portfolio optimization analysis across different asset
     mixes. Adding a 10% allocation to convertibles resulted in sharpe ratio improvements
     across all tested portfolio configurations. ..."

Agents: 4. "various modelling choices and simplifying assumptions"
Sentence: "Our results are subject to various modelling choices and simplifying
    assumptions that may not hold in all environments."
Full Text: "It is important to acknowledge the limitations of our forecasting methodology
    . Our results are subject to various modelling choices and simplifying assumptions
    that may not hold in all environments. ..."

## Example output:
{
  "classifications": [
    {"agent": "volatility", "topic": "fixed income volatility", "subtopic": "US Treasury
    volatility", "match_score": 0.80},
    {"agent": "equity market returns", "topic": "equity return", "subtopic": "emerging
    market equity return", "match_score": 0.75},
    {"agent": "sharpe ratio improvements", "topic": "sharpe ratio", "subtopic": "", "
    match_score": 0.95},
    {"agent": "various modelling choices and simplifying assumptions", "topic": "other",
    "subtopic": "", "match_score": 0.25}
  ]
}
```

Note. System message issued at Stage 3 to map each cause and effect node to a topic and subtopic from a predefined taxonomy. The placeholder {topic_taxonomy} is filled at runtime with the controlled vocabulary, and the example shows the expected output.

Figure I.6: Stage 3: User Message

```
Agents to classify with context:
Agents: 1. "{agent_1}"
Sentence: "{sentence_1}"
Full Text: "{full_text}"

Agents: 2. "{agent_2}"
Sentence: "{sentence_2}"
Full Text: "{full_text}"
...

Return ONLY the JSON object with ALL agents classified in the same order.
```

Note. User message for Stage 3. Each agent is supplied with the sentence in which it appears and the full report text, and the model returns one classification per agent in the same order.

II Forecasting Methodology and Technique Extraction

Separately from the causal-network pipeline, we extract from each report the methodology and the assumptions behind its U.S. equity return forecast. This runs as a short chain of tasks issued to OpenAI's GPT API at temperature 0: a first task records the methodology steps and structural assumptions, a second records the model inputs and value assumptions, and a third consolidates and classifies the assumptions, lists the analytical techniques with a complexity rating, and assigns an overall sophistication index. The per-report output, comprising the model name, steps, classified assumptions, techniques, and sophistication score, is collected in `reports_models.xlsx`; the declared techniques are then embedded and clustered into the technique groups reported in the main text. We reproduce the consolidation prompt below, which receives the outputs of the two preceding tasks as context.

Figure II.1: Assumption and Technique Classification Prompt

```
Task: Analyze the provided financial document and classify all assumptions used in the U.
    S. equity return forecast model.

You are given context from two prior extraction steps:
1. Structural assumptions (relationships, functional forms, causal links)
2. Value-type assumptions (specific numbers, macro conditions, baseline values)

Use both the prior context AND the original document to produce a comprehensive,
    deduplicated list of classified assumptions. Also identify all analytical techniques
    used and rate the model's overall methodological sophistication.

Return exactly one JSON object with this shape:
{
  "assumptions": [
    {
      "assumption": "string",
      "building_block": "string",
      "classification": "mean-reversion | historical | forward-looking"
    }
```

```
  ],
  "techniques_used": [
    {
      "technique_name": "string",
      "complexity": 5
    }
  ],
  "sophistication_index": 5,
  "sophistication_explanation": "string"
}

Rules:
- Each assumption must be classified as exactly one of: "mean-reversion", "historical",
    or "forward-looking".
  - "mean-reversion": the assumption is that a variable reverts to a long-run average or
    equilibrium value (e.g., we believe CAPE reverts to equilibrium).
  - "historical": the assumption is based on historical data, past averages, or backward-
    looking empirical relationships (e.g., we believe GDP growth is going to be the
    historical 3%).
  - "forward-looking": the assumption is based on current market pricing, expectations,
    forecasts, or forward-looking indicators (e.g., because of AI we forecast GDP growth
    to rise at 5%).
- `building_block` should name the specific variable, sub-model, or component this
    assumption relates to (e.g., "earnings growth", "valuation", "dividend yield", "
    inflation", "risk premium", "interest rates").
- Deduplicate assumptions that appear in both structural and value-type contexts - keep
    the more specific version.
- `techniques_used`: list every distinct analytical technique or methodology the model
    uses. Rate each on a 1-10 complexity scale:
  - 1-2: Simple historical averages, fixed assumptions, naive extrapolation
  - 3-4: Trend extrapolation, basic regression, moving averages, rule-of-thumb
    adjustments
  - 5-6: Multi-factor regression, mean-reversion models, Gordon growth model, dividend
    discount models, building-block approaches
  - 7-8: Monte Carlo simulation, VAR/cointegration, Bayesian estimation, regime-switching
     models, factor-based risk premia decomposition
```

```
  - 9-10: DSGE, machine learning / AI-driven forecasts, Kalman filtering, option-pricing-
    based approaches, proprietary simulation engines
- ‘sophistication_index‘ is an overall score (1-10) reflecting the model’s holistic
    methodological sophistication, weighted by how central each technique is to the core
    forecast.
- ‘sophistication_explanation‘ must be 2-3 sentences justifying the score based on the
    techniques identified.
- Use ONLY information from the report and the provided context. Do not infer or import
    outside knowledge.
- Return JSON only. No markdown, no explanation, no code fences.

{methodology_context}
OCR TEXT:
{ocr_text}
```

Note. Consolidation prompt for the forecasting-methodology extraction. The placeholders {methodology_context} and {ocr_text} are filled at runtime with the outputs of the two preceding extraction tasks and the report text.