From Numbers to Words: Breaking Down Institutional Beliefs*

Andrea Andolfatto† *Federico Bastianello*‡

[Link to the latest version]

July 17, 2026

Abstract

Using Capital Market Assumptions (CMA) reports, in which managers publish long-horizon forecasts and explain how they construct them, we show how they form beliefs about returns, volatilities, and correlations. They decompose return expectations into broad building blocks across asset classes, but assign different values to them. For U.S. equity, valuation change is the largest net source of disagreement. Growth expectations, however, partly offset valuation-change expectations because the two are strongly negatively correlated, so return expectations understate disagreement in the underlying views. In most non-U.S. equity markets, growth is the main source of disagreement. This matters for portfolios: equity allocations load about twice as strongly on growth expectations as on valuation-change expectations. Three mechanisms shape belief formation: managers anchor to peer consensus, use different modeling assumptions, and process information through narratives that differ in complexity and topic attention. Modeling assumptions such as mean reversion and historical calibration predict systematic forecast differences. Narrative complexity and topic attention predict responses to news: managers with more complex or valuation-focused narratives respond less to positive earnings news, whereas attention to dividend yield or downturn risks is associated with stronger responses. Comparisons with N-CSR shareholder letters show that CMA narratives reflect persistent institution-specific investment views. Volatility and correlation beliefs, by contrast, remain backward-looking.

*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, Claudio Tebaldi, and Philippe van der Beck (discussant) 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), FIRS (2026), and WFA (2026).

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

‡London Business School: fbastianello@london.edu

In the first quarter of 2021, PGIM and Amundi both projected that U.S. equities would return 5.8% per year over the following decade. Yet the two institutions arrived at the same forecast for very different reasons. PGIM expected real growth of 2.2%, less than half Amundi's 4.6%, but assumed higher income and a smaller decline in valuations. The two institutions agreed on the final number, but not on the economic forces behind it (Panel A in Table 1). This example illustrates a core tension highlighted in this paper. A published return forecast reveals the final number, but not the beliefs and assumptions that produce it. How much disagreement does the total return forecast conceal? What generates this disagreement, and does it matter for portfolio choice?

[Table 1 here]

In this paper, we aim to characterize the belief-formation framework used by large asset managers to produce long-horizon forecasts of returns, volatilities, and correlations across asset classes, and to use that framework to understand why beliefs differ across institutions and markets. By belief-formation framework, we mean the process that connects economic information to a published forecast.

We characterize four elements of this process: the accounting building blocks into which managers decompose return expectations, the quantitative assumptions used to populate them, the peer anchors that shape revisions, and the causal narratives through which managers interpret news. The building blocks reveal where disagreement lies; the remaining elements explain why it arises and how it changes. Volatility and correlation beliefs provide a contrasting case.

Understanding this process matters because these forecasts are inputs into strategic asset allocation and risk budgeting. Managers with the same total return forecast may respond differently to the same news and choose different portfolios when their underlying beliefs differ. The published total return forecast may therefore understate economically relevant disagreement.

Capital Market Assumptions (CMAs) make these layers observable. Large asset managers and investment consultants publish these documents to guide strategic asset allocation, risk budgeting, and long-horizon investment advice. CMAs report return expectations, volatility forecasts, and correlation forecasts across asset classes. Many also decompose return expectations into building blocks and explain the assumptions behind them. They therefore provide an unusual empirical setting in which we observe the forecast, its reported ingredients, and the text used to explain and justify it. This allows us to characterize large asset managers' belief-formation process using both numerical forecasts and the accompanying text.

We hand-collect CMAs from 83 institutions over 2008–2026 and read them in numbers and in words. In numbers, we extract return expectations, volatility and correlation forecasts, and reported building blocks across equity and fixed-income asset classes. In words, we extract disclosed modeling assumptions and causal narratives. We combine these data with market data, RavenPack earnings-news sentiment, Morningstar mutual fund allocations, machine-learning benchmarks for total returns and their components, shareholder-letter outlooks from SEC Form N-CSR filings, and publicly available biographical information on CMA authors.

Existing work tells us a great deal about the outputs of this process. Institutional subjective premia are countercyclical and move closely with simple real-time objective benchmarks (Dahlquist & Ibert 2024, 2026). CMAs display a systematic risk–return structure across asset classes (Couts et al. 2023, 2024), and institutional expectations pass through to portfolio allocations (Dahlquist & Ibert 2024, Andonov et al. 2026). The literature also provides an important explanation of disagreement in U.S. equity. Many institutions use a common building-block decomposition. Combining income and real earnings growth into a single cash-flow component, Dahlquist & Ibert (2026) use a covariance-share decomposition to show that repricing, or valuation change, accounts for most of the net cross-sectional dispersion in total U.S. equity return expectations.

These findings establish that institutional forecasts are disciplined, heterogeneous, and economically relevant. However, they leave four questions unanswered.

The first arises even within U.S. equity. A covariance-share decomposition identifies which components make the largest net contribution to dispersion in total return expectations. But when the building blocks offset one another, it does not reveal how much component-specific disagreement remains after the other components are held fixed. Nor can a combined income-and-growth block reveal the separate role of growth. If stronger growth beliefs are associated with less favorable repricing or lower income, substantial disagreement about growth can leave only a small trace in the dispersion of total return expectations. Do these offsetting component beliefs also explain how total return expectations move over the cycle?

The second question is whether the anatomy of U.S. equity disagreement generalizes. Does valuation change remain dominant outside the United States? Does the building-block logic extend to fixed income and, if so, which components account for disagreement?

The third question is whether the building blocks are merely an accounting decomposition, with meaningful dynamics only in the total return expectation, or whether the individual components

have dynamics of their own that carry through to the total forecast and to portfolio decisions.

The final question is what produces disagreement within the building blocks and heterogeneous responses to information. Do disclosed modeling assumptions explain why managers populate the same components differently? Do narrative complexity and topic attention explain why managers exposed to the same public news revise their forecasts differently? Relatedly, are volatility and correlation forecasts formed through the same process as return expectations?

Taken together, these questions allow us to characterize the belief-formation framework through which large asset managers form long-horizon beliefs about returns, volatilities, and correlations across asset classes, and to explain why those beliefs differ across institutions and markets.

To answer these questions, we begin with the anatomy of return expectations. First, within U.S. equity, valuation change accounts for 53% of the cross-sectional covariance-share decomposition and nominal growth for 22%. Once the other blocks are held fixed, however, the corresponding contributions are 43% and 35%. The difference reflects strong offsetting relationships: nominal growth is negatively correlated with valuation change, at -0.46, and with income, at -0.34. Thus, valuation change remains the largest net driver of U.S. equity disagreement, but the covariance-share decomposition understates the independent role of growth. The same offset appears over time. Valuation-change expectations are countercyclical: when valuation ratios are high, managers expect lower future repricing. Growth expectations are procyclical: when valuations and macroeconomic conditions are strong, managers expect stronger long-run growth. Total return expectations are countercyclical because valuation dominates on average. This provides a component-level explanation for the aggregate countercyclicality documented by Dahlquist & Ibert (2024): valuation change drives the pattern, while growth dampens it.

Second, U.S. equity is not the template for the rest of the world. Across most non-U.S. equity markets, growth expectations—not valuation change—are the central source of disagreement. Building-block decompositions also appear in fixed income. Income is the largest component, but valuation change and, in high yield, credit losses are economically meaningful. Managers therefore disagree along different margins in different markets.

Third, we show that the building blocks have dynamics of their own. Managers revise each component toward peer consensus, and this component-level convergence carries through to the total return forecast. Using an ex-ante statistical benchmark built only from information available before the CMA reference date, we further find that benchmark-relative deviations align with recent peer

views. The distinctions also matter for portfolio choice. Matching CMA providers to Morningstar mutual fund portfolios, we find that equity allocations load on growth and valuation-change beliefs, but not on income. In the component specification, the coefficient on growth is approximately twice that on valuation change. Component mixes that produce similar total return forecasts can therefore have different portfolio implications.

To answer the final question, we turn from the numbers in the CMAs to their words. The numerical evidence identifies where disagreement lies and whether it is associated with systematic benchmark deviations, but it does not explain why managers exposed to the same macro-financial environment—and often to the same public news—produce different component forecasts and revise them differently.

The reports first disclose the modeling assumptions behind the forecasts, which begin to explain why managers disagree and depart from objective benchmarks. Historical calibration is associated with more optimistic forecasts, whereas a mean-reversion assumption is associated with lower forecasts and more negative valuation-change assumptions. Here, disagreement traces to different inputs: managers who assume different dynamics arrive at different numbers. Because forecasts are produced by identifiable teams within institutions, we examine whether team composition matters. We find that educational composition is associated with component-level deviations from consensus and with the forecasting techniques managers disclose: teams with greater Master's/MBA and PhD representation more frequently report using model-based techniques.

The narratives take us further, revealing that managers also differ in how they process the same information. The reports do not merely state numbers; they make causal claims about how inflation, interest rates, growth, policy, technology, and downturn risks affect return components. We extract these claims as directed, signed causal networks using a large language model, which let us measure two features of managers' stated reasoning: narrative complexity and topic attention. Managers with more complex causal networks react less positively to favorable earnings news and produce more pessimistic forecasts relative to the benchmark. Attention to valuation change is associated with underreaction to positive news, while attention to dividend yield and downturn narratives is associated with overreaction. The structure of managers' stated reasoning thus helps explain how common public information is translated into heterogeneous beliefs.

We validate the interpretation of CMA narratives as institution-specific investment views by comparing them with shareholder-letter outlooks in N-CSR filings. Same-institution CMA–N-CSR

pairs are systematically more similar than cross-institution pairs. The house view expressed in CMAs also appears in the language used to communicate with fund shareholders. CMA narratives are therefore not merely standardized report text; they reflect persistent institution-specific investment views.

Finally, second-moment forecasts provide a contrast to return expectations. Return expectations are revised more, display richer component-level disagreement, and are tied to explicit economic assumptions. Reading the second-moment assumptions stated in the reports and comparing the reported forecasts with their history-based counterparts, we find that volatility and correlation forecasts are backward-looking: equity-volatility forecasts cluster around long-run norms, and pairwise correlation forecasts show limited disagreement, small year-to-year revisions, and close alignment with historical correlations.

Taken together, the evidence points to a simple interpretation. Sophisticated institutions share an accounting grammar, but not the content of their forecasts. The building blocks differ in relevance across markets and can conceal disagreement when they offset one another. Their content is supplied by heterogeneous quantitative assumptions, peer anchors, and causal narratives. A CMA forecast is therefore best understood as the output of a common accounting language combined with heterogeneous belief-formation frameworks.

Literature review

Relative to the closest CMA literature, our first contribution is to broaden what the building blocks reveal about disagreement. Dahlquist & Ibert (2026) show that valuation change is central for U.S. equity when income and growth are treated as a single cash-flow component. We separate growth from income and show that growth and valuation-change beliefs are negatively correlated in U.S. equity and in most non-U.S. equity markets in our sample, so they partly offset one another in the total return expectation. As a result, the total forecast understates how much managers disagree about growth. Growth is an important source of disagreement even within U.S. equity and is the main source of disagreement in most non-U.S. equity markets. We also systematically document and quantify the building-block approach in fixed income by collecting and harmonizing reported numerical decompositions and measuring disagreement over income, valuation changes, roll-down, and credit losses.

More broadly, existing CMA studies establish the cyclicality, cross-asset risk–return structure, and portfolio relevance of institutional beliefs (Dahlquist & Ibert 2024, Couts et al. 2023, 2024,

Andonov et al. 2026). We shift attention from the properties and implementation of these beliefs to the process that produces them. We show how accounting components, quantitative assumptions, peer consensus, and causal narratives jointly shape return expectations and their revisions. We also identify which components of return expectations are most closely associated with portfolio allocations and study how institutions form volatility and correlation forecasts. Our contribution is therefore to characterize how the belief vector studied in this literature is formed.

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.[1] 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 (Hillert et al. 2025) 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 literature by documenting substantial heterogeneity in asset managers' narratives and belief formation frameworks. 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 at-

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

tention 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 a novel 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 within and across asset classes.

Roadmap of the paper

The remainder of the paper proceeds as follows. Section 1 describes the data. Section 2 analyzes expectations in numbers: heterogeneity, building blocks, cyclicality, machine-learning benchmarks, and portfolio allocations. Section 3 examines expectations in words through disclosed modeling assumptions, causal narratives, attention, complexity, and the comparison with N-CSR shareholder letters. Section 4 turns to second-moment forecasts and the stated assumptions underlying them. Section 5 concludes.

1 Data

Our main data source is a hand-collected sample of publicly available CMA reports published by large asset managers and investment consultants. These reports provide long-horizon forecasts across asset classes, together with the accompanying discussion of the economic assumptions behind them. A key innovation of our dataset is that it links, within the same documents, asset managers' qualitative text—economic models, causal narratives, and analytical frameworks—to their quantitative forecasts of returns, volatilities, and correlations. We complement the CMA data with RavenPack news sentiment, Morningstar Direct mutual fund allocations, N-CSR shareholder-letter outlooks, and publicly available information from LinkedIn, personal websites, and company websites on the educational and professional backgrounds of CMA authors.

1.1 Capital Market Assumptions

The institutional significance of CMAs is well established (Dahlquist & Ibert 2024, 2026, Couts et al. 2023, 2024, Andonov et al. 2026). They are essential inputs for strategic asset allocation and long-term investment planning, and they serve as business tools produced independently by institutions—typically on an annual or quarterly basis—by dedicated teams of economists, quantitative analysts, and senior strategists who synthesize macroeconomic trends, asset-pricing models, and academic research.

We hand-collect CMA reports from publicly available sources for both asset managers and institutional investment consultants. For simplicity, we use the term asset managers to refer to both throughout the paper.[2] Our final sample comprises 83 distinct institutions: 59 asset managers and 24 investment consultants.[3] The CMA documents span the years 2008 to 2026, with coverage varying across institutions because of differences in data availability and publication frequency. For temporal analysis, we assign each CMA report a reference date corresponding to the data cutoff used in the institution's forecasting process, when explicitly stated in the report.[4]

Each institution-year CMA provides forecasts across a range of asset classes, including return expectations, volatilities, and pairwise correlations. We focus on major asset classes selected based on three criteria: relevance to institutional portfolios, consistency of appearance within institutions over time, and broad coverage across institutions. Because asset-class labels vary across managers—for example, "U.S. Large Cap Equity" versus "U.S. Equity"—we standardize them using a consistent naming convention.[5] We cover nine equity classes—U.S. Equity, Canada Equity, U.K. Equity, Europe Equity, Japan Equity, Australia Equity, China Equity, Developed ex-U.S. Equity, and Emerging Market Equity—and nine fixed-income classes—U.S. Government Bonds, U.S. TIPS, U.S. Aggregate Bonds, U.S. Investment Grade Corporate Bonds, U.S. High Yield Bonds, U.S. Municipal Bonds, Developed ex-US Government Bonds, Global ex-US Bonds, and Emerging

[2]Including both groups 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.

[3]Table A5 provides detailed information on the organizations included in the dataset along with their respective coverage periods.

[4]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.

[5]If a CMA specifies a benchmark, such as the S&P 500 or MSCI EM, we use that information to guide the mapping. When no benchmark is provided, we rely on textual descriptions of regional, sectoral, currency, maturity, duration, or style exposures. If a manager reports a short-term U.S. Treasury-bill line, such as a three-month bill, in place of an explicit U.S. Cash forecast, we map it to the U.S. Cash class.

Markets Debt. We also collect forecasts for alternative asset classes, including Commodities, Hedge Funds, Private Equity, and Real Estate. Our baseline empirical analysis uses geometric return expectations as reported in the CMA documents, without imposing a uniform forecast horizon; the median horizon is 10 years across asset classes.

1.2 Other Data

We complement the CMA reports with five external data sources. First, we use market data from CRSP, Robert Shiller, Bloomberg, FactSet, and the Goyal predictor dataset to construct realized returns, valuation ratios, market capitalizations, traded-index benchmarks, and the variables used in the forecasting exercises. Second, we use RavenPack earnings-news sentiment to measure the flow of fundamental news arriving between CMA vintages. Third, we match CMA providers to U.S.-domiciled open-end mutual funds in Morningstar Direct to study whether published beliefs are reflected in portfolio allocations. Fourth, we collect shareholder-letter and portfolio-manager outlooks from Certified Shareholder Reports filed with the U.S. Securities and Exchange Commission (SEC), primarily Forms N-CSR and N-CSRS, to test whether CMA narratives capture persistent institution-specific investment views. Fifth, we use publicly available information from LinkedIn, personal websites, and company websites to collect authors' educational attainment, job titles, and years of professional experience at the time of each report. Appendix A.1 describes the construction and matching procedures for each external data source.

2 First Moments: Evidence from Numbers

We start with a simple fact: asset managers disagree, and this disagreement is not confined to return forecasts. It appears across return expectations, volatility forecasts, and correlation forecasts. To map this heterogeneity, we use a forecast-specific fixed-effects decomposition, in the spirit of Dahlquist & Ibert (2026), and apply it to the broader set of CMA forecast variables in our sample: return expectations across major equity and fixed-income classes, volatilities, and pairwise correlations. The exercise is descriptive. It shows where disagreement is large and persistent before we open return forecasts into their building blocks and study the dynamics of second-moment forecasts.

Figure 1 implements this decomposition for return expectations, volatility forecasts, and correlation forecasts across the main equity and fixed-income asset classes in our CMA sample. The

exercise is deliberately forecast-specific. For each forecast variable—for example, the return expectation for U.S. equities, the volatility of U.S. High Yield Bonds, or the correlation between U.S. equities and U.S. Government Bonds—we ask whether variation is explained mainly by persistent differences across managers or by common time-series movements.

For each forecast variable k, we estimate:

$$x_{i,k,t} = \alpha_i^k + \tau_t^k + \varepsilon_{i,k,t} \tag{1}$$

where $x_{i,k,t}$ denotes asset manager i's forecast for variable k at time t, α_i^k are asset-manager fixed effects, and τ_t^k are year fixed effects. We then use the partial R^2 associated with each set of fixed effects to measure the share of variation explained by manager identity and by common time variation. The decomposition is estimated separately for each forecast variable, so it is not driven by variation across asset classes.

The equity forecast set includes U.S., Canada, U.K., Europe, Japan, Australia, China, Developed ex-U.S., and Emerging Markets. The fixed-income forecast set includes U.S. Government Bonds, U.S. TIPS, U.S. Aggregate Bonds, U.S. Investment Grade Corporate Bonds, U.S. High Yield Bonds, U.S. Municipal Bonds, Developed ex-US Government Bonds, Global ex-US Bonds, and Emerging Markets Debt.[6]

Panel A of Figure 1 reports the decomposition for return expectations. Equity-return forecasts display substantial persistent cross-manager heterogeneity. Asset-manager fixed effects explain between about 32% and 67% of the variation across equity markets, while year fixed effects explain between about 6% and 27%. Fixed-income forecasts display a related pattern. Asset-manager fixed effects explain between about 19% and 51% of the variation, while year fixed effects explain between about 37% and 64%. Fixed-income expectations appear more tightly connected to common time-series conditions, consistent with the idea that starting yields—and, for credit markets, spreads—provide strong observable anchors.

Panels B and C extend the same decomposition to second-moment forecasts. Panel B shows that volatility forecasts are strongly manager-specific. Across the main equity and fixed-income variables, asset-manager fixed effects explain between about 59% and 82% of the variation, while

[6]This set provides the richest universe of public asset classes for which we can collect repeated return expectations, volatility, and correlation forecasts. Figure A8 reports analogous decompositions for alternative asset classes—Commodities, Hedge Funds, Private Equity, and Real Estate.

year fixed effects explain between about 6% and 32%. Panel C shows a similar pattern for many pairwise correlation forecasts. Asset-manager fixed effects explain between 25% and 90% of the variation across asset-class pairs, while year fixed effects explain between 5% and 67%. Thus, persistent differences across managers are not limited to return expectations. They also appear in beliefs about volatilities and correlations, and this heterogeneity is pervasive across asset classes (as shown in Figure A8). This manager-specificity, however, conceals an important fact. As discussed in Section 4, most asset managers report constructing their volatility and correlation inputs from historical estimates. Taken at face value, this common reliance on history sits uneasily with the large asset-manager fixed effects documented above: if managers estimate second moments from broadly the same historical record, their forecasts should look alike. The tension is resolved once we recognize that reliance on history is not uniform—managers differ in how much weight they place on the distant versus the recent past. We show below that accommodating this heterogeneity through a simple exponentially weighted moving average (EWMA), in which the decay parameter is allowed to vary across managers, uncovers a strong relationship between the reported forecasts and a backward-looking benchmark. Under this reading, the manager fixed effects largely reflect heterogeneous—but still backward-looking—weighting schemes rather than genuinely idiosyncratic views about volatilities and correlations.

[Figure 1 here]

These facts organize the rest of the analysis. We first study return expectations. Sections 2.1, 2.2, 2.3, and 2.4 decompose return expectations into building blocks, study their cyclicality, compare them with machine-learning benchmarks, and relate them to portfolio allocations. Section 4 then turns to second-moment beliefs.

2.1 An Approximate Building Block Framework

A natural way to understand disagreement in long-horizon return expectations is to open the forecast into its components. For U.S. equities, Dahlquist & Ibert (2026) show that most institutions use a building-block approach and that disagreement about repricing, or valuation change, explains the largest share of cross-sectional disagreement. Their empirical decomposition groups income and real earnings growth into a single cash-flow component, which is natural for broad coverage but leaves open how much of that disagreement reflects expected growth rather than payout. We take these results as the starting point and use our harmonized component data to ask three further questions.

First, does growth matter as a separate source of disagreement once it is disentangled from income? Second, is U.S. equity special, or do other equity markets display a different decomposition of disagreement? Third, how do building-block decompositions help explain disagreement in fixed-income return forecasts?

CMA reports make this exercise possible because many managers report numerical decompositions of return expectations. We collect these decompositions across major equity and fixed-income asset classes and standardize the reported labels into common component categories. Building-block-style approaches also appear in reports for commodities, real estate, private equity, private credit, and hedge funds, but our main analysis focuses on equities and fixed income, where coverage and comparability are strongest.

The reported decompositions are expressed in raw, non-log units. We therefore work with approximate return identities in the same units. Appendix A.4 derives the corresponding identities. For equities, the approximate h-period return identity is:

$$\widetilde{\mathbb{E}}_{i,t}[R^{Eq}_{t,t+h}] \approx \widetilde{\mathbb{E}}_{i,t}[DY_{t,t+h}] + \widetilde{\mathbb{E}}_{i,t}[VC_{t,t+h}] + \widetilde{\mathbb{E}}_{i,t}[g_{real,t,t+h}] + \widetilde{\mathbb{E}}_{i,t}[\pi_{t,t+h}] \tag{2}$$

where DY is the payout or income component, commonly reported as "income," "net payout yield," "carry," or "dividend yield"; VC is valuation change, often reported as "valuation adjustment," "repricing," or "change in valuation"; g_{real} is real earnings or cash-flow growth; and π is inflation. When managers report growth and inflation jointly, we use the reported nominal-growth component. For fixed income, the corresponding approximate identity is:

$$\widetilde{\mathbb{E}}_{i,t}[R^{FI}_{t,t+h}] \approx \widetilde{\mathbb{E}}_{i,t}[Carry_{t,t+h}] + \widetilde{\mathbb{E}}_{i,t}[Roll_{t,t+h}] + \widetilde{\mathbb{E}}_{i,t}[Val.Change_{t,t+h}] - \widetilde{\mathbb{E}}_{i,t}[CreditLoss_{t,t+h}] \tag{3}$$

Here, $Carry$ captures coupon income and yield accrual, $Roll$ captures the return from rolling down the yield curve under an unchanged curve, $Val.Change$ captures mark-to-market effects from expected changes in risk-free rates and credit spreads, including convexity effects, and $CreditLoss$ captures expected default losses.

We first study cross-sectional disagreement within each year. To do so, we remove year fixed effects from both total return forecasts and reported building blocks. Figure 2 reports two complementary decompositions of the remaining dispersion. The blue bars show the covariance-share decomposition, which measures each block's net contribution to disagreement in the total return

forecast. The red bars show the Type II ANOVA contribution, which measures the disagreement specific to each block after holding fixed the other blocks. The blue bars answer which components drive dispersion in total return expectations. The red bars answer where forecasters disagree in the underlying decomposition. The gray numbers in the background of each panel report the corresponding within-year correlations among the building blocks.

When the components are uncorrelated, the two decompositions agree. When they are correlated, they diverge: a positive correlation raises a block's covariance share relative to its independent contribution, while a negative, offsetting correlation lowers it—so the covariance share can hide economically important disagreement. U.S. equity illustrates this point. Consistent with Dahlquist & Ibert (2026), the covariance-share decomposition assigns the largest role to valuation change: VC accounts for 53% of the dispersion in total U.S. equity return forecasts, while nominal growth accounts for 22%. The ANOVA decomposition gives a more balanced view of the underlying disagreement: valuation change still matters, with a contribution of 43%, but nominal growth rises to 35%. This difference is consistent with the correlations reported in the same panel. Valuation change and nominal growth are strongly negatively correlated, -0.46, and nominal growth is also negatively correlated with dividend yield, -0.34. Some managers therefore reach similar total return forecasts through opposite growth–valuation splits.

The U.S. case is not representative of the rest of the world. Across most non-U.S. equity markets, nominal growth is the main source of disagreement. Developed ex-U.S. equity, emerging markets, the U.K., Australia, and Japan all assign the largest covariance-share contribution to nominal growth. Thus, the main equity result is not simply that valuation matters. It is that the source of disagreement changes across markets: valuation change dominates U.S. equity disagreement, while growth is the central source of disagreement for most non-U.S. equity markets.

Figure 3 repeats the analysis with asset-manager fixed effects in place of year fixed effects, thereby isolating how each manager's expectations vary over time around their own average. The same pattern emerges within managers: both growth and valuation-change expectations account for a large share of the independent variation, yet their contributions to the dispersion of total return expectations partly offset one another, reflecting a negative within-manager correlation between the two. A given forecaster thus tends to revise growth and valuation expectations in opposite directions over time, mirroring the cross-manager trade-off documented above.

[Figures 2 and 3 here]

Figures A9 and A10 repeat the decomposition for fixed-income forecasts. Disagreement is more concentrated than in equities. Income is the largest component across U.S. Government, Investment Grade, High Yield, and Aggregate Bonds, but valuation-change disagreement is non-negligible, especially for government and aggregate bonds, while credit losses matter mainly for high yield. Within managers, fixed-income variation is dominated by income, which explains more than 80% of the variation.

Three facts emerge from the component analysis. First, for U.S. equity, valuation change remains the largest net driver of disagreement, but growth is a large separate source of underlying disagreement once it is measured separately from income. Second, outside the United States, equity disagreement is usually more about growth than valuation change. Third, fixed-income forecasts also contain economically meaningful building-block disagreement: income matters, but valuation-change and credit-loss components are also important. Managers therefore disagree along different margins across markets. We next ask whether those margins vary systematically over the cycle.

2.2 Campbell-Shiller Decomposition: Cyclicality

In line with studies examining 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 guide our analysis of asset managers' return-expectation building blocks. Following Campbell & Shiller (1987, 1988), log-linearizing the one-period return identity yields:

$$pd_t = \kappa + \Delta d_{t+1} - r_{t+1} + \rho\, pd_{t+1} \tag{4}$$

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 ($\widetilde{\mathbb{E}}_{i,t}$), taking the covariance of both sides with respect to pd_t, dividing by the variance of pd_t, and subtracting ρ from both sides, we obtain:

$$1-\rho \;=\; \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}], pd_t\right)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{\Delta d}, pd)}} - \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[r_{t+1}], pd_t\right)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{r}, pd)}} + \rho \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}], pd_t\right)}{\mathrm{Var}(pd_t)}}_{\beta^i_{(\widetilde{\Delta pd}, pd)}} \tag{5}$$

where Δpd_{t+1} is the one-period change in the log price–dividend ratio. The terms on the right-hand side are the regression coefficients of cash-flow growth expectations, returns expectations, and valuation-change expectations on the price–dividend ratio. 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. 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. Following the same procedure, we obtain:

$$1 - \rho = \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[\Delta x_{t+1}], px_t\right)}{\mathrm{Var}(px_t)}}_{\beta^i_{(\widetilde{\Delta x}, px)}} - \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[r_{t+1}], px_t\right)}{\mathrm{Var}(px_t)}}_{\beta^i_{(\widetilde{r}, px)}} + \rho \underbrace{\frac{\mathrm{Cov}\left(\widetilde{\mathbb{E}}_{i,t}[\Delta px_{t+1}], px_t\right)}{\mathrm{Var}(px_t)}}_{\beta^i_{(\widetilde{\Delta px}, px)}} \tag{6}$$

where Δpx_{t+1} is the one-period change in the log price–earnings ratio. Equations (5) and (6) hold at the individual or consensus level and motivate our empirical analysis: the cyclicality of return expectations must reflect the cyclicality of expected cash-flow growth, expected valuation changes, or both. We focus on institution-level regressions for managers with at least five years of data, thereby avoiding aggregation issues related to changes in sample composition.

We begin with U.S. equities. For each asset manager i and forecast component $b \in \{ER$ (Return Expectation), NG (Nominal Growth), VC (Valuation Change)$\}$, we estimate regressions of the form:

$$F^b_{i,t,h} = \alpha^b_i + \beta^b_{i,pd}\, pd_t + \varepsilon^b_{i,t,h} \tag{7}$$

$$F^b_{i,t,h} = \alpha^b_i + \beta^b_{i,px}\, px_t + \varepsilon^b_{i,t,h} \tag{8}$$

where $F^{ER}_{i,t,h}$ is the total equity return forecast, $F^{NG}_{i,t,h}$ is the nominal-growth component, $F^{VC}_{i,t,h}$ is the valuation-change component, pd_t is the log price–dividend ratio of the S&P 500 from Robert Shiller's website, and px_t is the log cyclically adjusted price–earnings ratio from Robert Shiller's website.

Figure 4 reports the results. Consistent with Dahlquist & Ibert (2024), 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, with the growth component dampening the overall cyclicality. Thus, the countercyclicality of total return expectations conceals an opposing procyclical component in growth beliefs.

[Figure 4 here]

We extend the cyclicality analysis to other equity markets and U.S. fixed-income asset classes. Figure 5 reports, for each market and building block, the share of managers with a positive loading on the relevant CAPE ratio; positive loadings indicate procyclicality, while negative loadings indicate countercyclicality. In every equity market shown, at least half of the managers have countercyclical total return expectations, with the share ranging from 50% in Japan to 82% in developed ex-U.S. equity. The components producing this pattern, however, differ substantially across markets. The share of managers with procyclical valuation-change expectations ranges from 6% in the U.S. to 44% in Europe, while the corresponding share for growth expectations ranges from 27% in developed ex-U.S. equity to 80% in Canada. For U.S. fixed income, between 75% and 83% of managers have countercyclical total return expectations. In each fixed-income class, the manager-level classification of the income component coincides with that of the total return forecast, indicating that income drives the common cyclical pattern. Thus, countercyclical total return expectations are widespread, but the equity components generating them vary markedly across markets; the fixed-income pattern is more uniform and concentrated in income.

[Figure 5 here]

The substantial heterogeneity in asset managers' equity forecasts and their underlying components raises a natural benchmark question. The Campbell–Shiller exercise shows how managers' forecasts move with valuation ratios, but it does not tell us whether those forecasts are well calibrated to a broader real-time information set (Bianchi et al. 2022, 2025).

We therefore compare CMA expectations with an ex-ante machine-learning benchmark based only on information available when asset managers formed their views. The recent nature of our sample makes this comparison feasible: the regularized, tree-based, and neural-network methods we use were already available during the period covered by the CMA data, allowing us to construct a benchmark that a real-time forecaster could plausibly have implemented.

We deliberately avoid simple benchmarks based on valuation ratios, which are closely related

to the reported valuation-change component and may therefore mechanically generate strong correlations with CMA forecasts that explicitly reference those ratios. Instead, we use non-linear machine-learning methods that combine a broader set of potentially relevant macro-financial predictors.

The benchmark is a real-time statistical yardstick for assessing whether CMA forecasts and their building blocks deviate systematically from the information that predicts realized returns, including whether managers place relatively more weight on some variables and less on others. This allows us to compare the predictors associated with managers' beliefs with those that forecast realized returns.

2.3 CMA Forecasts and Objective Benchmarks

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 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.5).[7] 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 improve robustness when no single specification is known to be dominant.

While total return expectations are the natural starting point, we also use the same methodology to construct a separate benchmark for each equity building block. Appendix A.5 shows that the machine-learning forecasts systematically outperform CMA forecasts for total returns.

We next use these benchmarks in two ways. First, we compare the predictors that explain CMA beliefs with those that predict realized returns. Second, we use the benchmarks to examine whether managers' forecasts are anchored to the recent CMA consensus.

[7]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.

2.3.1 Feature Importance: Objective vs. Subjective

To understand the source of differences between asset managers' return 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, $R_{t,t+h}$, thereby generating the objective benchmark $G_{t,h}$. Second, we use the same models to predict asset managers' total CMA return 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 the dynamics of CMA forecasts.

Figure 6 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 greatest importance to valuation variables and the investment-to-capital ratio.

Panel (b) compares the ensemble importance recovered from the realized-return models with that 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; the vertical axis reports the equal-weighted ensemble importance from the models fitted to total CMA return forecasts. 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 6 here]

The comparison suggests that the machine-learning models use different signals when predicting realized returns and 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. By contrast, the realized-return models assign relatively greater importance to valuation ratios, particularly the earnings–price and dividend–price ratios. Thus, somewhat surprisingly, although asset managers frequently re-

fer to valuation ratios, the machine-learning models that best approximate their forecasts place less weight on these variables than do the models forecasting realized returns. We therefore turn next to the dynamics of asset managers' return expectations to better understand the origins of their benchmark-relative deviations, $G_{t,h} - F^{ER}_{i,t,h}$ (Bianchi et al. 2022).

2.3.2 Forecast Dynamics: Convergence and Consensus Anchoring

We begin by asking whether managers whose forecasts lie away from the cross-sectional consensus subsequently revise toward it.

For each asset class with an explicit return decomposition, we consider both the total return expectation and each of its building blocks, indexed by b. For equities, these building blocks are valuation change, growth, and dividend yield; for fixed income, they are income, roll-down, credit loss, and valuation change. We estimate the following forecast-change specification separately for each forecast variable:

$$\Delta F^b_{i,t} \quad = \quad \alpha^b_i + \beta^b Dev^b_{i,t-1} + \delta^b \Delta \overline{F}^{\,b}_{-i,t} + \varepsilon^b_{i,t} \tag{9}$$

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

$$Dev^b_{i,t-1} = F^b_{i,t-1} - \overline{F}^{\,b}_{-i,t-1}, \qquad \overline{F}^{\,b}_{-i,t-1} = \frac{1}{N^b_{t-1} - 1} \sum_{\substack{j \in \mathcal{I}^b_{t-1} \\ j \neq i}} F^b_{j,t-1}.$$

Here, $\mathcal{I}^b_{t-1}$ denotes the set of asset managers reporting component b for the relevant asset class in year $t-1$, and $N^b_{t-1} = |\mathcal{I}_{t-1}|$. Thus, $\overline{F}^{\,b}_{-i,t-1}$ 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,

$$\Delta \overline{F}^{\,b}_{-i,t} = \overline{F}^{\,b}_{-i,t} - \overline{F}^{\,b}_{-i,t-1},$$

which captures information arriving between $t-1$ and t that is reflected in the forecasts of other managers.

The coefficient of interest is β^b. Under the null that a manager's lagged position relative to its

peers has no incremental predictive content for its subsequent revision, $\beta^b = 0$. 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 2 presents the results. Asset classes appear in columns and forecast variables in rows, with each asset-class–forecast-variable cell corresponding to a separate regression. The coefficient on the lagged deviation from the leave-one-out consensus is negative in every one of the forty-three estimated cells and statistically significant at the 5% level in thirty-seven. The magnitudes are economically large: the median coefficient is -0.837, implying that, conditional on the contemporaneous movement in consensus, roughly five-sixths of the prior gap is reversed within one year. Managers' relative positions therefore do not remain fixed. Forecasts systematically converge toward the center of the cross-sectional belief distribution, both for total return expectations and for the underlying building blocks.

The contemporaneous change in the leave-one-out consensus is positive in all forty-three cells and statistically significant in twenty-nine. Thus, individual revisions also comove positively with contemporaneous revisions by other managers, consistent with common information and possibly peer influence. Importantly, the convergence result survives after controlling for this common contemporaneous component. The negative estimates of β^b therefore capture more than common contemporaneous movements: a manager's prior position relative to the peer consensus contains incremental predictive power for its subsequent revision.

For total return expectations, convergence is stronger across fixed income: the median coefficient is -1.112 across the four fixed-income sectors, compared with -0.682 across the six equity markets. Within equities, convergence is strongest for valuation change and weakest for dividend yield. Dividend-yield revisions also display the weakest contemporaneous comovement with the peer consensus.

[Table 2 here]

Specification (9) documents convergence, but it does not by itself distinguish anchoring from rational responses to information embedded in peer forecasts. To sharpen the test, we exploit the long history of U.S. equity data to construct out-of-sample machine-learning benchmarks. Following the anchoring literature (Tversky & Kahneman 1974, Campbell & Sharpe 2009, Bastianello 2025), we ask whether managers' benchmark-relative deviations align with the wedge between a

salient anchor and the 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. Because the CMA reports decompose the total return forecast into building blocks, we estimate the same specification for the total return forecast and each component. Specifically, for each forecast variable $b \in \{ER, VC, g, DY, \pi\}$, we estimate:

$$G^b_{t,h} - F^b_{i,t,h} \;=\; \alpha^b_i + \beta^b \left(G^b_{t,h} - \overline{F}^{b,365d}_{t-1,h} \right) + \varepsilon^b_{i,t,h} \tag{10}$$

where $F^b_{i,t,h}$ is manager i's forecast for variable b at report date t and horizon h, $G^b_{t,h}$ is the machine-learning benchmark forecast for the same variable, reference date, and horizon, and $\overline{F}^{b,365d}_{t-1,h}$ is the rolling consensus constructed from horizon-matched CMA forecasts for the same variable issued during the previous 365 days. Setting $b = ER$ recovers the total-return specification.

The key parameter is β^b. If managers anchor on the recent CMA consensus, their forecasts should tend to lie on the same side of the machine-learning benchmark as the recent consensus. Hence, a positive and statistically significant β^b 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 3 presents the results. Panel A reports the total-return specification. Column (1) shows a large and statistically significant coefficient of 0.745, 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 benchmark-minus-consensus wedge 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 wedge, $G_{t-1,h} - \overline{F}^{ER,365d}_{t-1,h}$, 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. Thus, the lagged benchmark–consensus wedge predicts current benchmark-relative forecast deviations: 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.

Panel B decomposes this result across building blocks and reveals that the total-return coefficient masks pronounced heterogeneity. Anchoring is strongest for the two macroeconomic blocks: the coefficient is 0.958 for inflation and 0.817 for real growth, both significant at the 1% level, and the benchmark–consensus wedge alone accounts for 44.1% and 28.9% of the within-manager variation, respectively. The valuation-change coefficient is positive but less than half the total-return magnitude at 0.331, with a within-R^2 of only 0.061. The dividend-yield coefficient is negative, at -0.134, marginally significant, and explains little within-manager variation, with a within-R^2 of 0.015; it therefore provides no evidence of positive consensus anchoring for this block.

The weaker valuation-change result is consistent with heterogeneity in the assumptions underlying this component. As shown in Section 3.1, managers differ in whether they impose mean reversion in valuations. Managers who do not impose mean reversion may project little valuation change regardless of the recent consensus, reducing their sensitivity to the anchor. Because the pooled estimate averages across managers with different sensitivities to the anchor, this slope heterogeneity can attenuate β^{VC} and contribute to the low within-R^2. Asset-manager fixed effects absorb level differences across managers, but not this slope heterogeneity.

The total-return coefficient of 0.745 therefore lies between the strong macroeconomic coefficients and the weaker valuation-change estimate, while dividend yield shows no comparable positive anchoring. This pattern is consistent with the total return forecast inheriting consensus anchoring primarily through its macroeconomic inputs. Column (2) confirms that this ranking is not an artifact of contemporaneous benchmark variation: using the lagged benchmark, the macro blocks retain the largest coefficients (0.720 for inflation and 0.312 for real growth), valuation change attenuates to 0.178, and the dividend-yield coefficient becomes statistically indistinguishable from zero.

[Table 3 here]

The two exercises point to a coherent view of forecast dynamics. Across markets and building blocks, managers systematically revise toward the peer consensus. Relative to the machine-learning benchmark, their deviations also align with the recent consensus. Taken together, these results are consistent with peer consensus serving as an economically meaningful anchor rather than merely a statistical average of individual forecasts. We next ask whether heterogeneity in the building blocks matters for institutional investors' portfolio choices.

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:

$$w_{f,t} = \mu_f + \tau_t + \boldsymbol{\beta}' \mathbf{X}_{i(f),t} + \varepsilon_{f,t} \tag{11}$$

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 4 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: valuation change, nominal growth, and dividend yield. 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 0.24, 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 com-

ponents 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 4 here]

3 First Moments: Evidence from Text

Having established the first-order facts from the numerical return-expectation forecasts in Section 2, we now turn to the text of the reports to uncover the mechanism behind the return heterogeneity in asset managers' forecasts.

We convert each CMA report into a structured digital format through an optical-character-recognition and machine-learning workflow that preserves reading order, recovers document hierarchy, and extracts embedded visuals. The resulting data undergo an initial preprocessing stage and then enter our large language model (LLM) pipeline (Appendix A.2).[8]

Building on the numerical analysis, we proceed in four steps. First, we examine disclosed modeling assumptions and the composition of authoring teams. Second, we use free-text modeling-assumption phrases to identify recurring forecasting approaches. Third, we extract causal networks from CMA narratives and measure their complexity and topic attention. Fourth, we relate these textual features to news updating and benchmark-relative forecast errors, and validate their institution-specific content using N-CSR shareholder letters.

As a final validation, we compare the economic content of CMA narratives with the shareholder-letter outlooks that affiliated fund managers provide in N-CSR filings.

[8] In Appendix A.3.1, we analyze word counts over time and across asset managers, revealing disparities in reporting practices.

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 valuation change, growth, dividend yield, and inflation. Together, these methodological choices and team characteristics provide a window into the forecasting process behind the published numbers.

Preset Modeling Assumptions. We extract managers' disclosed modeling assumptions from the CMA reports in two complementary ways. This paragraph describes the first, *preset* approach: we fix a short list of forecasting techniques in advance and use LLM-based text classification to detect, for each report, which of these preset techniques the text discloses. The second, *free-text* approach imposes no preset list and is described below. Specifically, we first classify the report text and then define manager-year preset technique indicators equal to one if at least one report by asset manager i in year t is classified as disclosing the corresponding preset technique.[9] Let:

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

denote the set of preset 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

[9] Appendix A.2.7 details the extraction methodology.

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 7 reports the prevalence of forecasting techniques across CMA reports. Historical 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 7 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 total U.S. equity return forecast and its main reported components: ER denotes the total 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; VC denotes valuation change, g real growth, DY dividend yield, and π inflation. For the total return expectation and each forecast component, $b \in \{ER, VC, g, DY, \pi\}$, we estimate:

$$Dev_{i,t}^{b} = \alpha^{b} + \tau_{t}^{b} + \sum_{k \in \mathcal{K}} \beta_{k}^{b} \mathbf{1}\{\text{technique}_{k,i,t}\} + \varepsilon_{i,t}^{b} \tag{12}$$

The year fixed effects τ_t^b 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 5 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 total 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 5 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.3.2 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 a graduate degree or an MBA. 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 both to the assumptions embedded in published CMA forecasts and to the forecasting techniques used to produce them.

We focus on educational composition as one observable dimension of team human capital. Panel (a) of Figure 8 reports the Pearson correlation between the number of authors whose highest degree is a Bachelor's, Master's/MBA, or PhD and the manager's deviation from the contemporaneous leave-one-manager-out consensus across six U.S. equity components: equity return, real growth, inflation, nominal growth, dividend yield, and valuation change. These correlations are computed at the report level. Because the education variables are counts, the panel 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 in Panel (a) is that PhD-trained authors are associated with a distinct set of deviations from consensus. Bachelor's and Master's/MBA counts are positively correlated with deviations in real and nominal growth, 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's or Master's/MBA authors tend to deviate below consensus, whereas reports with more PhD authors tend to deviate above consensus.

Panel (b) performs an analogous exercise for the preset forecasting techniques declared by managers in their equity-return forecasts. For this analysis, observations are collapsed to the manager level by averaging author counts and technique indicators across all of each manager's reports. The panel reports the correlation between managers' average educational composition and their propensity to disclose the use of regressions, simulations, historical averages, mean reversion, trend-based methods, or other techniques. Managers with more Master's/MBA- or PhD-trained authors are generally more likely to report the use of explicitly model-based techniques, particularly regressions and mean-reversion methods. PhD representation is most strongly associated with the use of regressions, while both Master's/MBA and PhD representation are negatively associated with techniques classified as Other. Bachelor's representation displays weaker and less systematic relationships, with the clearest positive association appearing for trend-based methods.

Taken together, the two panels suggest that educational composition, and in particular the presence of PhD-trained authors, is related not only to how asset managers allocate return expectations across underlying building blocks, but also to the forecasting methods they disclose.

[Figure 8 here]

Free-Text Modeling Assumption Clusters and Return Expectations. The five preset technique indicators above capture broad categories of forecasting techniques. We now turn to the second, *free-text* approach: a more granular analysis of the actual modeling descriptions used in CMA reports, imposing no preset list of techniques. Rather than recording only whether a report discloses one of the preset techniques (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.[10] 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.

[10] Appendix A.2.7 details the extraction methodology.

Figure 9 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 6 defines the seven clusters and reports their prevalence across reports.[11] Let:

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

denote the set of free-text 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 9 and Table 6 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 7 relates the technique clusters to managers' positions relative to their peers. For the total 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.

[11] We apply the same extraction procedure to fixed income (government bonds, high yield, and investment grade), commodities, private equity, hedge funds, and real estate. Figure A11 and Table A6 show that the techniques used across these asset classes are broadly similar to those used for equities.

Managers whose reports disclose "Mean-reversion" techniques report total 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 7 here]

Table 8 asks whether the same technique clusters are associated with deviations from the objective machine-learning benchmark. The dependent variable is the manager's total U.S. equity return forecast minus the objective benchmark forecast, matched at the same reference date and horizon. We estimate:

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

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 8 here]

3.2 Causal Networks

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 their Capital Market Assumption (CMA) reports, articulating how economic variables transmit to returns. We extract these stated causal mechanisms and represent them as directed networks, allowing us to quantify heterogeneity in managers' belief formation frameworks—both which topics receive attention and how complex the causal reasoning structure is. Our extraction pipeline uses a large language model to identify explicit causal claims in each report and to map them onto a standardized financial taxonomy; the full procedure is described in Appendix A.2 and illustrated in Appendix Figure A1.

The network representation allows us to construct two families of text-based measures. The

first captures *narrative complexity*: how indirect and interconnected a manager's causal reasoning is. The second captures *topic attention*: which economic topics receive the most emphasis. 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.

Within narrative complexity we distinguish two dimensions. *Chain complexity* captures whether managers explain equity returns through direct links or through multi-step causal chains, and is measured by the average path length between topics and by the indirect connection ratio—the share of links belonging to multi-step paths leading to or from asset class return expectations. *Interconnected complexity*, measured by network transitivity, captures the extent to which the causal narrative forms closed triangles of relationships and is therefore more locally interconnected. High transitivity can reflect a richer belief formation framework, but need not imply more accurate forecasts; whether such complexity improves or distorts belief formation is an empirical question. Appendix A.2.3 gives formal definitions and a schematic illustration.

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, capturing how much attention a manager devotes to it regardless of whether the overall network is simple or complex.

We refer to these directed networks as "causal networks" and use them throughout the rest of the paper. By "causal," we mean that the network encodes directional links stated by managers in their narratives, rather than causal effects in the econometric sense. Return identities such as the Campbell–Shiller decomposition—together with structural asset-pricing models—imply a natural set of links among subjective expected returns, expected cash-flow growth, and expected valuation changes. Appendix A.2.5 develops this conceptual foundation and illustrates, using a consumption-based asset-pricing model, how economic structure enriches the baseline accounting network (Figure A4). Appendix A.2.4 documents the substantial heterogeneity in the attention and complexity measures across asset managers' causal networks, and highlights that the main topics asset managers refer to when discussing return expectations are valuation change and growth, consistent with our evidence on the building blocks.

Figure 10 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 path-

ways connecting policy and regulations, interest rates, and dividend yield. This structural mapping enables systematic comparison across institutions to identify where sophisticated investors differ in their views of market mechanics. In this network, 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.

[Figure 10 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:

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_W^m \left(C_{i,t}^m - \overline{C}_i^{\,m}\right) + \gamma_W^m \left(C_{i,t}^m - \overline{C}_i^{\,m}\right) \times S_t + \gamma_B^m \overline{C}_i^{\,m} \times S_t + \varepsilon_{i,t}^m \tag{13}$$

where $F_{i,t}^{ER}$ denotes asset manager i's total 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.2.4 for definitions of all network measures).

Table 9 shows that causal-network complexity is not directly associated with the level of return expectations: the coefficients on the within component $C_{i,t}^m - \overline{C}_i^m$ 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, $(C_{i,t}^m - \overline{C}_i^m) \times S_t$, is weakly significant only when complexity is measured by network transitivity. Conversely, the interaction between structural, between-manager complexity and sentiment, $\overline{C}_i^m \times S_t$, 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 9 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 $A_{i,t}^q$ denote asset manager i's topic-attention measure in year t, and let:

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

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

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) + \gamma_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) \times S_t + \gamma_B^q \overline{A}_i^q \times S_t + \varepsilon_{i,t}^q \tag{14}$$

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 10 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 10 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 10 here]

We next examine whether the structural complexity of managers' causal networks is associated with how beliefs change over time. Complex belief formation frameworks 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

$$F_{i,t}^{ER} - F_{i,t-1}^{ER} = \mu_i + \tau_t + \beta_W^m \left(C_{i,t}^m - \overline{C}_i^{\,m}\right) + \gamma_W^m \left(C_{i,t}^m - \overline{C}_i^{\,m}\right) \times S_t + \gamma_B^m \overline{C}_i^{\,m} \times S_t + \varepsilon_{i,t}^m \tag{15}$$

where $F_{i,t}^{ER} - F_{i,t-1}^{ER}$ is the year-over-year change in asset manager i's total 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 11 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 11 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:

$$F_{i,t}^{ER} - F_{i,t-1}^{ER} = \mu_i + \tau_t + \beta_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) + \gamma_W^q \left(A_{i,t}^q - \overline{A}_i^q\right) \times S_t + \gamma_B^q \overline{A}_i^q \times S_t + \varepsilon_{i,t}^q \tag{16}$$

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

Table 12 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 12 here]

To investigate whether network complexity predicts systematic deviations from objective fore-

casts, we employ a within-between decomposition. For each complexity measure m, we estimate:

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) + \gamma^m_W \left(C^m_{i,t} - \overline{C}^m_i\right) \times S_t + \gamma^m_B \overline{C}^m_i \times S_t + \varepsilon^m_{i,t,h} \tag{17}$$

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 13 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 13 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 devi-

ations from the objective benchmark. For each topic q, we estimate:

$$G_{t,h} - F^{ER}_{i,t,h} = \mu_i + \tau_t + \beta^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) + \gamma^q_W \left(A^q_{i,t} - \overline{A}^q_i\right) \times S_t + \gamma^q_B \overline{A}^q_i \times S_t + \varepsilon^q_{i,t,h} \tag{18}$$

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 14 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 14 here]

3.3 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 11 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 11 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 Second Moments: Evidence from Numbers and Text

Having studied how managers form return expectations, we now turn to second-moment forecasts and examine their drivers.

By reading the CMA reports across our providers, we find that the frameworks behind second-moment forecasts are far simpler and more homogeneous than those used for first-moment (return expectation) forecasts. Disclosed methodologies fall into three broad categories: simple historical estimates; historical estimates with small adjustments (mostly to reflect scenario-dependent assumptions or regime overlays); and calibration exercises (for example, historical regressions or GARCH models). Table 15 reports the methodological approach each asset manager uses for its volatility and correlation forecasts. Among providers that disclose a methodology, 33% use purely

historical estimates for volatility, 41% use historical estimates with small adjustments, and 27% use calibration-based methods. The corresponding shares for correlation are 35%, 38%, and 27%. Taken together, roughly three-quarters of asset managers rely on historical estimates—whether unadjusted or lightly adjusted—rather than formal calibration when forming their second-moment forecasts. A minority of providers do not disclose their methodology; we group these separately and label the three disclosed approaches *Estimates* (purely historical), *Estimates** (historical with adjustment), and *Calibrations* (model-based).

[Table 15 here]

The disclosed methodologies provide initial evidence that second-moment forecasts are largely history-based. At first glance, this is hard to square with the pervasive manager-specific heterogeneity in second moments documented earlier: if managers estimate volatilities and correlations from broadly the same historical record, their forecasts should look alike. The reconciliation is that reliance on history is not uniform—managers differ in how much weight they place on the distant versus the recent past. We test this implication by comparing each manager's volatility and correlation forecasts with a per-manager exponentially weighted moving average (EWMA) benchmark. The forecasts cover eight asset classes—three equity and five fixed-income—for volatility, and their pairwise combinations for correlation.[12] For manager i, the benchmark is $\beta_i^* \cdot \mathrm{EWMA}(W_i^*, \lambda_i^*)$, where the lookback window W, decay λ, and scale β are selected separately for each manager (the manager-specific window W_i^* and decay λ_i^* are the locus of cross-manager heterogeneity—they govern how heavily each manager weights recent versus distant observations) by minimizing, across all of the manager's asset class–date (or pair–date) observations, the mean squared difference between the subjective forecast and the scaled EWMA, with each squared error divided by a class-specific level ℓ_c and the intercept fixed at zero. The normalization by ℓ_c places forecast errors on a comparable footing across classes of very different magnitude, so the benchmark is fit to the manager's forecasts across all classes rather than only where the numbers are largest. The window and decay are searched over a grid—$W \in \{5, \ldots, 20\}$ years and $\lambda \in [0.80, 1.00]$—together with the scale $\beta \in \{0.1, \ldots, 1.5\}$ in steps of 0.02, because the window and decay enter the EWMA weights nonlinearly and cannot be recovered by linear regression.

Figure 12 plots each manager's subjective forecast against this benchmark—volatility in Panel (a)

[12]The asset classes are U.S. Large Cap Equity, Developed ex-US Equity, and Emerging Market Equity; and U.S. Aggregate Bonds, U.S. Government Bonds, U.S. Investment-Grade Corporate Bonds, U.S. High-Yield Bonds, and U.S. TIPS. The eight classes yield 28 distinct pairs for the correlation analysis.

and pairwise correlation in Panel (b)—with both axes demeaned within asset class (Panel a) and within asset pair (Panel b) to remove level differences across series. For each disclosure group and the pooled sample, we report the OLS slope $\hat{\beta}_i$ from regressing the demeaned subjective forecast on the demeaned benchmark, together with the R^2; a slope near one and a concentration of observations along the 45° line indicate that the subjective forecast moves nearly one-for-one with the mechanical rule.

Managers' subjective forecasts track this benchmark closely for both volatility and correlation. The relationship is tight in both panels: the pooled R^2 is 0.58 for volatility and 0.59 for correlation, and the observations concentrate along the 45° line. The pooled slopes are 0.84 and 0.95, respectively. The correspondence is consistent across stated methodologies: managers reporting purely historical estimates, historical estimates with adjustment, and model-based calibrations all display slopes in a narrow range, roughly 0.81 to 0.97, with comparable fit. Thus, the close agreement with a simple time-series rule does not depend on how managers describe their approach to setting second moments. The slopes lie slightly below one, as expected: because the benchmark is an estimated and therefore noisy regressor, classical measurement error attenuates the estimated slope toward zero, so these values understate the strength of the underlying comovement rather than reflecting any real dampening by managers.[13]

Taken together, the results indicate that a parsimonious exponentially weighted moving average captures the bulk of the systematic variation in managers' subjective second-moment forecasts, across both volatility and correlation and across disclosed methodologies. This resolves the tension raised above: the large asset-manager fixed effects in second moments largely reflect heterogeneous—but still backward-looking—weighting of a common historical record, encoded in manager-specific windows and decay rates.

[Figure 12 here]

Figures 13 and 14 turn from cross-sectional fit to the time series, and show a complementary pattern: not only do forecasts track a backward-looking rule, they respond little to current conditions. Figure 13 shows that subjective volatility forecasts for large U.S. equities remain highly stable over time, even as EWMA realized-volatility measures with different decay parameters exhibit considerable variability. The narrow range and stability of these forecasts further suggests that managers

[13]The Undisclosed group is too small to interpret, with five managers and 40 observations for volatility and two managers and 38 for correlation, and yields unstable estimates across the two panels; we therefore do not interpret these estimates.

anchor their volatility forecasts to long-term historical norms rather than responding strongly to current market conditions. Figure 14 compares the cross-manager distribution of subjective correlations between U.S. large-cap equity and six other asset classes with the corresponding realized correlations: U.S. Small-Cap Equity, Developed ex-U.S. 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-U.S. equity, while the distribution for emerging-market equity is lower and more dispersed, centered near 0.71. Subjective equity–bond correlations also track their realized historical counterparts closely.

[Figures 13 and 14 here]

To further highlight the contrast between return expectations and second-moment beliefs, Figure 15 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 15 here]

Taken together, the evidence points to a sharp contrast in belief formation: return expectations are revised more and draw on richer forward-looking assumptions and narratives, whereas volatility and correlation beliefs are revised less and remain more closely tied to historical estimates.

5 Conclusion

A published return expectation is the endpoint of a belief-formation framework, not a sufficient description of it. Table 1 makes this concrete. PGIM and Amundi both report a ten-year U.S. equity return expectation of 5.8%, yet Panel A reveals different beliefs about growth, income, inflation, and valuation, while Panel B reveals differences in author composition, causal-network complexity,

and narrative attention. The same number can therefore emerge from different economic views and different institutional processes. Our results show that this is systematic: managers may agree on the total return expectation while disagreeing about the forces that produce it, the assumptions used to construct it, and the way new information should change it.

CMAs allow us to follow this belief-formation process from the final number backward. The building-block decomposition supplies institutions with a common accounting language, but not with a common view of the economy. Once we open the total return expectation, apparent agreement often dissolves. In U.S. equities, valuation change makes the largest net contribution to dispersion in total return expectations, but its negative relationship with growth conceals substantial disagreement about long-run fundamentals. For U.S. equity, these distinctions also explain the countercyclicality of return expectations: favorable economic conditions raise growth expectations, but high valuations reduce future repricing expectations, and the latter effect dominates. They matter for behavior as well, because growth and valuation-change beliefs—but not dividend-yield beliefs—are associated with equity allocations. Across most non-U.S. equity markets, growth is the main divide. Fixed-income return expectations have a different anatomy, organized around income, roll-down, valuation change, and credit losses.

The next question is why managers assign different values to the same accounting components. The evidence points beyond a single model used with different parameter values. Institutions adopt different assumptions, emphasize different signals, and interpret the same information through different causal structures. Mean-reversion and historical-calibration approaches are associated with systematic differences relative to both peer consensus and statistical benchmarks. Models fitted to managers' forecasts place more weight on macro-financial conditions and less on valuation ratios than models forecasting realized returns. Forecast formation also has a social dimension: managers revise toward peer consensus, and their benchmark-relative deviations align with recent peer views. Team composition is also associated with how managers distribute return expectations across components and with the forecasting methods they disclose. These associations are descriptive rather than causal, but they reinforce the interpretation of CMA forecasts as institutional products.

The narratives reveal how these forecasting frameworks process information. Positive earnings news can be read as evidence of stronger future cash flows, but also as a sign that valuations have already increased and that less future repricing remains. Managers therefore need not translate the same news into the same revision. More complex and valuation-focused narratives are associated

with underreaction to favorable news, whereas dividend-yield and downturn-focused narratives are associated with overreaction. The recurrence of these narratives in shareholder letters indicates that they reflect persistent house views rather than standardized language written around forecast tables. Volatility and correlation beliefs provide a useful contrast: they remain more closely tied to historical estimates and are revised less.

For research using expectations data, the implication is that agreement in forecast levels is not equivalent to agreement in beliefs. A scalar forecast compresses component expectations, modeling choices, institutional inputs, and rules for responding to new information. Two institutions that agree today may therefore react differently to tomorrow's news and choose different portfolios. The published number is the endpoint of belief formation, not the belief itself.

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., Gonçalves, A. S., Loudis, J. A. & 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., Décaire, P. H. & Guenzel, M. (2024), 'Mental models and financial forecasts', *Working Paper* .

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.

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.

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

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. & 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.

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.

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.

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

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

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. (2026), 'Institutions' return expectations across assets and time', *Journal of Financial Economics* **175**, 104188.

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.

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.

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.* .

Harvey, D., Leybourne, S. & Newbold, P. (1997), 'Testing the equality of prediction mean squared errors', *International Journal of Forecasting* **13**(2), 281–291.

Hillert, A., Niessen-Ruenzi, A. & Ruenzi, S. (2025), 'Mutual fund shareholder letters: Flows, performance, and managerial behavior', *Management Science* **71**(5), 4453–4473.

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. (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'. Working Paper.

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: Sources of Variation in Subjective Return Expectations, Volatility Forecasts, and Correlation Forecasts. This figure reports variance decompositions based on partial R^2 from saturated fixed-effects models with asset-manager and year fixed effects. Panels A and B report decompositions for asset-class-specific return expectations and volatility forecasts, respectively. The equity asset classes are U.S. (US Eq), Canada, U.K., Europe, Japan, China, Developed ex-U.S. (Dev ex-US Eq), and Emerging Markets (EM Eq). The fixed-income asset classes are U.S. Government Bonds (US Govt), U.S. TIPS, U.S. Aggregate Bonds (US Agg), U.S. Investment Grade Corporate Bonds (US IG Corp), U.S. High Yield Bonds (US HY), U.S. Municipal Bonds (US Muni), Developed ex-US Government Bonds (Dev ex-US Govt), Global ex-US Bonds (Global ex-US Bd), and Emerging Markets Debt (EM Debt). Each stacked bar shows the percentage of total variation explained by asset-manager fixed effects, year fixed effects, and the residual component. Panel C reports decompositions for pairwise correlation forecasts among broad equity and fixed-income asset classes. For each asset-class pair, the lower-left triangle of the heatmap reports the percentage of variation explained by asset-manager fixed effects (in blue), while the upper-right triangle reports the corresponding percentage explained by year fixed effects (in yellow). Diagonal entries are omitted. All decompositions are estimated separately for each forecast object. Sample period: 2008–2026.

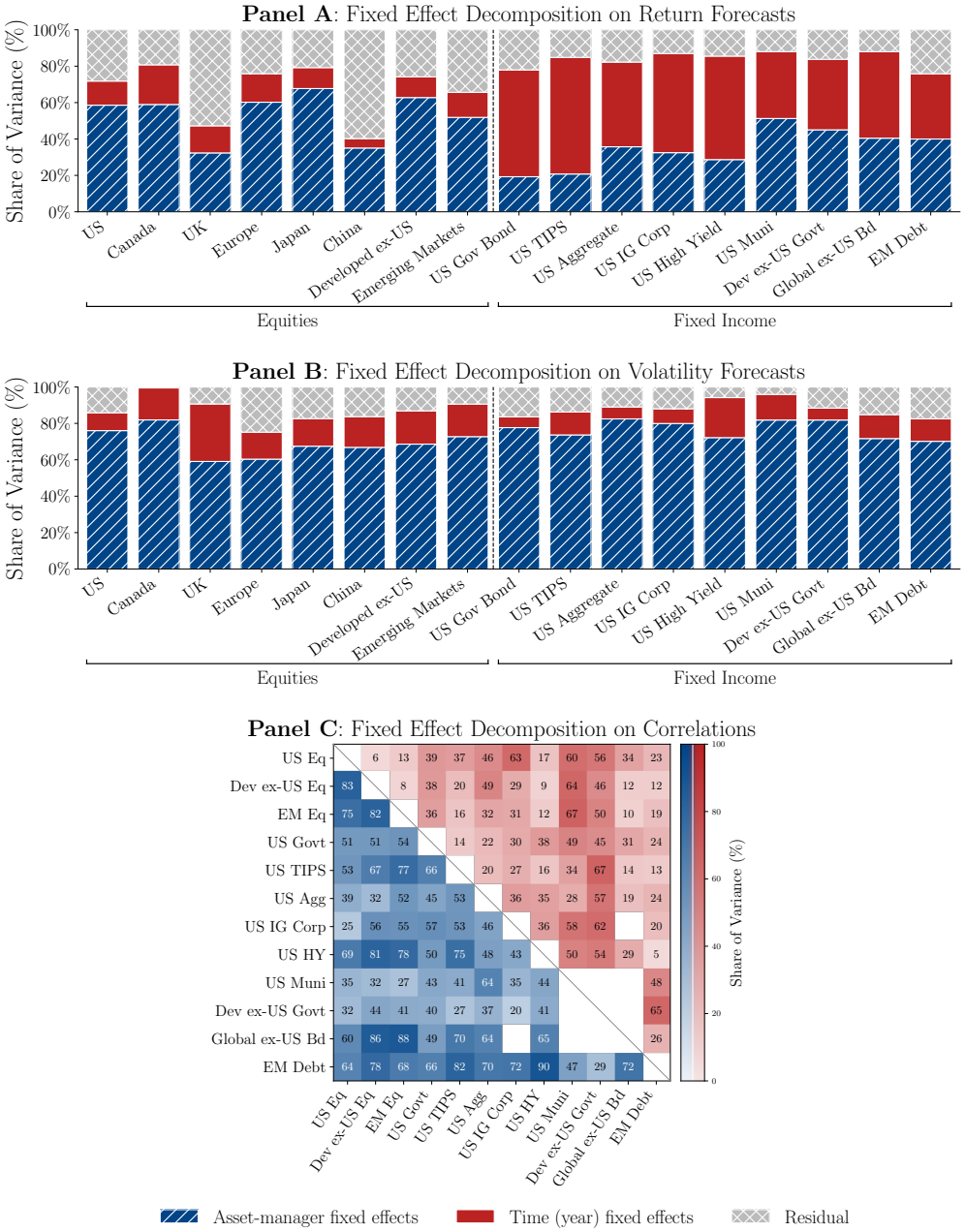


Figure 2: Within-Year Decomposition of Disagreement in Equity Return Expectations. This figure decomposes the cross-sectional dispersion in total equity return expectations into three reported components—valuation change, nominal growth, and dividend yield—across eight equity markets: the United States, developed markets excluding the United States, Europe, the United Kingdom, Australia, Japan, China, and emerging markets. All variables are residualized using year fixed effects, equivalently demeaned within survey year, so the plotted variation reflects disagreement across forecasters rather than common time-series movements. Blue bars report the covariance-share decomposition, which measures each block's net contribution to dispersion in the total return expectation. Red bars report the Type II ANOVA contribution, which measures the contribution specific to each block after holding the other blocks fixed. Gray numbers in the background report pairwise correlations among the residualized building-block forecasts. Sample period: 2008–2026.

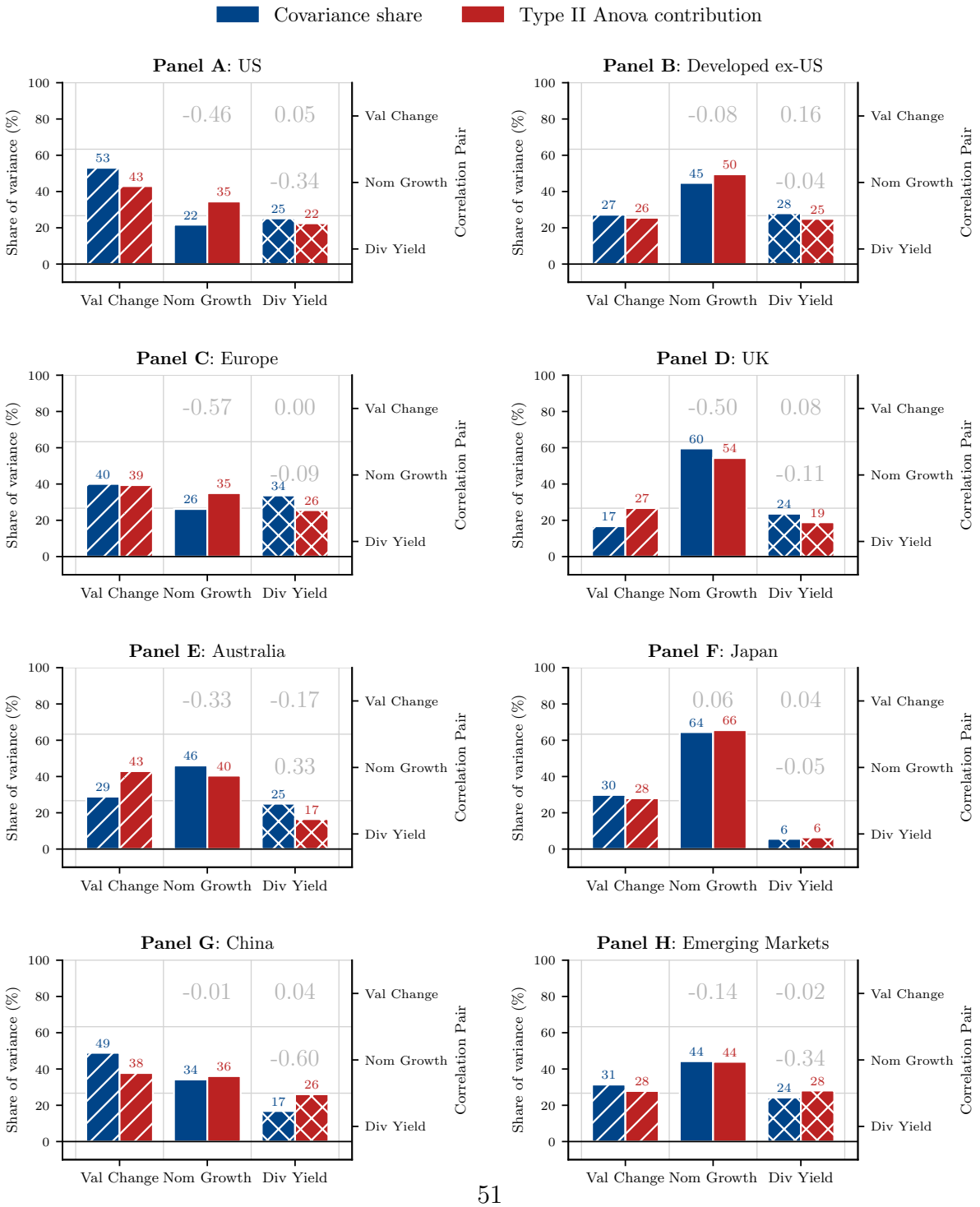


Figure 3: Within-Manager Decomposition in Equity Return Expectations. This figure decomposes the within-manager variation in total equity return expectations into three reported building blocks—valuation change, nominal growth, and dividend yield—across eight equity markets: the United States, developed markets excluding the United States, Europe, the United Kingdom, Australia, Japan, China, and emerging markets. All variables are residualized with manager (provider) fixed effects, equivalently demeaned within manager, so the plotted variation reflects how a manager's expectations move over time around their own average, rather than differences across managers. Blue bars report the covariance-share decomposition, which measures each block's net contribution to the variance of the total-return forecast. Red bars report the Type II ANOVA decomposition, which measures the contribution specific to each block after holding the other blocks fixed. Gray numbers in the background report pairwise (within-manager) correlations among the residualized building-block forecasts. Sample period: 2008–2026.

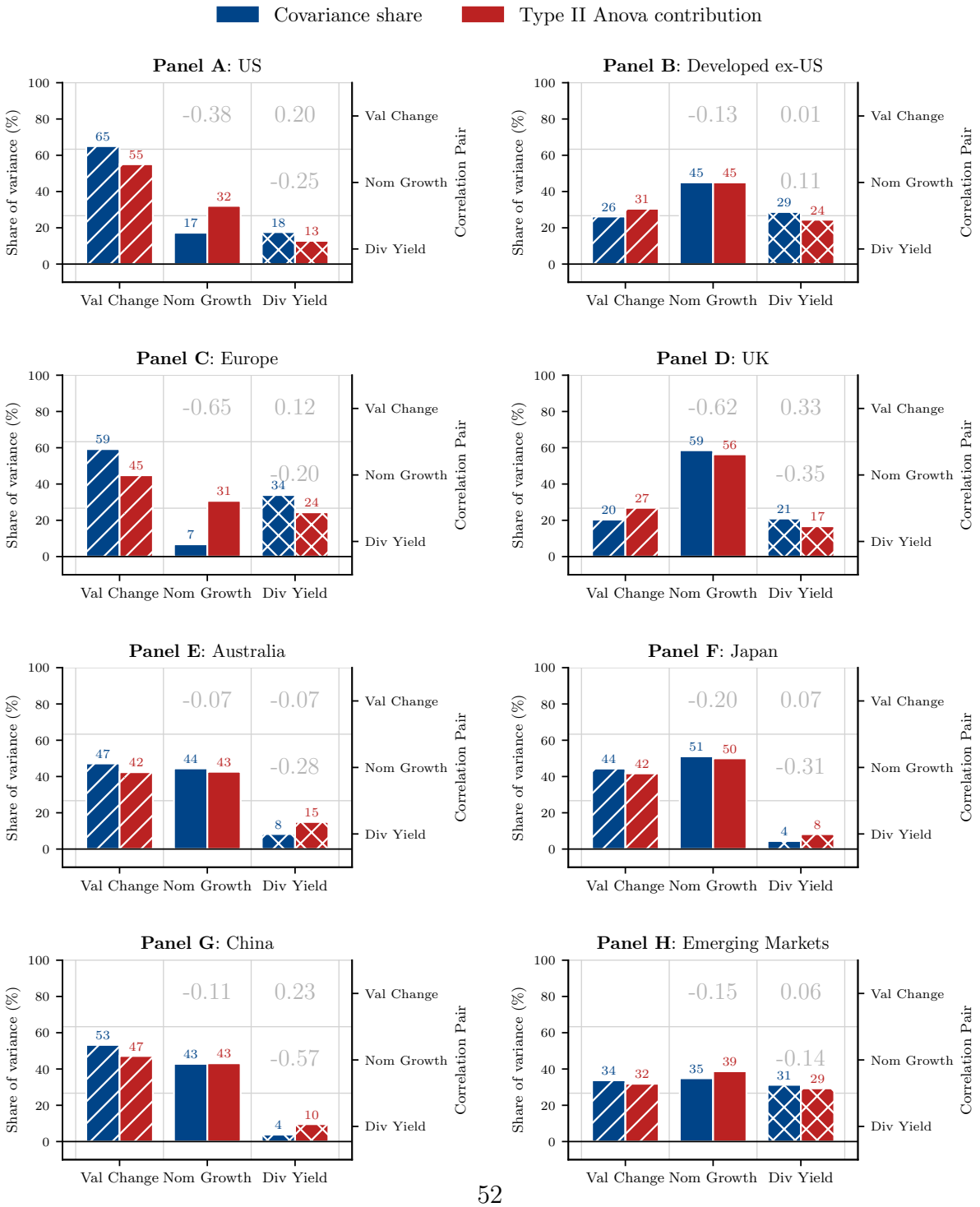
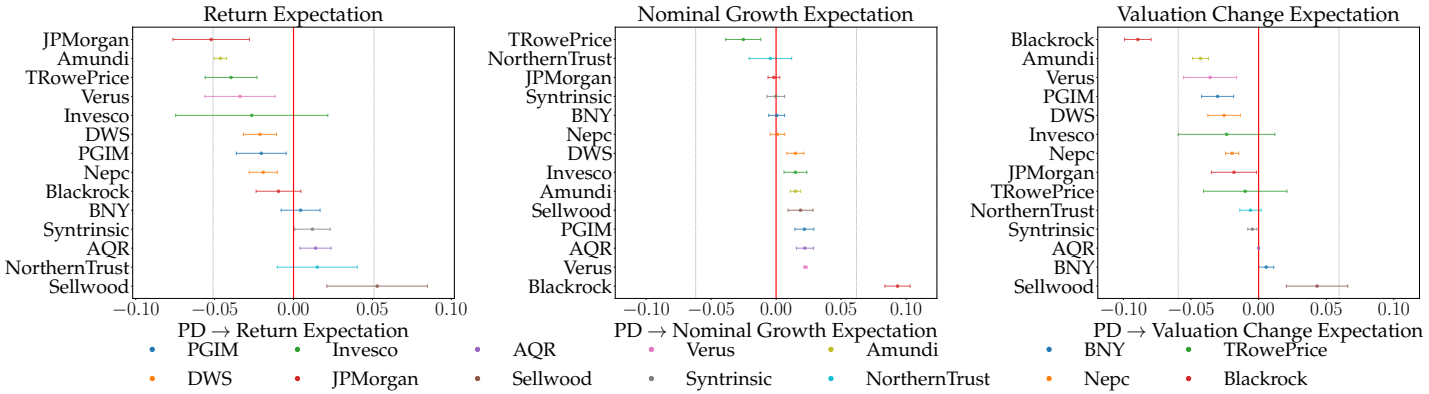


Figure 4: Campbell-Shiller Decomposition: Asset-Manager-Level Loadings on Valuation Ratios. This figure shows regression coefficients from asset-manager-level specifications in which CMA forecast components are regressed on valuation ratios. In each panel, $F^b_{i,t,h}$ denotes asset manager i's forecast for component b at CMA reference date t and horizon h, where $b \in \{ER, NG, VC\}$. The three columns report, respectively, the total 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.

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

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



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

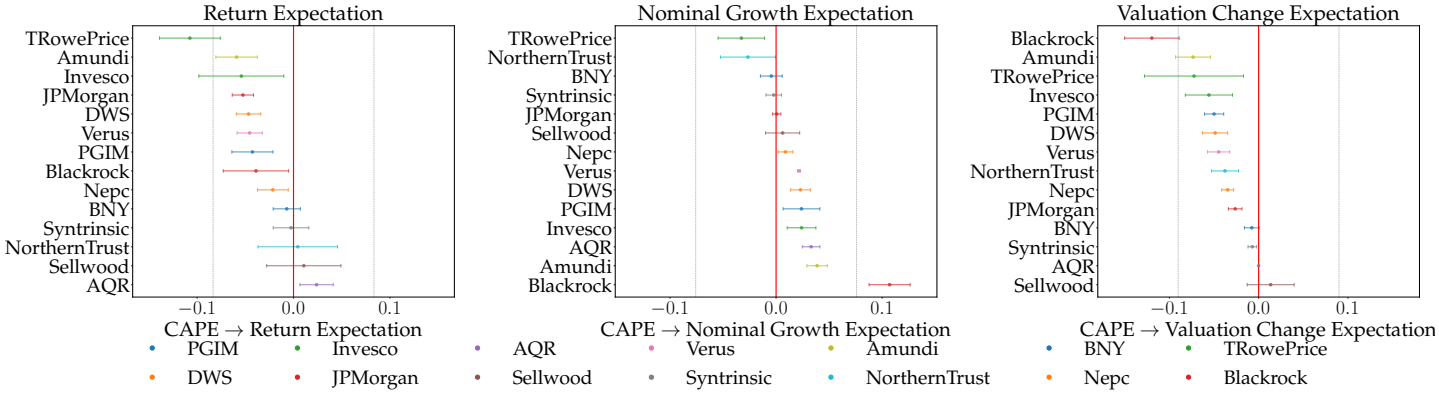


Figure 5: Cyclicality on CAPE: Share of Positive Slopes (Equity and Fixed Income). This figure reports, for each asset class and building block, the share of asset managers whose per-manager slope of that block on the log cyclically-adjusted price–earnings ratio (CAPE) is positive. Rows above the horizontal rule are equity, ordered U.S., Canada, U.K., Europe, Japan, China, Developed ex-U.S. and Emerging Markets; rows below it are fixed income, ordered U.S. Government Bonds, U.S. Investment Grade Bonds, U.S. High Yield Bonds, and U.S. Aggregate Bonds. Each equity market is regressed on the CAPE of its own reference index, while U.S. equity and every fixed-income class are regressed on the U.S. CAPE. Columns are the total return expectation followed by its building blocks: Valuation Change, Growth and Dividend Yield for equity, and income, roll-down and credit loss for fixed income, with valuation change common to both. Within each asset class the return expectation is set equal to the sum of the blocks reported in the remaining columns, so a block a manager does not report is read as that manager assuming it is zero, and the first column regresses the total return expectation on CAPE while the remaining columns regress its components. Color tracks the share on a diverging scale (deep red at 0%, white at 50%, deep green at 100%): green where most managers raise the block as valuations rise, red where most lower it. Equity and fixed income are decomposed into different blocks, so the grid has a block structure rather than a full grid; columns that do not apply to a class are shaded grey, while blocks common to both groups populate every row. A cell is drawn only where at least three managers are available. Sample period: 2008–2026.

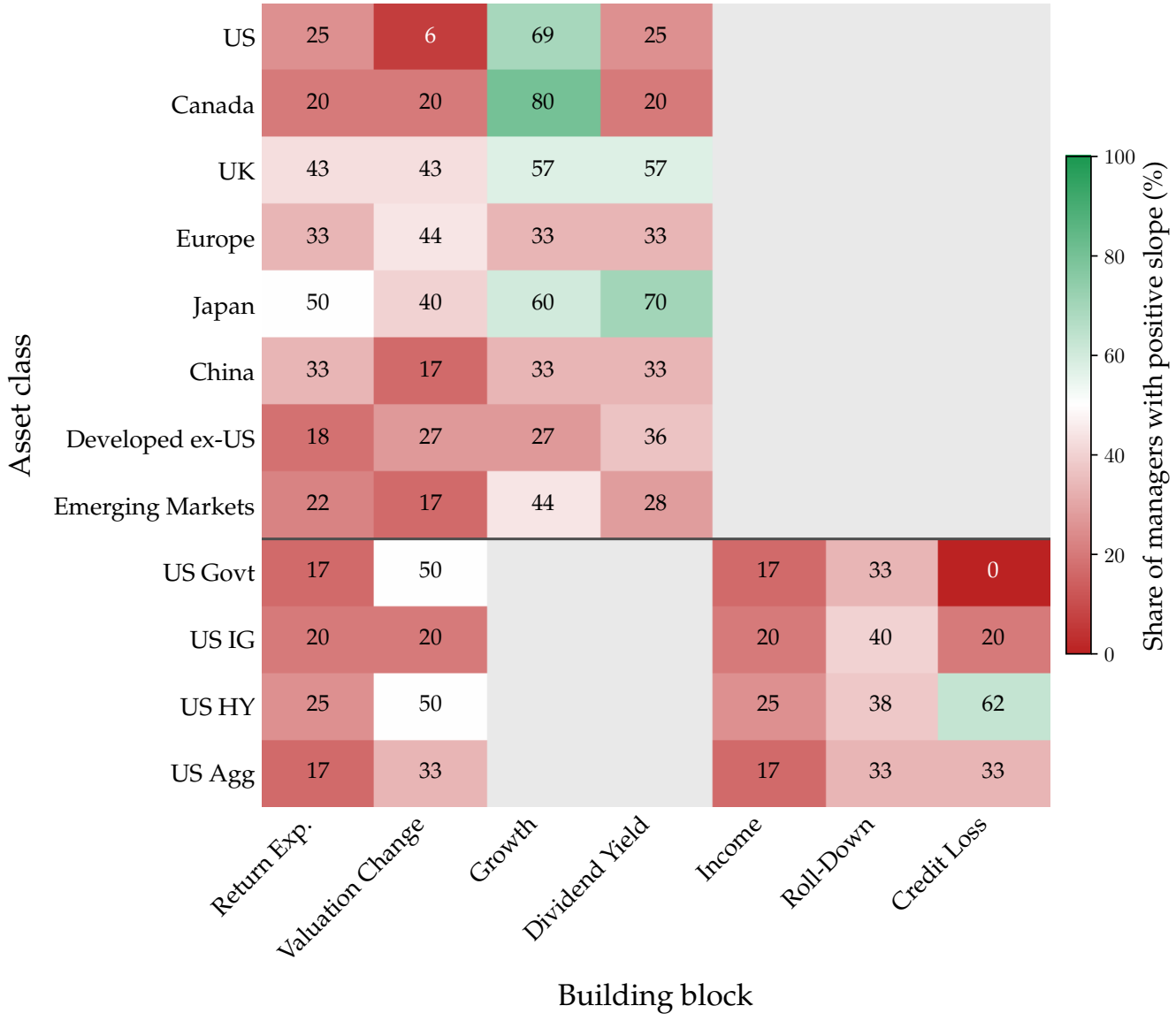
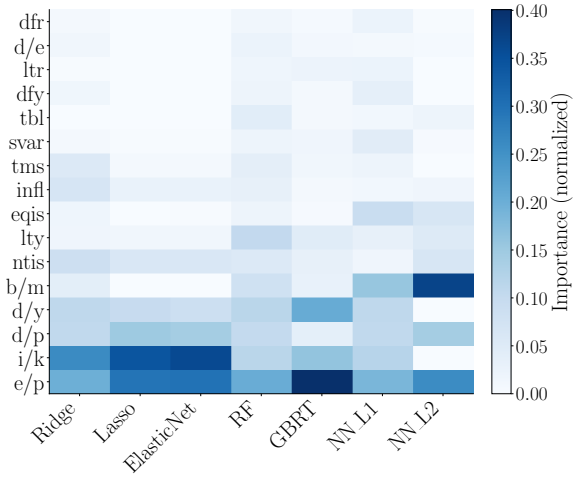


Figure 6: Model feature importance and the manager-versus-machine comparison. Panel (a) reports normalized permutation importance for the 16 Welch & Goyal (2008) predictors in models trained to predict annualized realized U.S. equity returns, $R_{t,t+h}$, the target used to construct the objective benchmark $G_{t,h}$. 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' total CMA return 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 A3. 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.

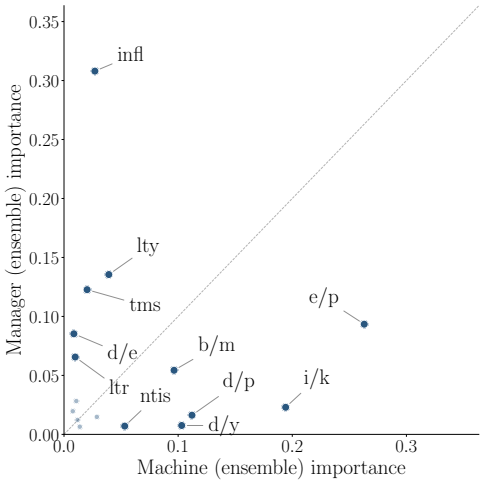


Figure 7: Prevalence of Preset Forecasting Techniques across Reports. This figure reports the share of CMA reports in which each of six preset forecasting techniques is utilized: regressions, simulations, historical averages, mean reversion, trend hypotheses, and others. The indicators are non-exclusive: a single report can declare more than one technique. Sample: 2010–2025.

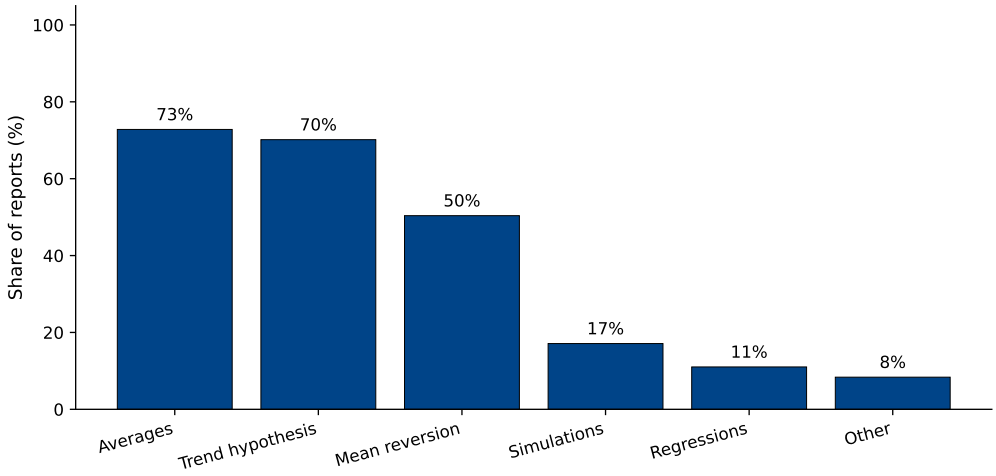


Figure 8: Author Education, Forecast Decomposition, and Preset Forecasting Techniques. Each cell reports the Pearson correlation between the number of authors in the report holding the indicated highest degree and a characteristic of the manager's U.S.-equity forecast. Authors are classified from LinkedIn profiles into Bachelor, Master / MBA, and PhD. Panel (a) uses, as the characteristic, the manager's deviation from the same-year leave-one-out cross-sectional consensus, $Dev_{i,t,b} = F_{i,t,b} - \overline{F}_{-i,t,b}$, on each equity building block (equity return, real growth, inflation, nominal growth, dividend yield, and valuation change), computed at the report level. Reports in which a given building block is not reported are excluded from both the consensus average and the correlation for that cell. Panel (b) uses the manager's preset forecasting techniques for equity returns, each coded as an indicator for whether the report reports using regressions, simulations, historical averages, mean reversion, a trend hypothesis, or other techniques. Observations are collapsed to one per manager, averaging author counts and technique indicators across the manager's reports. Positive (red) values indicate that more authors of a given degree go with a higher value of the characteristic. Sample period: 2010 to 2024.

(a) Deviations from consensus on the equity building blocks (report level)



(b) Preset forecasting techniques for equity returns (manager level)

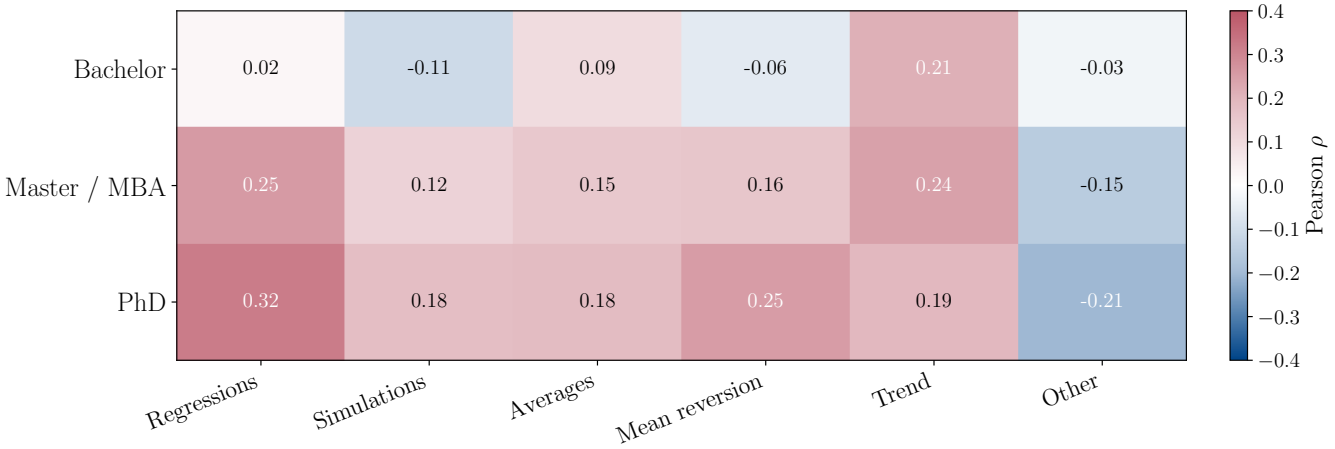


Figure 9: The Free-Text Modeling Assumption Embedding Space for Equities. This figure shows a two-dimensional t-SNE projection of the 4,145 distinct modeling-technique phrases collected from the equity sections of 589 CMA reports. Each point is one technique phrase, positioned by the similarity of its 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.

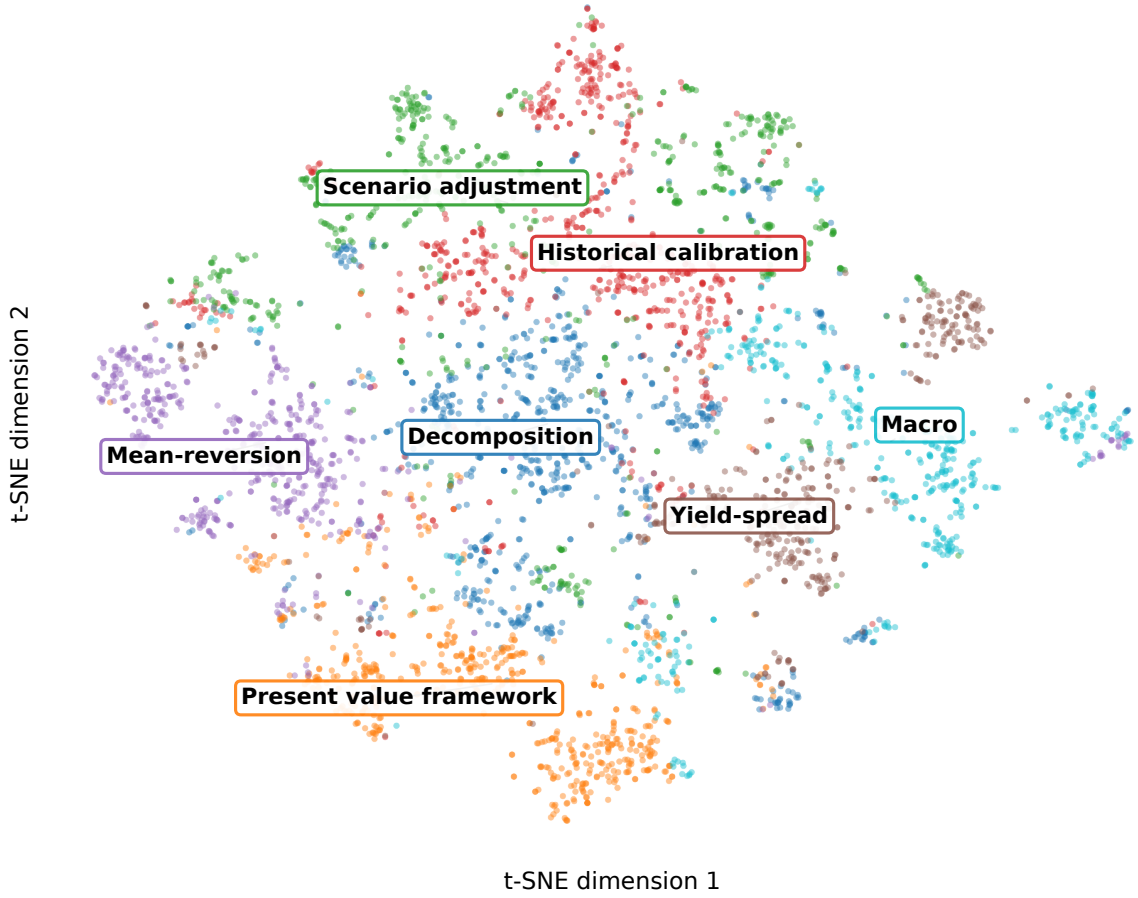


Figure 10: Asset Manager Causal Network of U.S. Equity Return Expectations. This figure shows the directed network of causal relationships extracted from DWS Capital Market Assumptions. Nodes represent topics identified in the text, with size reflecting centrality (number of connections). Edges indicate 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.

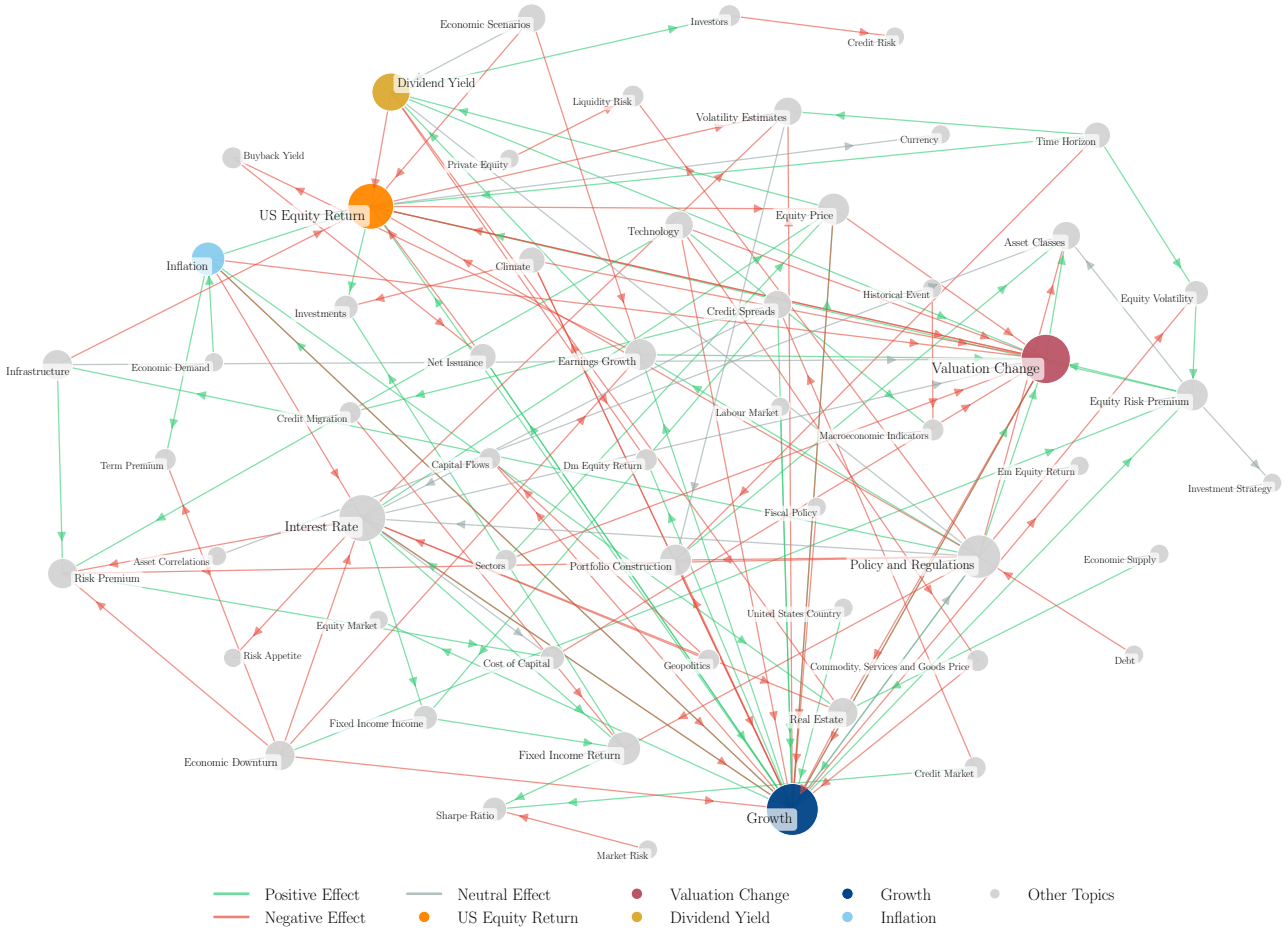


Figure 11: Cumulative Distribution of Letter–CMA Cosine Similarities. This figure shows the empirical cumulative distribution function of cosine similarities between the outlook section of each asset manager's N-CSR letter and the paragraphs of a contemporaneous capital-markets-assumption (CMA) report. 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.

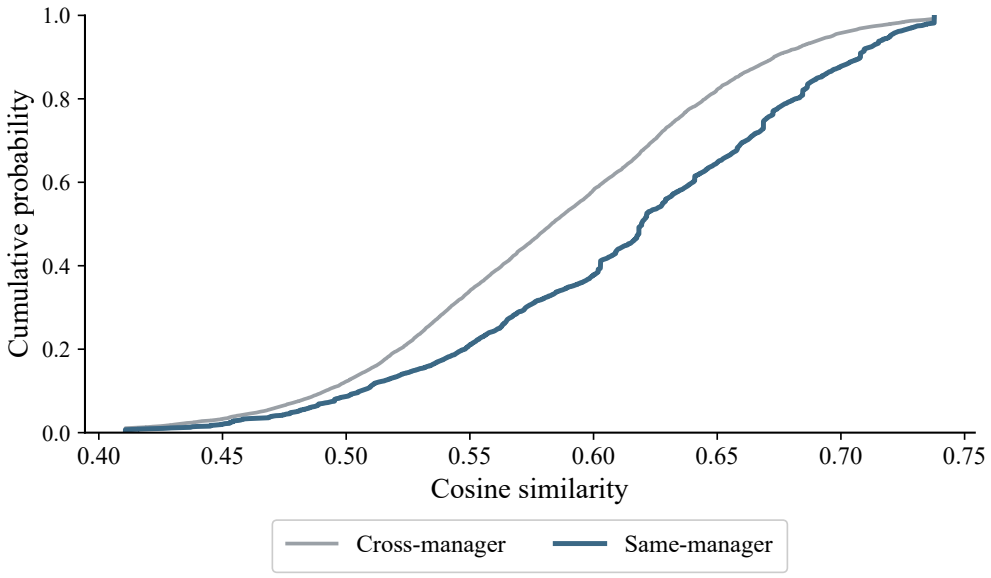
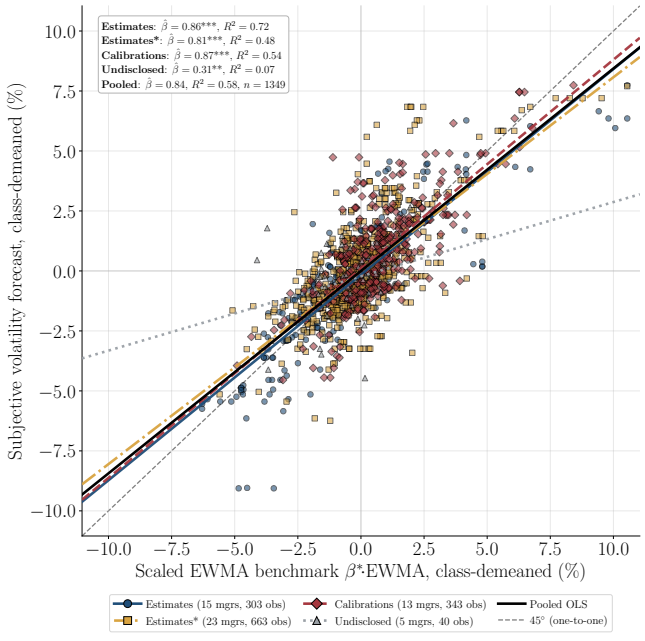


Figure 12: Second Moment Assumptions and Past Realizations. Each panel plots asset managers' subjective forecasts (vertical axis) against a per-manager exponentially weighted moving average (EWMA) benchmark (horizontal axis): cross-asset volatility in Panel A (in %) and pairwise correlation in Panel B (unitless). For each manager the benchmark is $\beta^* \cdot \text{EWMA}(W^*, \lambda^*)$, with the lookback window W, decay λ, and scale β selected jointly on a grid—$W \in \{5, \ldots, 20\}$ years, λ from 0.80 to 1.00 in steps of 0.01, and $\beta \in \{0.1, \ldots, 1.5\}$ in steps of 0.02—by minimizing the level-normalized pooled mean-square forecast error against realized outcomes (each squared error scaled by the class-specific level ℓ_c), with the intercept fixed at zero ($\alpha \equiv 0$). To remove level differences, both axes are demeaned within asset class (Panel A) and within asset pair (Panel B). Markers denote how each manager discloses its second-moment assumptions—Estimates (purely historical), Estimates* (historical with adjustment), Calibrations (model-based), and Undisclosed (method not stated)—with the number of managers and forecast observations shown in the legend; the groups partition the sample. The inset reports the OLS slope $\hat{\beta}$ and R^2 from regressing the subjective forecast on the scaled benchmark, within each group and pooled. Here $\hat{\beta}$ (the fitted subjective-on-benchmark slope) is distinct from β^* (the scale used to *build* the benchmark): $\hat{\beta} \approx 1$ indicates that a manager's subjective forecast moves one-for-one with the mechanical EWMA. The solid line is the pooled OLS fit and the dashed diagonal the 45° (one-to-one) line. ***, **, and * denote significance at 1%, 5%, and 10%. Note, in Panel B the Undisclosed group contains only 2 managers (38 observations), so its near-unit slope and high R^2 should be read with caution. The asset classes covered are: U.S. Large Cap Equity, Developed ex-US Equity, Emerging Market Equity, U.S. Aggregate Bonds, U.S. Government Bonds, U.S. Investment-Grade Corporate Bonds, U.S. High-Yield Bonds, and U.S. TIPS. Panel a reports volatility forecasts across asset classes, while Panel b reports correlation forecasts.

(a) Volatility Forecasts.



(b) Correlation Forecasts.

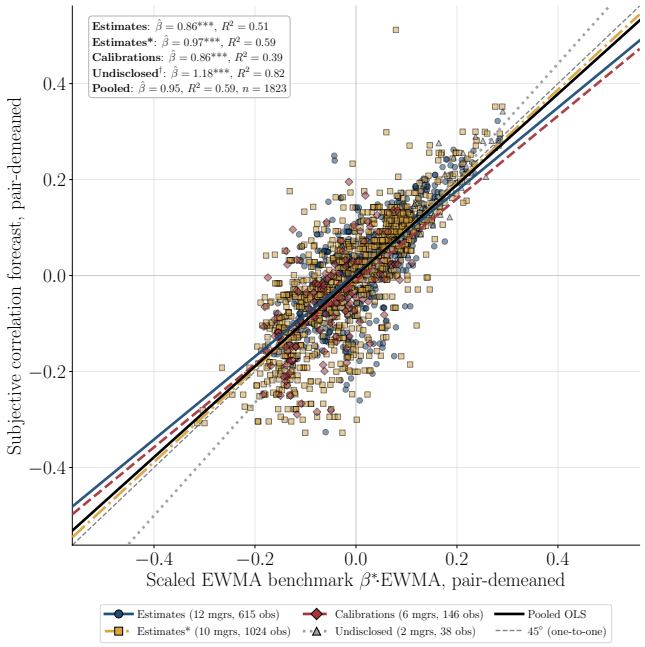


Figure 13: Subjective Volatility Forecasts versus Realized Volatility Measures for U.S. Large-Cap Equity. This figure compares asset managers' subjective volatility forecasts (blue box plots) with exponentially-weighted moving averages (EWMA) of trailing realized volatility for U.S. large-cap equity (S&P500), 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. Sample period: 2012–2026.

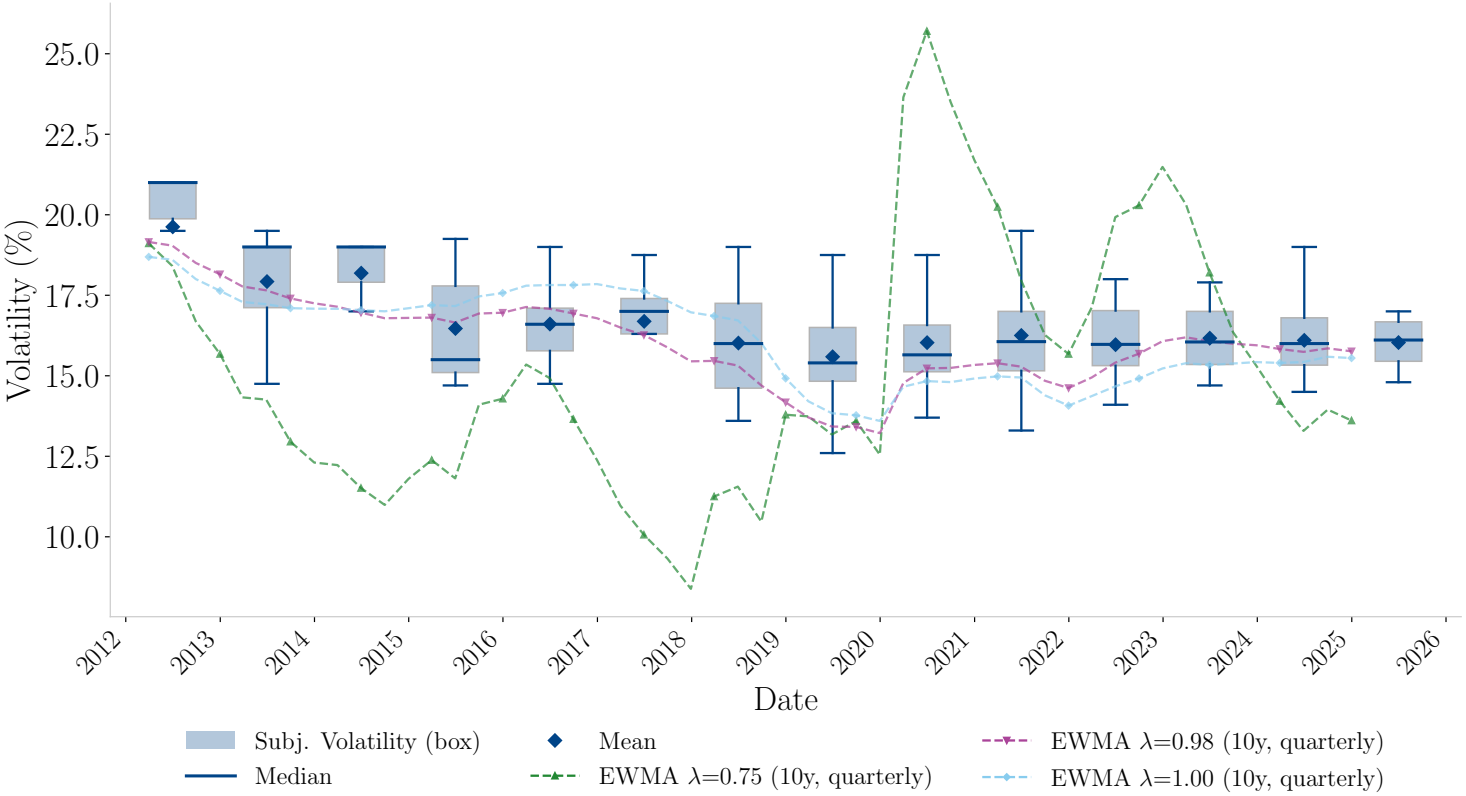


Figure 14: Subjective and Realized Correlations with U.S. Large-Cap Equity. For each counterparty class, the boxes show the cross-manager distribution of the subjective correlation between U.S. large-cap equity and that class, one box per vintage year. The dashed lines show the realized correlation between 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-U.S.), MSCI EM (emerging markets), the Bloomberg U.S. Aggregate (aggregate bonds), U.S. Treasury Intermediate (government bonds), and U.S. Corporate High Yield (high-yield corporates). Sample period: 2016–2025.

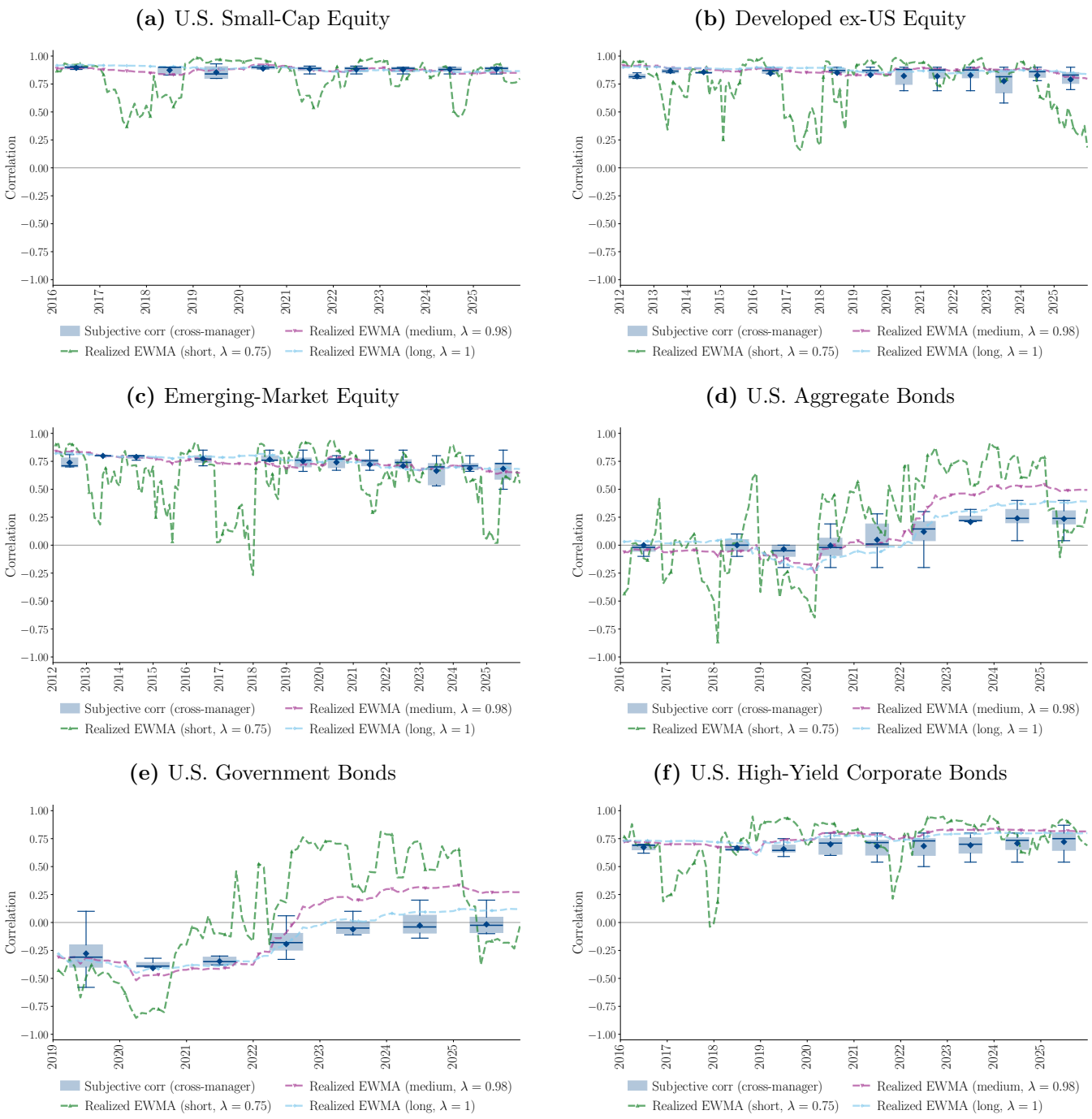


Figure 15: Year-over-Year Revision of U.S.-Equity Return, Volatility, and Correlation Beliefs. For each asset manager, the horizontal bar stacks the typical absolute year-over-year revision of three U.S. large-cap-equity beliefs: the return expectation, the return volatility, and the average correlation 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.

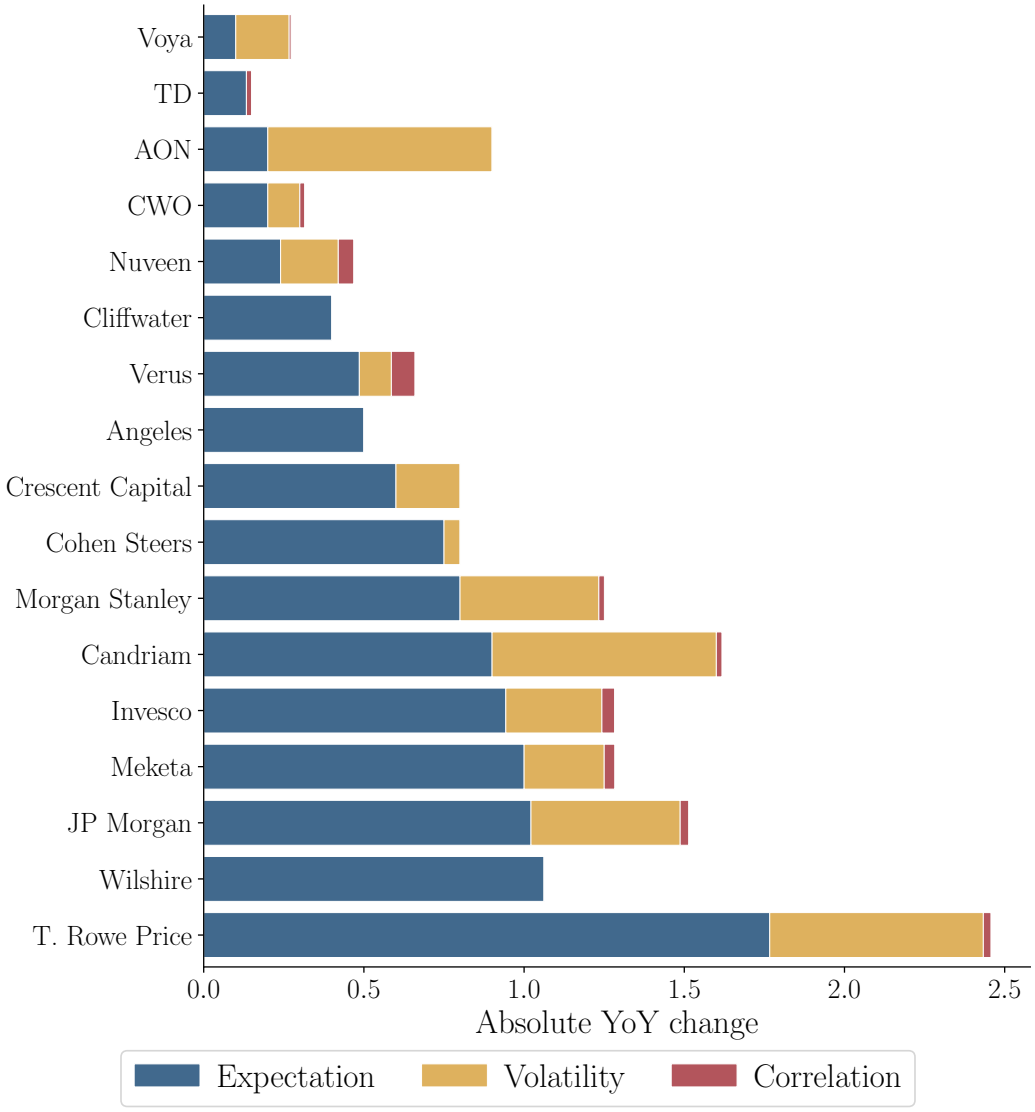


Table 1: Same Return Forecast, Different Component Beliefs. This table compares the annualized ten-year U.S. equity return forecasts reported by PGIM and Amundi in their Capital Market Assumptions reports in the first quarter of 2021. Panel A decomposes each institution's total return forecast into four approximately additive building blocks: Valuation Change, Real Growth, Income, and Inflation. Although both institutions forecast a total return of 5.8%, they reach this figure through markedly different assumptions about the individual components. The reported components are harmonized into these four common categories, and their sum yields the total return forecast up to rounding. Panel B reports characteristics of the same two reports, with all variables measured at the report level. For each report, it shows the share of degree-holding authors who hold a PhD. It also reports a complexity t-statistic for the null hypothesis $H_0 : C_{\text{PGIM}} = C_{\text{Amundi}}$, where complexity is measured by the average path length of the report's causal network and the difference is standardized against the report-level cross-section. The % Attention columns report the raw share of causal-network link endpoints associated with Valuation Change and Real Growth, with all remaining attention grouped under Other. The p-value for the complexity test is reported in parentheses. The construction of the causal-network measures is described in Section 3.2.

Panel A: Return Forecast Decomposition

Manager	Total Return	Valuation Change	Real Growth	Income	Inflation
PGIM	5.8%	−2.1%	2.2%	3.5%	2.3%
Amundi	5.8%	−3.1%	4.6%	2.3%	2.0%

Panel B: Manager Framework's Characteristics

	PhD	Complexity difference	% Attention (manager level)		
Manager	in Team	$H_0 : C_{\text{PGIM}} = C_{\text{Amundi}}$	Valuation Change	Real Growth	Other
PGIM	66.7%	-1.68* (0.09)	40.0%	20.0%	40.0%
Amundi	21.4%		4.7%	13.3%	82.0%

Table 2: Forecast Change Dynamics and Convergence Toward Consensus. This table reports panel regression estimates of forecast changes by asset managers for each asset class with an explicit return decomposition (Equities: U.S., U.K., Europe, Japan, Developed ex-U.S., Emerging Markets; Fixed Income: U.S. Government Bonds, U.S. Investment Grade Bonds, U.S. High Yield Bonds, U.S. Aggregate Bonds). For each forecast variable b, where ER denotes the total return expectation and the remaining variables denote the reported CMA building-block components (Valuation Change, Growth, and Dividend Yield for equity, and income, roll-down, credit loss, and valuation change for fixed income), 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)–(10) report the six equity markets and four U.S. fixed-income sectors. The main regressor is the manager's lagged deviation from the leave-one-out consensus, $Dev^b_{i,t-1} = F^b_{i,t-1} - \overline{F}^b_{-i,t-1}$, where $\overline{F}^b_{-i,t-1}$ is the leave-one-out peer consensus, computed as the average value of component b among all other asset managers in year $t-1$. The regressions also control for the contemporaneous change in the leave-one-out consensus, $\Delta\overline{F}^b_{-i,t} = \overline{F}^b_{-i,t} - \overline{F}^b_{-i,t-1}$. Panels A and B report β^b and δ^b, and Panel C the within-R^2. 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: 2008–2026.

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

	Equity						Fixed Income			
	(1) US	(2) UK	(3) Europe	(4) Japan	(5) Dev ex-US	(6) EM	(7) US Govt	(8) US IG	(9) US HY	(10) US Agg
Panel A: Convergence, $Dev^b_{i,t-1}$										
Return Exp.	-0.725***	-0.569***	-0.627***	-1.176***	-1.346***	-0.638***	-1.062***	-0.895***	-1.209***	-1.163***
	(0.171)	(0.120)	(0.109)	(0.164)	(0.273)	(0.128)	(0.131)	(0.308)	(0.099)	(0.112)
Valuation Change	-0.945***	-0.531***	-0.814***	-0.923***	-0.877***	-0.847***	-0.684*	-0.020	-1.070***	-0.919***
	(0.181)	(0.105)	(0.175)	(0.211)	(0.037)	(0.125)	(0.355)	(0.134)	(0.200)	(0.143)
Growth	-0.709***	-0.713***	-0.799***	-1.200***	-1.270*	-0.482***				
	(0.089)	(0.124)	(0.182)	(0.128)	(0.633)	(0.161)				
Dividend Yield	-0.603***	-0.334***	-1.091***	-0.976***	-0.276	-0.316**				
	(0.159)	(0.092)	(0.185)	(0.282)	(0.170)	(0.153)				
Income							-1.082***	-0.901***	-0.948***	-1.212***
							(0.132)	(0.187)	(0.106)	(0.086)
Roll-Down							-0.837***	-0.681***	-0.615***	-1.358***
							(0.162)	(0.114)	(0.064)	(0.158)
Credit Loss								-0.189	-0.368***	-0.467
								(0.180)	(0.125)	(0.390)
Panel B: Herding, $\Delta\bar{F}^b_{-i,t}$										
Return Exp.	0.911***	0.205	0.989***	0.854***	0.982**	0.849***	0.847***	0.898***	0.964***	1.170***
	(0.128)	(0.238)	(0.111)	(0.185)	(0.384)	(0.215)	(0.088)	(0.181)	(0.103)	(0.046)
Valuation Change	0.921***	0.634***	1.021***	0.876***	0.474***	0.898***	0.464	0.170	0.587***	0.545***
	(0.132)	(0.166)	(0.142)	(0.210)	(0.033)	(0.182)	(0.338)	(0.144)	(0.129)	(0.161)
Growth	0.514***	0.372*	0.435	0.404	0.805	0.207				
	(0.178)	(0.188)	(0.402)	(0.344)	(0.469)	(0.279)				
Dividend Yield	0.737***	0.238***	0.285	0.387**	0.130	0.531***				
	(0.179)	(0.070)	(0.167)	(0.163)	(0.197)	(0.149)				
Income							1.117***	0.885***	1.008***	1.199***
							(0.085)	(0.099)	(0.077)	(0.075)
Roll-Down							0.326	0.254**	0.393*	0.587***
							(0.203)	(0.117)	(0.213)	(0.096)
Credit Loss								0.168	0.206**	0.236
								(0.159)	(0.079)	(0.204)
Panel C: R^2 **(within)**										
Return Exp.	0.464	0.286	0.485	0.538	0.591	0.441	0.840	0.582	0.703	0.897
Valuation Change	0.553	0.305	0.496	0.381	0.515	0.422	0.449	0.152	0.716	0.816
Growth	0.383	0.258	0.311	0.627	0.512	0.181				
Dividend Yield	0.412	0.144	0.607	0.474	0.069	0.260				
Income							0.787	0.650	0.744	0.858
Roll-Down							0.429	0.351	0.316	0.844
Credit Loss								0.217	0.195	0.266
Asset Manager FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	139	40	37	46	26	71	27	36	44	24

Table 3: CMA Forecast Anchoring: Total Returns and Building Blocks. This table reports panel regression estimates testing whether individual CMA forecast deviations from the machine-learning benchmark (described in Section 2.3) are related to the distance between a salient consensus anchor and the same benchmark. The dependent variable is $G^b_{t,h} - F^b_{i,t,h}$, where $F^b_{i,t,h}$ is asset manager i's forecast of building block b—return forecast (ER), valuation change (VC), real growth (g), income or dividend yield (DY), and inflation (π)—at report date t and horizon h, and $G^b_{t,h}$ is the out-of-sample machine-learning benchmark forecast for the same year and horizon. Panel A uses the total U.S. equity return ($b = ER$); Panel B estimates the same specification block by block. The regressor in column (1) is $G^b_{t,h} - \overline{F}^{b,365d}_{t-1,h}$, where $\overline{F}^{b,365d}_{t-1,h}$ is the rolling consensus forecast constructed from horizon-matched component forecasts issued during the previous 365 days. Column (2) replaces the contemporaneous benchmark in the consensus deviation with the lagged benchmark, $G^b_{t-1,h} - \overline{F}^{b,365d}_{t-1,h}$, while keeping the dependent variable $G^b_{t,h} - F^b_{i,t,h}$. 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. N is the same in both columns; the within R^2 is reported for column (1). *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero slope coefficient. Sample period: 2008–2026.

	(1)	(2)		
b	$G^b_{t,h} - \overline{F}^{b,365d}_{t-1,h}$	$G^b_{t-1,h} - \overline{F}^{b,365d}_{t-1,h}$	N	R^2 (within)
Panel A: Total equity return				
ER	0.745***	0.400***	450	0.232
	(0.090)	(0.082)		
Panel B: Building blocks				
VC	0.331**	0.178***	177	0.061
	(0.152)	(0.065)		
g	0.817***	0.312***	150	0.289
	(0.123)	(0.076)		
DY	−0.134*	−0.060	204	0.015
	(0.074)	(0.087)		
π	0.958***	0.720***	182	0.441
	(0.083)	(0.066)		
Asset Manager FE	Y	Y		

Table 4: Portfolio Allocations and Return Expectations. This table reports regressions of U.S. equity allocations on CMA return expectations 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 U.S. equity return expectations is decomposed into Valuation Change ($VC_{i(f),t}$), Nominal Growth ($NG_{i(f),t}$), and Dividend Yield ($DY_{i(f),t}$). 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)	
Valuation Change			0.237**
			(0.111)
Nominal Growth			0.458***
			(0.145)
Dividend Yield			-0.362
			(0.286)
Fund FE	Y	Y	Y
Year FE	Y	Y	Y
R^2	0.063	0.036	0.034
Observations	3251	25765	25765

Table 5: Preset 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 total U.S. equity return forecast and each forecast component, $b \in \{ER, VC, g, DY, \pi\}$, 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, VC, g, DY, and π, respectively. Manager-years for which no peer forecast is available are excluded. Regressions are estimated at the manager-year level. Preset technique indicators are extracted and constructed from CMA report text using LLM-based classification. For each $k \in \mathcal{K} = \{$Regression, Simulation, Historical Average, Mean Reversion, Trend Hypothesis, Other$\}$, $\mathbf{1}\{\text{technique}_{k,i,t}\}$ 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.

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

	(1)	(2)	(3)	(4)	(5)
k	ER	VC	g	DY	π
Regression	-0.002	-0.003	0.007*	-0.004*	0.000
	(0.003)	(0.003)	(0.004)	(0.003)	(0.001)
Simulation	0.001	-0.003	0.004	0.002	0.000
	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)
Historical Average	0.003	0.002	0.001	0.003*	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)
Mean Reversion	-0.004**	-0.005***	-0.001	0.003	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Trend Hypothesis	0.001	-0.005**	-0.001	0.004	-0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)
Other	0.001	0.000	-0.003	0.008***	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)
Year FE	Y	Y	Y	Y	Y
R^2	0.041	0.112	0.108	0.120	0.041
Observations	343	158	168	188	168

Table 6: The Seven Free-Text Modeling Assumption Clusters for Equities. This table defines the seven clusters of modeling assumptions 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 7: Free-Text Modeling Assumption Clusters and Deviations from the Leave-One-Out Peer Consensus, by Forecast Component. This table shows how disclosed modeling assumption clusters relate to managers' signed deviations from the leave-one-out yearly peer consensus, separately for the total equity return forecast and each building-block component. For each component$b \in \{ER, VC, g, DY, \pi\}$, 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, VC, g, DY, 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. Free-text modeling assumption cluster indicators are extracted and constructed from CMA report text using LLM-based classification. Let $\mathcal{C}$ = {Decomposition, Historical Calibration, Scenario Adjustment, Present Value Framework, Mean Reversion, Macro, Yield Spread} denote the set of modeling assumption clusters. 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)
k	ER	VC	g	DY	π
Decomposition	-0.400	-0.381*	+0.163	-0.051	+0.011
	(0.285)	(0.195)	(0.177)	(0.131)	(0.071)
Historical Calibration	+0.610**	+0.456**	-0.276*	+0.302**	+0.049
	(0.247)	(0.182)	(0.147)	(0.153)	(0.039)
Scenario Adjustment	-0.126	-0.068	+0.197	-0.217	+0.046
	(0.214)	(0.168)	(0.160)	(0.146)	(0.047)
Present Value Framework	+0.146	-0.171	+0.290*	+0.109	+0.029
	(0.248)	(0.194)	(0.165)	(0.210)	(0.053)
Mean Reversion	-0.926***	-0.766***	+0.116	-0.060	-0.060
	(0.323)	(0.242)	(0.207)	(0.197)	(0.058)
Macro	+0.496	+0.175	+0.226	+0.183	+0.028
	(0.329)	(0.195)	(0.167)	(0.124)	(0.056)
Yield-spread	-0.283	+0.015	-0.020	-0.122	-0.075*
	(0.242)	(0.158)	(0.217)	(0.149)	(0.039)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.159	0.163	0.080	0.058	0.045
Observations	188	188	168	188	168

Table 8: Free-Text Modeling Assumption Clusters and Deviations from the Objective Model Forecast. This table shows how disclosed modeling assumption 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 total 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) matched to the same reference date and horizon. Regressions are estimated at the manager-year-horizon level. Free-text modeling assumption cluster indicators are extracted and constructed from CMA report text using LLM-based classification. Let $\mathcal{C} = \{\text{Decomposition, Historical Calibration, Scenario Adjustment, Present Value Framework, Mean Reversion, Macro, Yield Spread}\}$ denote the set of modeling assumption clusters. 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.

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

	$F^{ER}_{i,t,h} - G_{t,h}$
Decomposition	-0.539**
	(0.273)
Historical Calibration	+0.573**
	(0.223)
Scenario Adjustment	-0.164
	(0.201)
Present Value Framework	+0.175
	(0.309)
Mean Reversion	-0.630**
	(0.264)
Macro	+0.448*
	(0.257)
Yield Spread	-0.264
	(0.211)
Year fixed effects	Yes
R^2	0.369
Observations	179

Table 9: 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 $F_{i,t}^{ER}$, asset manager i's total 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, $C_{i,t}^{m}$ 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 $C_{i,t}^{m} - \overline{C}_i^{\,m}$ captures within-manager variation in complexity, while $\overline{C}_i^{\,m}$ 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.

$$F_{i,t}^{ER} = \mu_i + \tau_t + \beta_w(C_{i,t}^{m} - \overline{C}_i^{\,m}) + \gamma_w(C_{i,t}^{m} - \overline{C}_i^{\,m}) \times S_t + \gamma_b \overline{C}_i^{\,m} \times S_t + \varepsilon_{i,t}$$

m	(1) APL	(2) ICR	(3) NT
$C_{i,t}^{m} - \overline{C}_i^{\,m}$	0.254	-0.289	-0.405
	(0.477)	(0.366)	(0.250)
$(C_{i,t}^{m} - \overline{C}_i^{\,m}) \times S_t$	17.884	-20.420	30.498*
	(20.137)	(16.419)	(17.043)
$\overline{C}_i^{\,m} \times S_t$	-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 10: 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 total 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.

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

	(1)	(2)	(3)	(4)	(5)
q	VC	DY	g	π	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 11: 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 total 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 12: 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 total 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 13: 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), and $F^{ER}_{i,t,h}$ is asset manager i's total 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 14: 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), and $F^{ER}_{i,t,h}$ is asset manager i's total 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.

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

	(1)	(2)	(3)	(4)	(5)
q	VC	DY	g	π	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

Table 15: Second Moment Assumptions. This table reports the methodological approach each asset manager uses to forecast volatility (Methodology Vol.) and correlation (Methodology Corr.), based on our reading of each provider's Capital Market Assumptions report. We classify approaches into three categories. *Estimates* denotes forecasts based on simple historical estimates; *Estimates** denotes historical estimates with small adjustments (e.g., scenario-dependent assumptions or regime overlays); and *Calibrations* denotes calibration-based approaches (e.g., historical regressions or GARCH models). A dash (–) indicates that the provider does not disclose a methodology for the corresponding second moment.

Provider	Methodology Vol.	Methodology Corr.	Provider	Methodology Vol.	Methodology Corr.
ACG	Estimates	Estimates	Janney	-	-
Advent	Calibrations	Calibrations	Janus Henderson	-	-
AIMCo	Calibrations	Calibrations	JP Morgan	Estimates*	Estimates*
Allianz	Calibrations	Calibrations	KKR	Estimates	-
American Century	Estimates*	Estimates*	Mackenzie	-	Estimates*
Amundi	Calibrations	Calibrations	Manulife	-	Estimates*
Angeles	Estimates*	Estimates*	Meketa	Estimates*	Estimates*
AON	Calibrations	Calibrations	Merrill Lynch	Calibrations	Calibrations
AQR	-	-	MFS	Estimates	Estimates
Baillie Gifford	-	-	Morgan Stanley	Estimates*	Estimates*
Barclays	-	-	NEPC	Calibrations	Calibrations
Benjamin F.E.	Estimates*	Estimates*	Neuberger	Estimates*	Estimates*
BlackRock	Calibrations	Calibrations	Ninety One	-	-
BMO	Estimates*	Estimates*	Northern Trust	Estimates*	Estimates*
BNP	Calibrations	-	Nuveen	Estimates	Estimates
BNY Mellon	Estimates	Estimates	OneAscent	Estimates	Estimates
Callan	Calibrations	Calibrations	PGIM	Estimates*	Estimates*
CalPERS	-	-	PIMCO	Calibrations	Calibrations
Candriam	Estimates	Estimates	PineBridge	Estimates*	-
Capital Group	Estimates	Estimates	RAM	Estimates*	Estimates*
CIBC	Estimates	-	RBC	Estimates	-
CIGAM	Calibrations	-	Robeco	Estimates	-
Cliffwater	Estimates	Estimates	Rockefeller	Estimates*	-
Cohen Steers	Estimates*	Estimates	RVK	Estimates*	Estimates*
Columbia Thread.	-	-	Schroders	Estimates	-
Credit Suisse	Estimates*	-	SEI	Estimates*	Estimates*
Crescent Capital	Estimates	Estimates	Sellwood	Estimates*	Estimates
CWO	Estimates	Estimates	SIPA	Estimates	Estimates
Deutsche Bank	-	-	SJS	Estimates*	-
DSA	Estimates	Estimates	State Street	Estimates*	-
DWS	-	-	Syntrinsic	-	-
Edward Jones	Estimates	Estimates	T. Rowe Price	Estimates*	Estimates*
EFG	Calibrations	Calibrations	TD	Estimates	Estimates
Ellwood	-	Estimates	UBS	Estimates*	Estimates*
Envestnet	Calibrations	Calibrations	Vanguard	Calibrations	Calibrations
FI3	-	-	Verus	Estimates*	Estimates*
Fidelity	Calibrations	Calibrations	Voya	Calibrations	Calibrations
Fiducient	-	-	Wealthspire	-	-
Franklin Temp.	Estimates*	Estimates*	Wellington	Estimates	Estimates
Graystone	Estimates*	Estimates*	Wells Fargo	-	-
HSBC	-	-	Wilshire	Estimates*	Estimates*
Invesco	Estimates	Estimates			

Appendix A

A.1 External Data Construction

This appendix describes the construction of the external datasets used in the paper. These data complement the CMA reports by providing realized market outcomes, news flows, portfolio allocations, and shareholder communications.

A.1.1 Market Data and Forecasting Predictors

We use CRSP to compute realized U.S. equity returns and market capitalizations. We use Robert Shiller's data to obtain earnings, dividends, stock returns, and valuation ratios, including the CAPE ratio. We use the Goyal predictor dataset to obtain standard monthly return-predictor variables. We use Bloomberg and FactSet to obtain traded total-return indices used to benchmark realized returns, valuation ratios, volatilities, correlations across asset classes.

A.1.2 RavenPack Earnings-News 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. 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 focus on earnings-related news events. Earnings announcements represent frequent and systematically scheduled firm-specific information releases that provide regular updates on corporate fundamentals. Using this earnings-focused approach, we assess news sentiment at the U.S. equity-market level rather than at the individual-firm level, aggregating firm-level news to the S&P 500 index level, including only companies identified through CRSP monthly market cap data.

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).

We first aggregate news at the parent company level and then perform fuzzy matching between company names in RavenPack and CRSP. We define a value-weighted sentiment measure:

$$\overline{S}_{\tau}^{(\text{wtd})} = \frac{\sum_{\ell \in \mathcal{I}_\tau} w_{\ell,\tau} S_{\ell,\tau}}{\sum_{\ell \in \mathcal{I}_\tau} w_{\ell,\tau}} \tag{A1}$$

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 CRSP market-capitalization weights. In constructing news horizons aligned with the CMA report date, we aggregate news based on their publication dates at the daily frequency, following equation (A1). Subsequently, using the resulting daily time series, we calculate average news measures over specified time intervals. In the empirical specifications below, S_t denotes the annual value-weighted earnings-news sentiment measure constructed from $\overline{S}_{\tau}^{(\text{wtd})}$ and matched to CMA reference year t.

A.1.3 Morningstar Fund Allocations

We investigate portfolio-allocation decisions by analyzing 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 (Dahlquist & Ibert 2024). 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 track net fund allocations in U.S. equity markets.[14]

[14]We match CMA reference dates to fund portfolio observations within defined time windows. Our baseline uses a symmetric three-month window around the CMA embedded date.

A.1.4 SEC Shareholder-Letter Outlooks

We augment the analysis with shareholder communications from mutual fund annual and semi-annual reports filed with the U.S. Securities and Exchange Commission (SEC) 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 SEC EDGAR filings over the period 2008–2026, retaining shareholder-report form types, primarily N-CSR, N-CSR/A, and N-CSRS. For each filing, we 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.

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 Extraction

We extract the body text from each CMA report and strip asset manager identifiers from the corpus, so that the extraction model does not key on institution names; a bigram pass catches manager-name fragments that a token-level strip would miss. The final corpus comprises 528 reports, which we pass to the LLM causal-extraction pipeline described in Appendix A.2. Because the LLM operates on running text rather than a bag of words, no tokenization, lemmatization, stopword removal, or n-gram construction is required.

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.

[Figure A1 here]

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 Network Complexity Measures

The pipeline above yields, for each manager–year, a directed network whose nodes are financial topics and whose edges are the extracted causal links. Our representation shares the intuition of directed acyclic graph (DAG) frameworks, which formalize causal reasoning as nodes connected by directed edges, but departs from them in two ways suited to our setting. First, we extract *stated* causal claims from managers' narratives rather than inferring causal structure from data, so each edge records what a manager asserts rather than an identified causal effect. Second, we do not impose acyclicity: because managers routinely articulate feedback loops, the extracted objects are directed graphs that may contain cycles rather than DAGs in the strict sense. Preserving these cycles is deliberate—it captures heterogeneity in how institutions conceptualize the same forecasting problem and is precisely what the transitivity- and path-based measures below are designed to quantify.

Chain complexity. *Average path length* is the average number of causal steps connecting topics in the network; a higher value means the reasoning relies on longer chains. 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* is the share of links belonging to multi-step causal paths leading to, or from, asset class return expectations. The two are not mathematically identical, but both measure the extent to which reasoning relies on intermediate mechanisms rather than direct relationships.

Interconnected complexity. *Network transitivity* measures the extent to which the causal narrative forms closed triangles. If the network contains

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

then A, B, and C form a transitive triangle. A high-transitivity narrative is not necessarily longer, but it is more locally interconnected.

Figure A2 illustrates both dimensions of narrative complexity.

[Figure A2 here]

A.2.4 Empirical Network Measures

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 A.2.2.

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

$$\text{Avg. Path Length} = \frac{1}{|P|} \sum_{(i,j) \in P} d(i, j) \tag{A2}$$

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:

$$\text{Indirect Ratio} = \frac{|\mathcal{E}indirect|}{|\mathcal{E}total|} \tag{A3}$$

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:

$$\text{Transitivity} = \frac{3 \times \text{number of triangles}}{\text{number of connected triples}} \tag{A4}$$

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).

Descriptive Statistics. Table A1 reports the average number of edges and nodes per report for each asset manager in our dataset. Table A2 reports summary statistics for both families of attention and complexity measures. 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 exceed 1. Asset managers thus differ substantially both in how they structure causal reasoning and in which economic topics they emphasize, which allows us to study why sophisticated investors disagree despite observing similar macro-financial information.

[Tables A1 and A2 here]

Figure A3 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 the sum of economic growth and earnings growth appear in roughly than 60% of asset manager-years, respectively.

[Figure A3 here]

A.2.5 A Conceptual Network Approach

Network representations arise naturally in asset-pricing settings because return identities link expected returns, expected cash-flow growth, and expected valuation changes. As an example, we 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.

Starting from the Campbell–Shiller approximation,

$$r_{t+1} = \kappa + \Delta d_{t+1} - pd_t + \rho\, pd_{t+1} \tag{A5}$$

$$= \kappa + \Delta d_{t+1} - (1-\rho) pd_t + \rho\, \Delta pd_{t+1} \tag{A6}$$

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

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] = \kappa + \widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] - (1-\rho)\, pd_t + \rho\, \widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}] \tag{A7}$$

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.[15] Graphically, these directional accounting relationships can be represented as the network shown in Figure A4, 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

[15] Formally, the directional accounting links implied by equation (A7) are obtained by solving the identity for each object in turn. They are given by:

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

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

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

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

model that agent i might use as a belief formation framework. The agent considers a perfectly competitive endowment economy with complete markets. In his mind, the representative agent has lifetime utility over consumption:

$$U_t = \widetilde{\mathbb{E}}_{i,t}\left[\sum_{j=0}^{\infty} \beta^j \frac{C_{t+j}^{1-\gamma}}{1-\gamma}\right] \tag{A8}$$

subject to the budget constraint:

$$\begin{aligned} W_{t+1} &= \omega_t'(\mathbf{1} + \mathbf{R}_{t+1})(W_t - C_t) \\ \omega_t'\mathbf{1} &= 1 \end{aligned} \tag{A9}$$

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:

$$\Delta d_t = \Delta c_t = (1-\phi)\mu + \phi \Delta c_{t-1} + \epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1) \tag{A10}$$

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 (A7) can then be written as functions of consumption growth:

$$pd_t = A + B\,\Delta c_t, \tag{A11}$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta d_{t+1}] = (1-\phi)\mu + \phi \Delta c_t, \tag{A12}$$

$$\widetilde{\mathbb{E}}_{i,t}[\Delta pd_{t+1}] = B(1-\phi)(\mu - \Delta c_t), \tag{A13}$$

$$\widetilde{\mathbb{E}}_{i,t}[r_{t+1}] = \kappa + (\rho - 1)A + (1-\phi)\mu(1+\rho B) + [\phi(1+\rho B) - B]\,\Delta c_t \tag{A14}$$

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 A4. 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 A4 here]

A.2.6 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.2.7 Methodology Extraction

Preset Modeling Assumptions Extraction. We extract the preset 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.

Free-Text Modeling Assumptions Extraction. The seven free-text technique clusters behind asset managers' equity return forecasts 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.

A.3 CMA Reports: Text Length and Authors

We provide summary statistics describing the length of the textual components of CMA reports and the characteristics of their authors.

A.3.1 CMA Text Length

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 A5 summarizes these patterns. Panel (a) shows substantial variability in CMA report lengths over time, with no clear trend in the median, which varies around 5000 words. Panel (b) highlights considerable variation in typical report length, pointing to differences in communication strategies, which we explore with our text analysis.

[Figure A5 here]

A.3.2 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 A6 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 publi-

cation. 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 A7 reports the resulting manager-level distribution. This dispersion indicates that the human-capital composition behind CMA forecasts differs systematically across institutions.

[Figures A6 and A7 here]

A.4 Deriving Approximate Return Identities

We derive approximate return identities for equity and fixed-income asset classes that align with the building-block return decompositions used by institutional asset managers in CMA reports.

A.4.1 Equities

Starting with a h-period identity:

$$1 + R_{t+h} = \frac{D_{t+h} + P_{t+h}}{P_t} = \frac{D_{t+h}}{P_t} + \frac{P_{t+h}}{P_t} \tag{A15}$$

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

$$\begin{aligned} 1+R_{t+h} &= \frac{D_{t+h}}{P_{t+h}}\frac{(P/X)_{t+h}}{(P/X)_t}\frac{X_{t+h}}{X_t} + \frac{(P/X)_{t+h}}{(P/X)_t}\frac{X_{t+h}}{X_t} && \text{(A16)} \\ &= (1+DY_{t+h})\,(1+VC_{t+h})\,(1+g_{nominal,t+h}) && \text{(A17)} \\ &= (1+DY_{t+h})\,(1+VC_{t+h})\,(1+g_{real,t+h})\,(1+\pi_{t+h}) && \text{(A18)} \end{aligned}$$

where $DY_{t+h} = D_{t+h}/P_{t+h}$ (commonly referred to as "income", "net payout yield", "carry", or "dividend yield"), $1+VC_{t+h} = (P/X)_{t+h}/(P/X)_t$ (commonly referred to as "valuation adjustment" or "change in valuation"), $1+g_{nominal,t+h} = X_{t+h}/X_t$ (commonly captured under "EPS growth" or combined with inflation under "nominal growth"), and π_{t+h} is the inflation rate. Taking a first-order approximation by dropping higher-order interaction terms yields:

$$R_{t+h} \approx DY_{t+h} + VC_{t+h} + g_{real,t+h} + \pi_{t+h} \tag{A19}$$

This identity highlights how the total return can be expressed as a linear function of valuation changes, growth rates, dividend yields, and inflation.[16]

A.4.2 Fixed Income

Consider a default-free coupon bond with deterministic promised cash flows $\{CF_i\}_{i=1}^N$ paid at future times $\{\tau_i\}_{i=1}^N$. Using continuously compounded yield y_t, its ex-coupon price is

$$P_t(y_t) = \sum_{i=1}^{N} CF_i e^{-y_t \tau_i}. \tag{A20}$$

After the holding period, the aged bond price as a function of yield is

$$P_{t+h}(y) = \sum_{\tau_i > h} CF_i e^{-y(\tau_i - h)}. \tag{A21}$$

[16] Some asset managers provide finer decompositions that separately report margins, buybacks, and issuance. We harmonize these decompositions by assigning margin expansion to the real earnings-growth component and by treating dividends, buybacks, and issuance as components of total payout yield, following Boudoukh et al. (2007).

For a short holding period h, the simple holding-period return of default-free bonds is

$$R_{t,t+h}^{Default\,Free} = \frac{P_{t+h}(y_{t+h}) + CF_{t,t+h} - P_t(y_t)}{P_t(y_t)}, \tag{A22}$$

where $CF_{t,t+h}$ denotes coupon or principal cash flows received during the holding period, measured at time $t+h$.

Add and subtract $P_{t+h}(y_t)$, the time-$t+h$ price that would prevail if the yield were unchanged:

$$R_{t,t+h}^{Default\,Free} = \frac{CF_{t,t+h} + P_{t+h}(y_t) - P_t(y_t)}{P_t(y_t)} + \frac{P_{t+h}(y_{t+h}) - P_{t+h}(y_t)}{P_t(y_t)}. \tag{A23}$$

The first term is the return earned from coupon income plus the deterministic aging of the bond under an unchanged yield. This is commonly called *carry plus roll-down*. The second term is the mark-to-market return caused by the yield change.

Define

$$\Delta y_{t+h} = y_{t+h} - y_t. \tag{A24}$$

Then the second term in (A23) can be approximated by a Taylor expansion.

Take a second-order Taylor expansion of $P_{t+h}(y_t + \Delta y_{t+h})$ around y_t:

$$P_{t+h}(y_t + \Delta y_{t+h}) \approx P_{t+h}(y_t) + \frac{\partial P_{t+h}(y)}{\partial y}\bigg|_{y=y_t} \Delta y_{t+h} + \frac{1}{2}\frac{\partial^2 P_{t+h}(y)}{\partial y^2}\bigg|_{y=y_t} (\Delta y_{t+h})^2. \tag{A25}$$

Substituting (A25) into the valuation-change term in (A23) gives

$$\frac{P_{t+h}(y_{t+h}) - P_{t+h}(y_t)}{P_t(y_t)} \approx \frac{1}{P_t(y_t)}\frac{\partial P_{t+h}(y)}{\partial y}\bigg|_{y=y_t} \Delta y_{t+h} + \frac{1}{2P_t(y_t)}\frac{\partial^2 P_{t+h}(y)}{\partial y^2}\bigg|_{y=y_t} (\Delta y_{t+h})^2. \tag{A26}$$

Define the initial-price-normalized duration and convexity coefficients by

$$D_{t,h}^* = -\frac{1}{P_t(y_t)}\frac{\partial P_{t+h}(y)}{\partial y}\bigg|_{y=y_t}, \qquad C_{t,h}^* = \frac{1}{P_t(y_t)}\frac{\partial^2 P_{t+h}(y)}{\partial y^2}\bigg|_{y=y_t}. \tag{A27}$$

Then

$$\frac{P_{t+h}(y_{t+h}) - P_{t+h}(y_t)}{P_t(y_t)} \approx -D_{t,h}^* \Delta y_{t+h} + \frac{1}{2} C_{t,h}^* (\Delta y_{t+h})^2. \tag{A28}$$

The convexity term appears because the price–yield relation is nonlinear: the duration coefficient is proportional to the first derivative of price with respect to yield, while the convexity coefficient

is proportional to the second derivative.

Combining (A23) and (A28) and taking subjective expectations,

$$\widetilde{\mathbb{E}}_t[R^{Default\,Free}_{t,t+h}] \approx \underbrace{\widetilde{\mathbb{E}}_t\left[\frac{CF_{t,t+h}}{P_t(y_t)}\right]}_{\text{Carry}_{t,h}} + \underbrace{\widetilde{\mathbb{E}}_t\left[\frac{P_{t+h}(y_t)-P_t(y_t)}{P_t(y_t)}\right]}_{\text{Roll}_{t,h}} - \underbrace{\widetilde{\mathbb{E}}_t\left[D^*_{t,h}\Delta y_{t+h}\right]}_{\text{Duration Component}_{t,h}} + \underbrace{\widetilde{\mathbb{E}}_t\left[\frac{1}{2}C^*_{t,h}(\Delta y_{t+h})^2\right]}_{\text{Convexity Component}_{t,h}}. \tag{A29}$$

For bonds carrying credit risk, two features must be added to (A29): the yield is no longer purely a risk-free rate, and cash flows are subject to default. It is convenient to resolve the yield into a risk-free level and a spread. Let f_t be the short risk-free rate, let r_t be the risk-free yield at the bond's maturity, and let $s_t = y_t - r_t$ be the credit spread. The yield then admits the anatomy

$$y_t = \underbrace{f_t}_{\text{Short Risk-Free Rate}_t} + \underbrace{(r_t - f_t)}_{\text{Duration Spread}_{t,h}} + \underbrace{(y_t - r_t)}_{\text{Credit Spread}_{t,h}}, \tag{A30}$$

Because a bond held at an unchanged yield earns exactly that yield, the Cash Yield—carry plus roll—inherits the same three-way split,

$$\text{Cash Yield}_{t,h} = \text{Short Risk-Free Rate}_t + \text{Duration Spread}_{t,h} + \text{Credit Spread}_{t,h}. \tag{A31}$$

The same decomposition applies to the yield change. Since $y = r + s$,

$$\Delta y_{t+h} = \Delta r_{t+h} + \Delta s_{t+h}, \tag{A32}$$

where Δr_{t+h} is the change in the (maturity-matched) risk-free yield and Δs_{t+h} is the change in the spread. The duration component in (A29) therefore separates into a rate-driven and a spread-driven revaluation, $-D^*_{t,h}\Delta y_{t+h} = -D^*_{t,h}\Delta r_{t+h} - D^*_{t,h}\Delta s_{t+h}$.

Introducing a credit-loss term $L_{t,t+h}$ for the expected loss from default over the holding period, the defaultable expected return becomes

$$\widetilde{\mathbb{E}}_t[R^{Credit}_{t,t+h}] \approx \underbrace{\underbrace{\widetilde{\mathbb{E}}_t\left[\frac{CF_{t,t+h}}{P_t(y_t)}\right]}_{\text{Carry}_{t,h}} + \underbrace{\widetilde{\mathbb{E}}_t\left[\frac{P_{t+h}(y_t)-P_t(y_t)}{P_t(y_t)}\right]}_{\text{Roll}_{t,h}}}_{\text{Cash Yield}_{t,h}}$$

$$
\underbrace{
\underbrace{-\,\widetilde{\mathbb{E}}_t\big[D^*_{t,h}\Delta r_{t+h}\big]}_{\text{Rate Revaluation}_{t,h}}
\underbrace{-\,\widetilde{\mathbb{E}}_t\big[D^*_{t,h}\Delta s_{t+h}\big]}_{\text{Spread Revaluation}_{t,h}}
\underbrace{+\,\widetilde{\mathbb{E}}_t\big[\tfrac{1}{2}C^*_{t,h}(\Delta y_{t+h})^2\big]}_{\text{Convexity Component}_{t,h}}
}_{\text{Valuation Change}_{t,h}}
\underbrace{-\,\widetilde{\mathbb{E}}_t\left[\frac{L_{t,t+h}}{P_t(y_t)}\right]}_{\text{Credit Loss}_{t,h}}. \tag{A33}
$$

Equation (A33) is the general form. The "Cash Yield" collects income and deterministic aging and, through (A31), resolves into a short risk-free rate, a duration (term) spread, and a credit spread. The "Valuation Change" collects the mark-to-market from expected moves in the risk-free curve and in the spread, together with the second-order convexity gain. The convexity gain depends on the variance of the yield change and therefore remains positive even in the absence of a directional view; it also absorbs the rate–spread cross term, which is why it is kept as a single term rather than split into rate and spread pieces. The "Credit Loss" subtracts expected default losses.[17]

A.5 Objective Machine Learning Benchmark: Details

This appendix details the construction of the objective benchmarks introduced in Section 2.3 and their evaluation. We build a benchmark for the total U.S. equity return and for each of the four building blocks of the decomposition in (A18), and we evaluate all five against the corresponding CMA forecasts. Throughout, let $b \in \{ER, DY, VC, g_{real}, \pi\}$ index the target, let $Y^b_{t,t+h}$ denote its annualised realised value from t to $t+h$, and let $F^b_{i,t,h}$ denote asset manager i's corresponding CMA forecast.

Targets. The realized total-return target is the annualised geometric total return:

$$
Y^{ER}_{t,t+h} = \left(\frac{P_{t+h}}{P_t}\right)^{1/h} - 1 \tag{A34}
$$

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

The four block targets are the realised counterparts of the components on the right-hand side of (A18), of which the CMA reports state the subjective expectation. The realised income component

[17] "Duration" enters in two distinct roles: as the "Duration Spread" in (A31), a yield component rewarding position on the curve, and as the coefficient $D^*_{t,h}$ in the revaluation terms, a mark-to-market sensitivity to a yield change.

$Y^{DY}_{t,t+h}$ is the dividend yield over the horizon, taken from the Goyal et al. (2024) dataset. The realised valuation change $Y^{VC}_{t,t+h}$ is the change in the cyclically adjusted price-earnings ratio CAPE; the realised real growth $Y^{g_{real}}_{t,t+h}$ is the change in the smoothed real earnings E^{10} = RealPrice/CAPE that sit in the denominator of that ratio, so that the price part of the decomposition holds exactly; and the realised inflation component $Y^{\pi}_{t,t+h}$ is the change in the consumer price index. These three series are taken from Robert Shiller's website and are annualised over the horizon. All level series are sampled at year end, so that a horizon-h target spans December of year t to December of year $t+h$.

Predictors. 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 A3. 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. The same 16 predictors, with the same one-year lag, are used for the total return and for each block.

[Table A3 here]

Estimation. For each target b, the objective forecast $G^b_{t,h}$ is the equal-weighted average of $M = 7$ regularised forecasting models:

$$G^b_{t,h} = \frac{1}{M}\sum_{m=1}^{M} \phi_b\left(\hat{Y}^{b,m}_{t,t+h}\right) \tag{A35}$$

where $\hat{Y}^{b,m}_{t,t+h}$ is the forecast of base model m and ϕ_b imposes the Campbell & Thompson (2008) non-negativity constraint, $\phi_b(x) = \max(0, x)$, on each base forecast before averaging. The constraint is applied only to the targets that are bounded below by zero by construction, namely the total return and income; for valuation change, real growth and inflation, whose realised values are routinely negative, ϕ_b is the identity.

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 a real-time expanding window with a minimum of $T_{\min} = 30$ training observations: an observation enters the training set only if its horizon has elapsed before

the forecast date, so that no realised value observed after that date informs the estimation. 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. The estimation is otherwise identical across the five targets; only the left-hand-side variable and the constraint ϕ_b change.

Evaluation. We test for differences in predictive accuracy between $G^b_{t,h}$ and $F^b_{i,t,h}$ using the Diebold & Mariano (1995) test. For each target b, define the squared-error loss differential:

$$d^b_{i,t,h} = \left(e^b_{\mathrm{CMA},i,t,h}\right)^2 - \left(e^b_{\mathrm{G},t,h}\right)^2$$

where the two forecast errors are measured against the realised value of that target:

$$e^b_{\mathrm{CMA},i,t,h} = Y^b_{t,t+h} - F^b_{i,t,h}, \qquad e^b_{\mathrm{G},t,h} = Y^b_{t,t+h} - G^b_{t,h}$$

Let $\overline{d}^b_h$ be the average of $d^b_{i,t,h}$ across the manager–year observations available, and let n_h denote the number of forecast pairs. The Diebold–Mariano statistic is:

$$\mathrm{DM}^b_h = \frac{\overline{d}^b_h}{\sqrt{\widehat{V}(\overline{d}^b_h)/n_h}} \tag{A36}$$

where $\widehat{V}(\overline{d}^b_h)$ is the Newey–West HAC long-run variance estimator with lag $h-1$ (Newey & West 1987). A positive value of DM^b_h indicates that the objective benchmark has lower squared forecast errors than the CMA forecasts. Note that every manager writing in a given year and horizon shares the same objective forecast, and hence the same objective error.

For the total return we run the test horizon by horizon. For the blocks we collect forecast pairs over horizons $h \in \{3, \ldots, 10\}$ and, for each block, stack them into a single test. We do not test each block–horizon cell separately, because a realised component is known only for the reports whose horizon has elapsed, which leaves at most a handful of pairs per cell. For the same reason the evaluable samples are an order of magnitude smaller than for the total return, between 14 and 19 pairs per block. We therefore read the ordering of the root mean squared errors as the substantive result and treat the significance of the test as indicative, and we report the block statistics with the small-sample correction of Harvey et al. (1997).

Table A4 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), and the per-block comparison in Panel (c). At the level of the total return, 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. Across building blocks the comparison is consistent. The benchmark is more accurate than managers for valuation change, where its root mean squared error is 3.40% per year against 4.92%, for real growth, 2.05% against 2.59%, for dividend yield, 0.82% against 0.88%, and for inflation, 0.89% against 1.47%.

[Table A4 here]

Figure A1: 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.

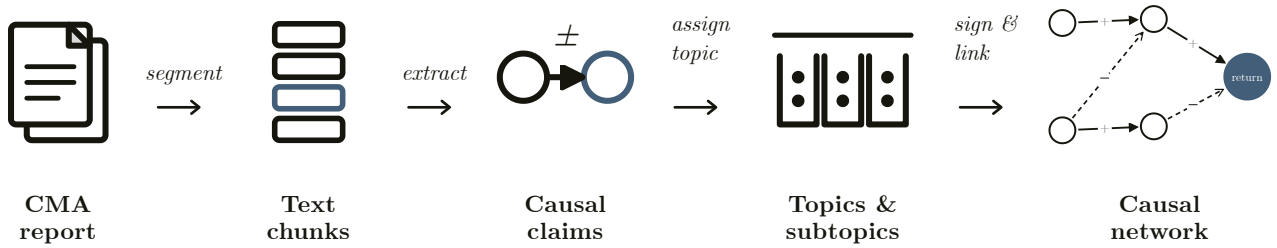


Figure A2: Schematic illustration of the main network-complexity measures. Panel (a) shows chain complexity: average path length and the indirect connection ratio are closely related measures of whether the manager explains a focal outcome through direct links or through multi-step causal chains. Panel (b) shows interconnected complexity: network transitivity measures whether the manager's causal reasoning forms closed triangles of relationships.

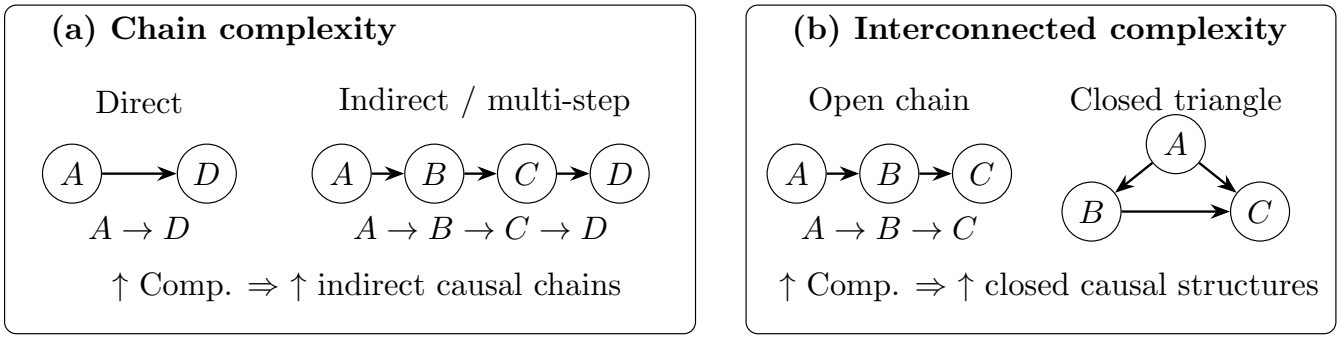
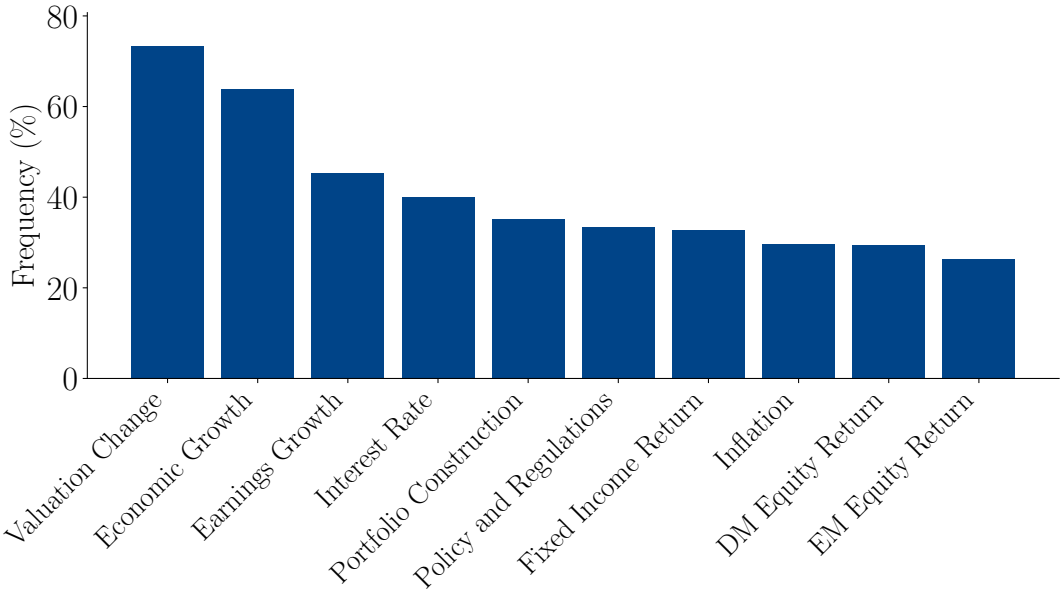


Figure A3: Causal Network Topic Composition. This figure shows the prevalence of topics in asset managers' causal networks. Panel (a) shows the percentage of asset manager-year observations where each topic appears in any causal relationship (as cause or effect). Panel (b) shows the percentage of asset 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

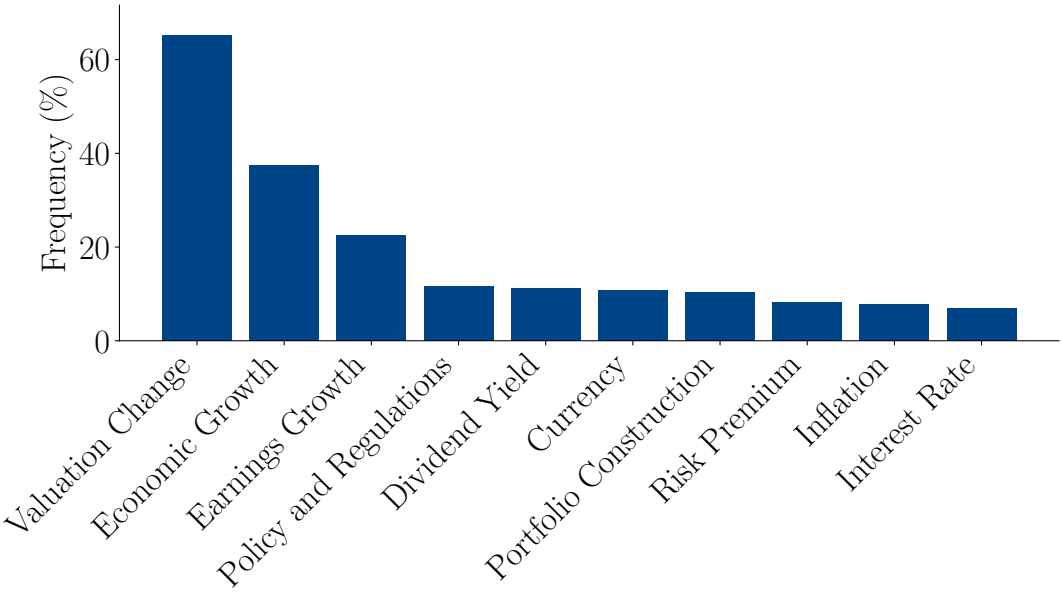


Figure A4: Causal network derived from a Campbell–Shiller (CS) approximation and a classic consumption-based asset-pricing model. Each node represents either a component of the subjective CS approximation or a state variable in the asset-pricing model of agent i, while each directed edge corresponds 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.

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

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

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

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

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

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

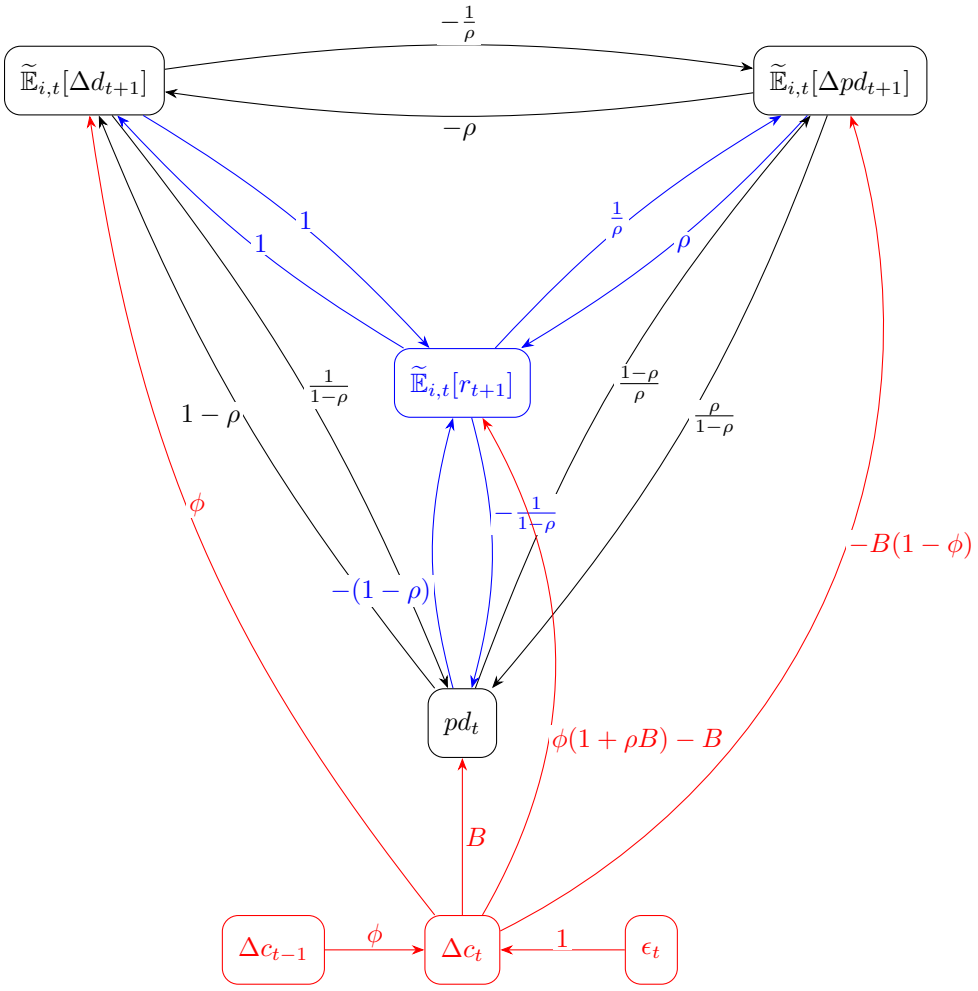
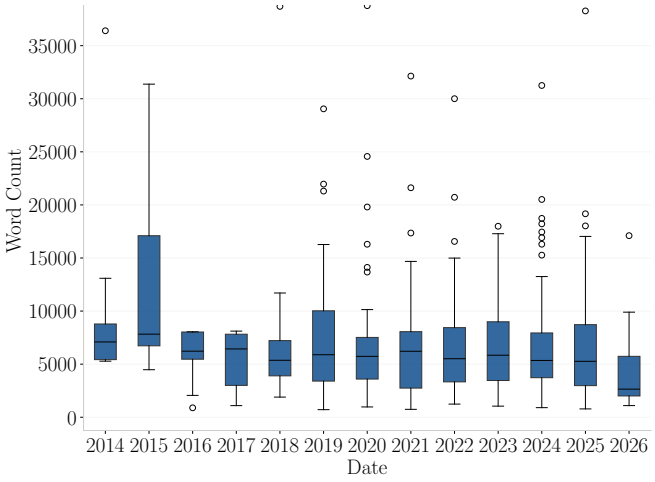


Figure A5: Evolution of CMA Report Textual Characteristics. This figure shows the evolution of CMA report lengths measured by word count. Panel (a) displays the temporal evolution of CMA report lengths from 2014 to 2026. Panel (b) shows the distribution of report word counts; for readability, word counts are capped at 40,000. Sample: CMA reports, 2014–2026.

(a) CMAs Texts Length Over Time



(b) CMAs Texts Length

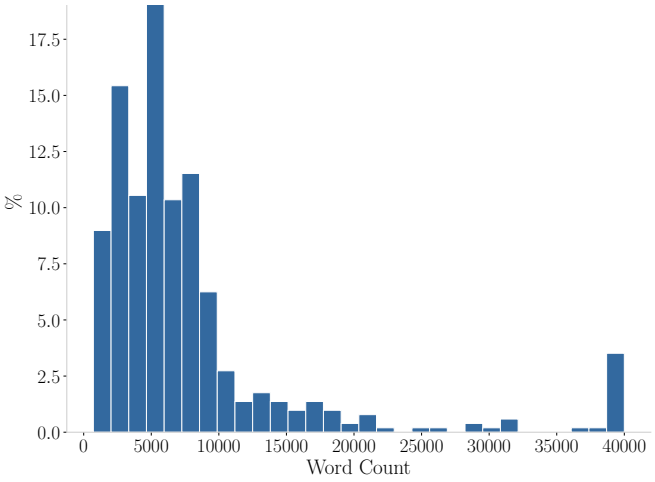


Figure A6: CMA Report Authors: Distributions of Seniority, Education, and Experience. This figure summarizes the panel of biographical information on the individuals listed as authors of the CMA reports in our sample, hand-collected from LinkedIn, personal websites, and company websites. Panel (a) shows 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: 2008–2024.

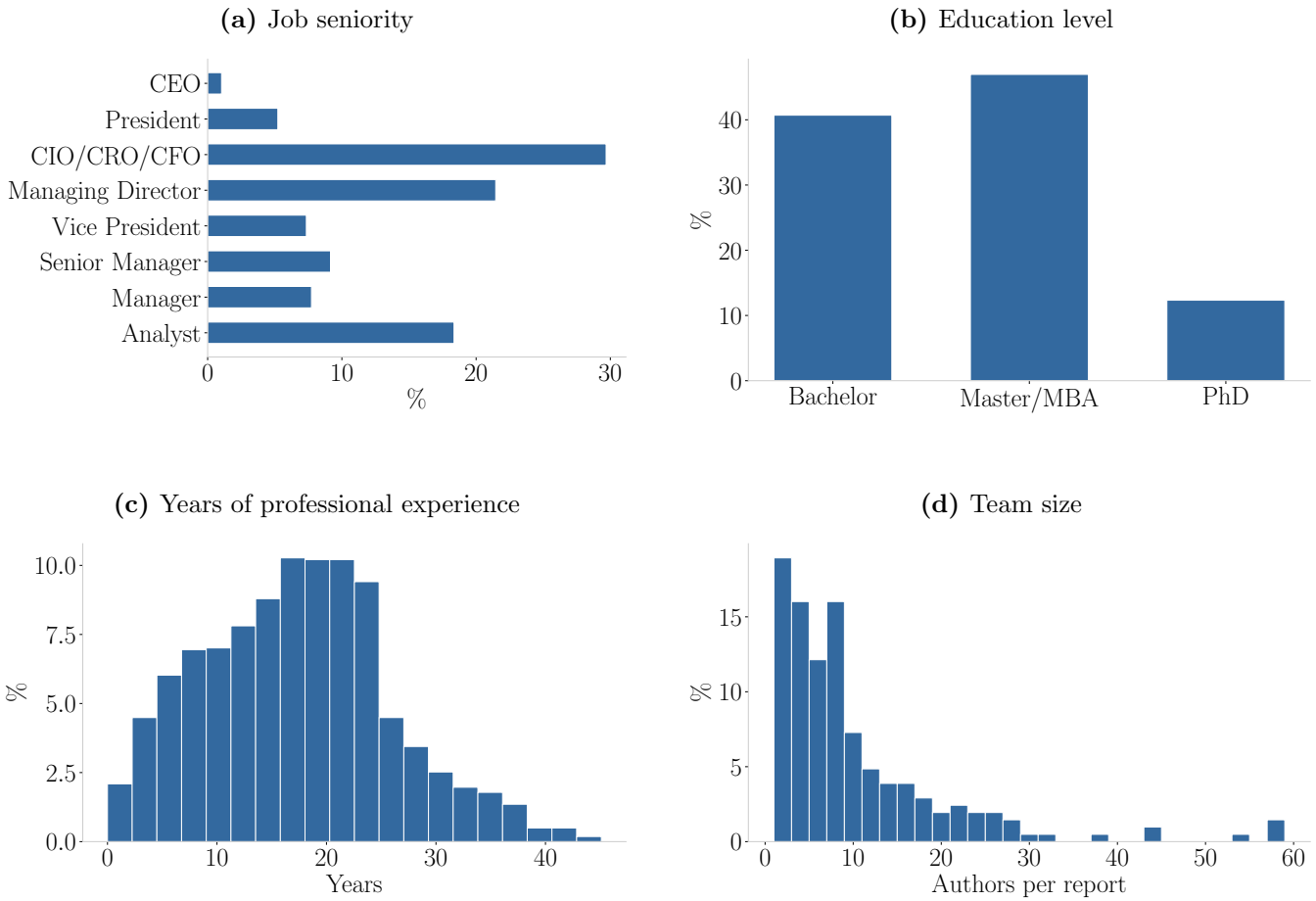


Figure A7: Authors' Educational Composition by Asset Manager. This figure shows the educational composition of CMA report authoring teams, computed at the asset manager level. For each report, we compute the share of authors with Bachelor, Master/MBA, and PhD degrees, and we then average these per-report 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.

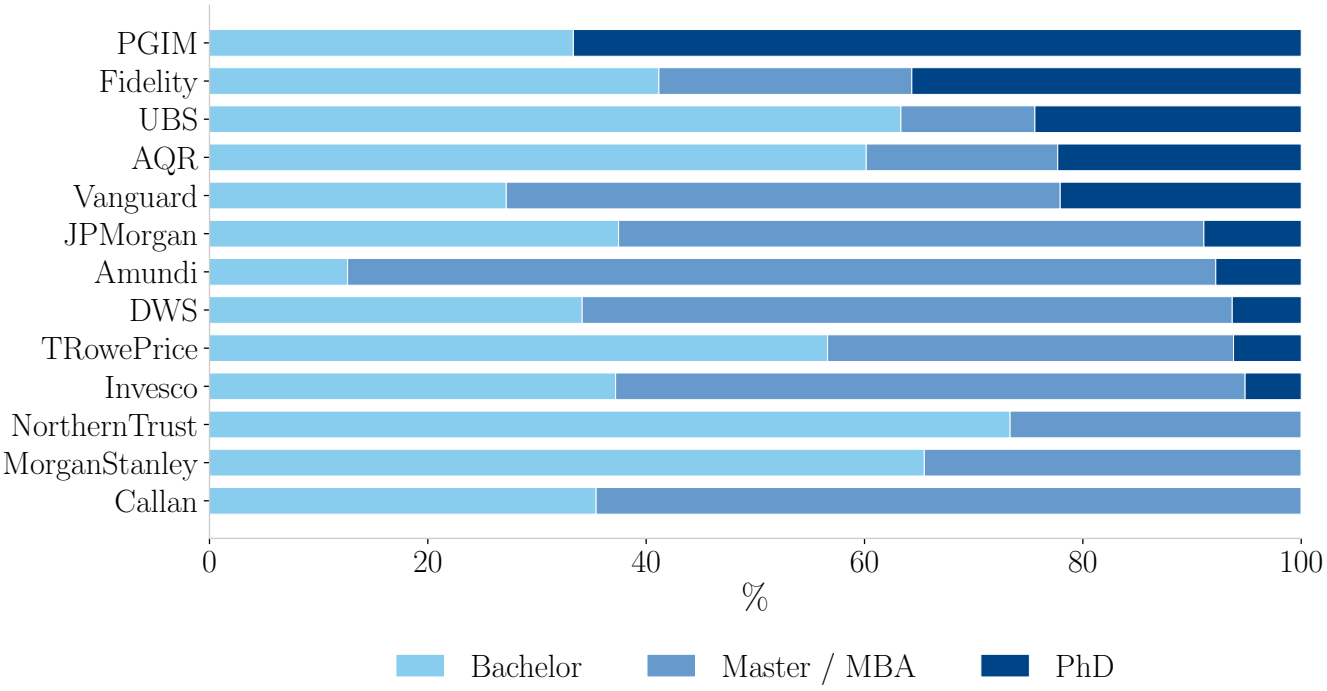


Figure A8: Sources of Variation in Subjective Return Expectations, Volatility Forecasts, and Correlation Forecasts for Alternative Asset Classes. This figure reports variance decompositions based on partial R^2 from saturated fixed-effects models with asset-manager and year fixed effects. Panels A and B report decompositions for asset-class-specific return expectations and volatility forecasts, respectively, for four alternative asset classes: Commodities, Hedge Funds, Private Equity, and Private Real Estate. Each stacked bar shows the percentage of total variation explained by asset-manager fixed effects, year fixed effects, and the residual component. Panel C reports decompositions for pairwise correlation forecasts among these alternative asset classes and three public-market reference asset classes: U.S. Large Cap Equities (US LC Eq), U.S. Aggregate Bonds (US Agg), and U.S. High Yield Bonds (US HY). For each asset-class pair, the lower-left triangle of the heatmap reports the percentage of variation explained by asset-manager fixed effects (in blue), while the upper-right triangle reports the corresponding percentage explained by year fixed effects (in red). Sample period: 2008–2026.

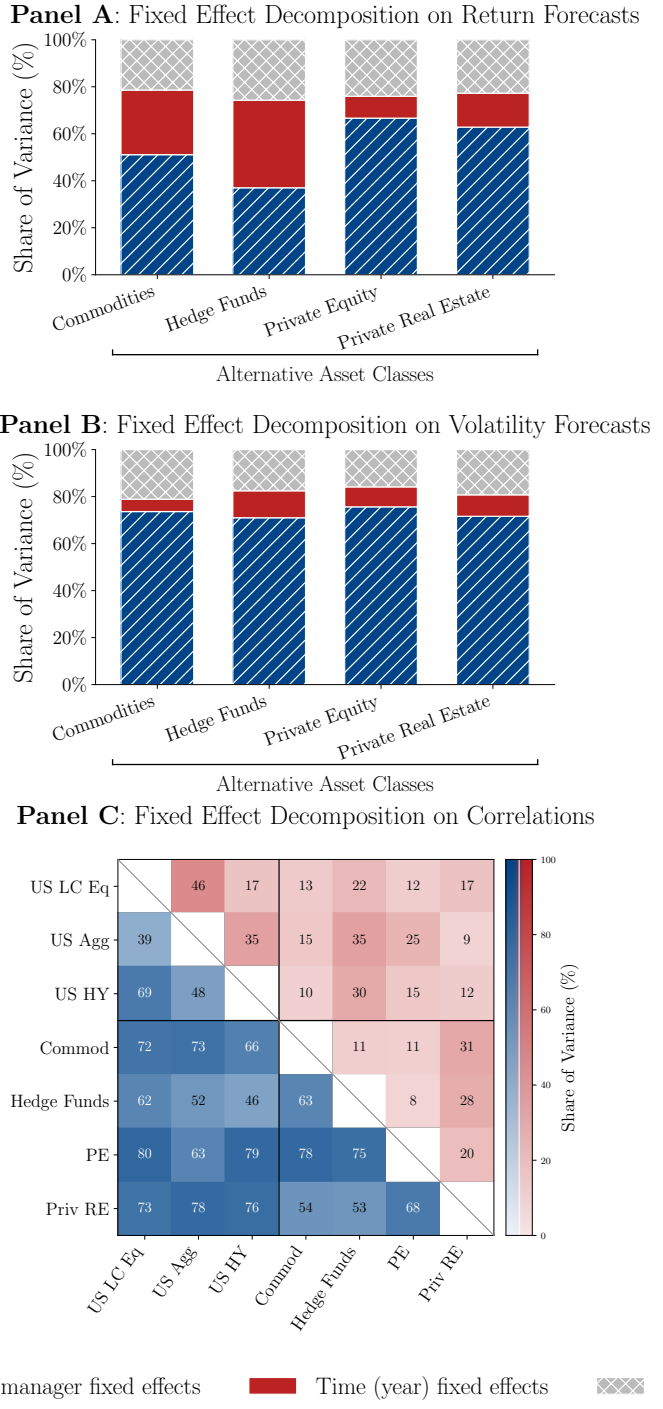


Figure A9: Within-Year Decomposition of Disagreement in Fixed-Income Return Expectations. This figure decomposes the cross-sectional dispersion of total fixed-income return expectations into their reported building blocks—income (inclusive of inflation), valuation change, credit loss for credit sectors, and roll-down—for U.S. government bonds, U.S. investment-grade bonds, U.S. high-yield bonds, and U.S. aggregate bonds. Government bonds have three blocks; investment grade, high yield, and the aggregate add a credit-loss block. All variables are residualized with year fixed effects, equivalently demeaned within survey year, so the plotted variation reflects disagreement across forecasters rather than common time-series movements. Blue bars report the covariance-share decomposition, which measures each block's net contribution to dispersion in the total-return forecast. Red bars report the Type II ANOVA decomposition, which measures each block's independent contribution holding the other blocks fixed. Gray numbers in the background report pairwise correlations among the residualized building-block forecasts. Sample period: 2008–2026.

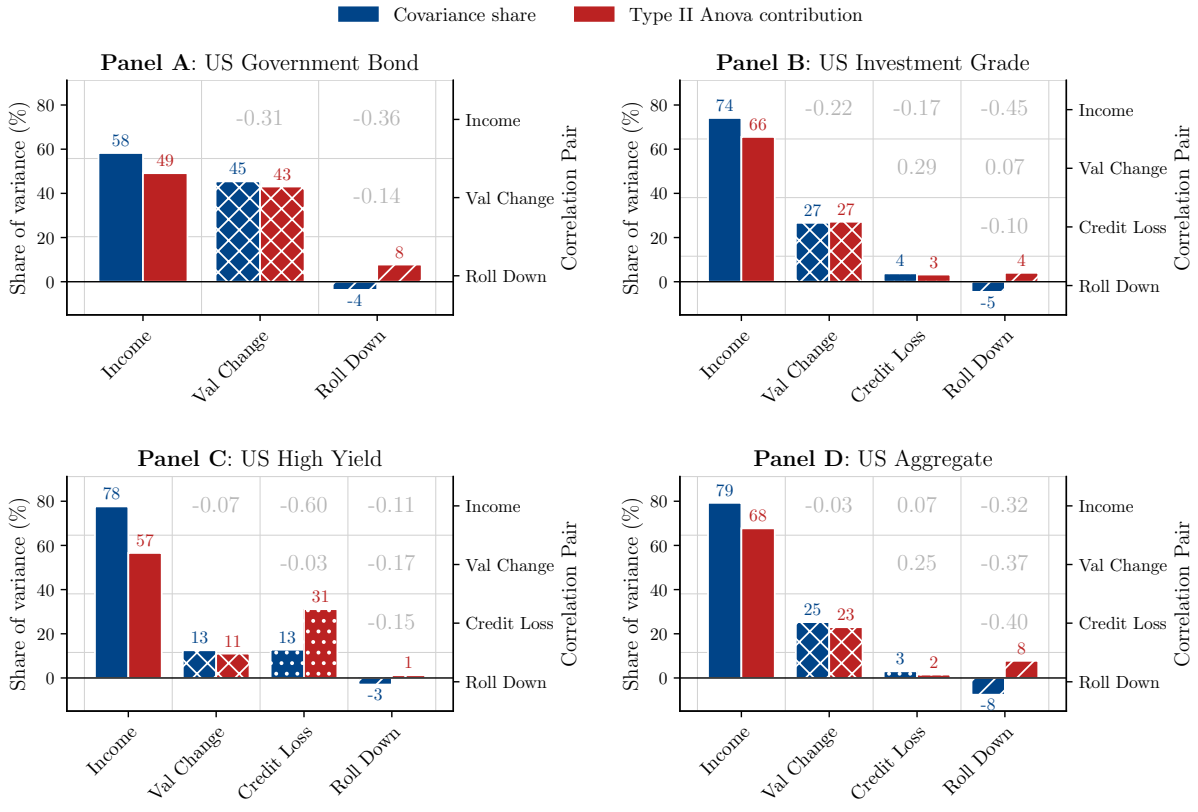


Figure A10: Within-Manager Decomposition in Fixed-Income Return Expectations. This figure decomposes the within-manager variation in total fixed-income return expectations into their reported building blocks—income , valuation change, credit loss for credit sectors, and roll-down—for each asset class. Government bonds have three blocks; investment grade, high yield, and the aggregate add a credit-loss block. All variables are residualized with manager (provider) fixed effects, equivalently demeaned within manager, so the plotted variation reflects how a manager's expectations move over time around their own average, rather than differences across managers. Blue bars report the covariance-share decomposition, which measures each block's net contribution to the variance of the total-return forecast. Red bars report the Type II ANOVA decomposition, which measures each block's independent contribution holding the other blocks fixed. Gray numbers in the background report pairwise (within-manager) correlations among the residualized building-block forecasts. Sample period: 2008–2026.

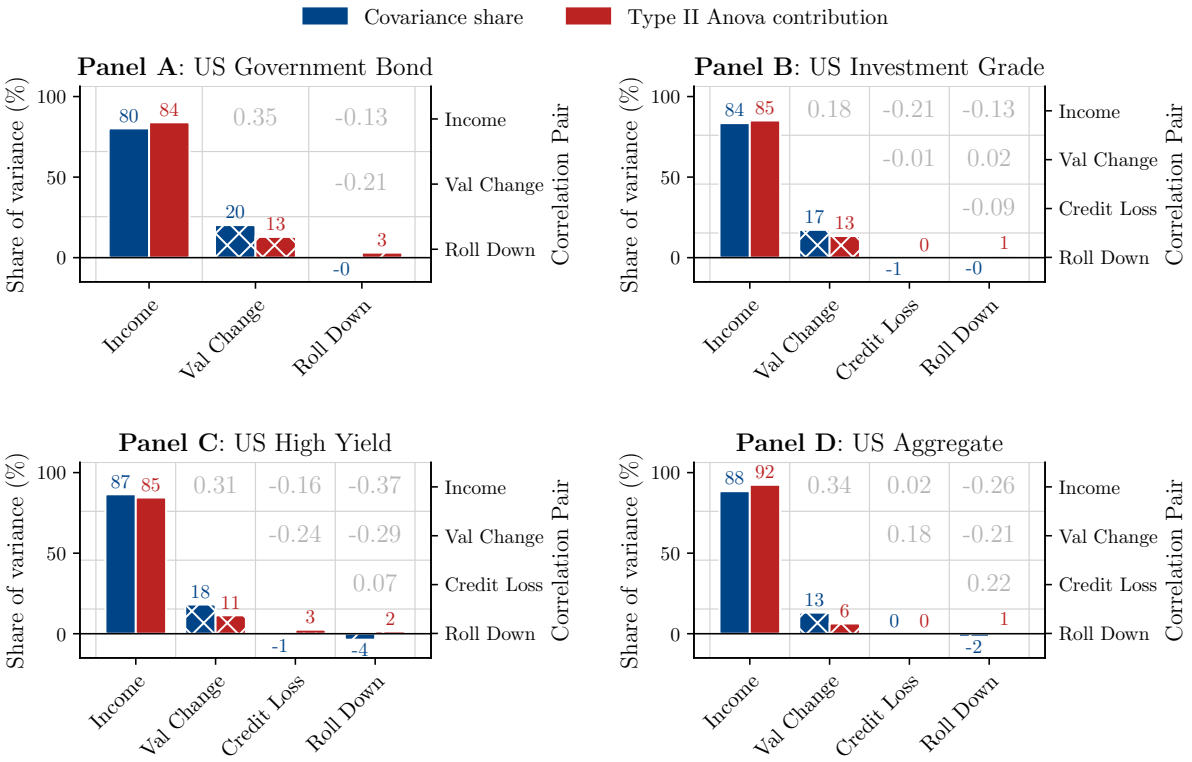
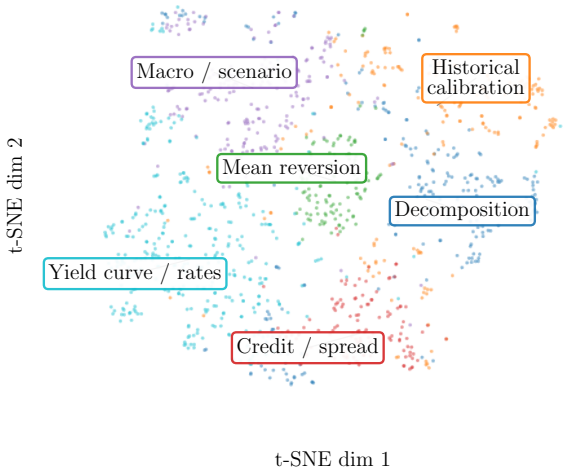
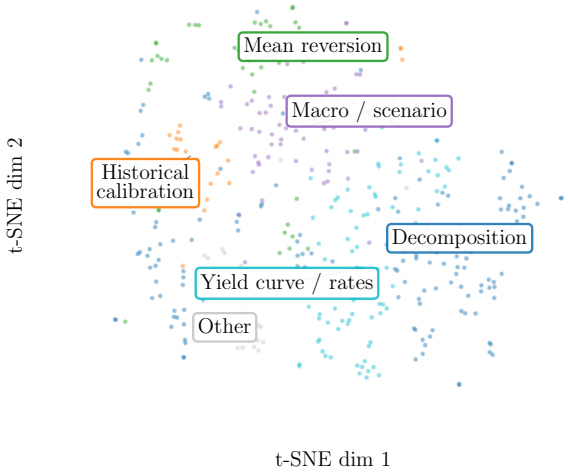


Figure A11: Free-Text Modeling Assumption Embedding Spaces Across Asset Classes. This figure shows two-dimensional t-SNE projections of the distinct modeling-technique phrases collected from the non-equity sections of CMA reports, with one panel for each asset class. Each point is one technique phrase, positioned by the similarity of its meaning and coloured by the cluster to which it is assigned; labels mark the cluster centroids. Panel (a) pools the four U.S. fixed-income segments and projects the resulting 1,199 technique phrases. Panels (b)–(h) show the corresponding projections for the remaining non-equity asset classes. Sample period: 2008–2026.

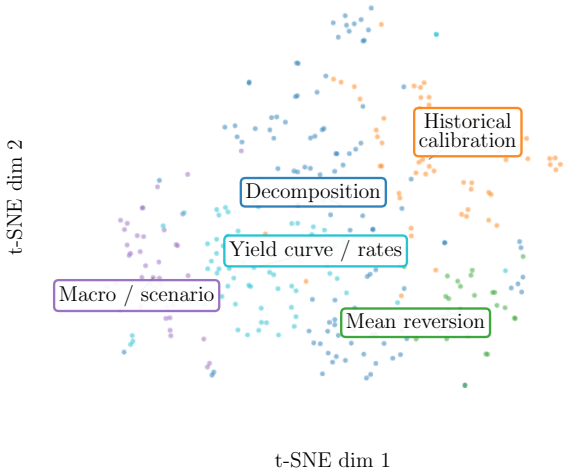
(a) Fixed Income



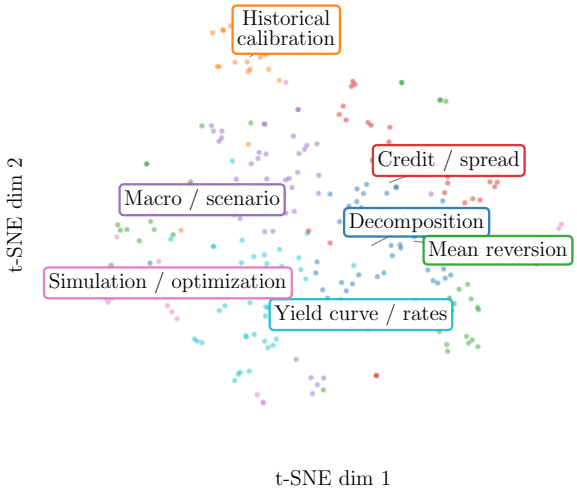
(b) U.S. Government Bond



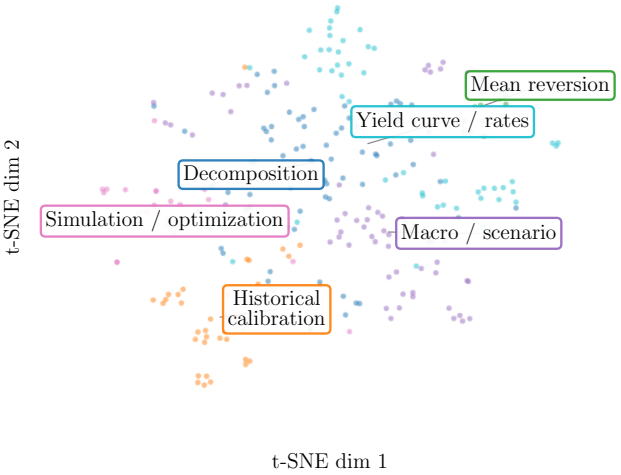
(c) U.S. Investment-Grade Corporate



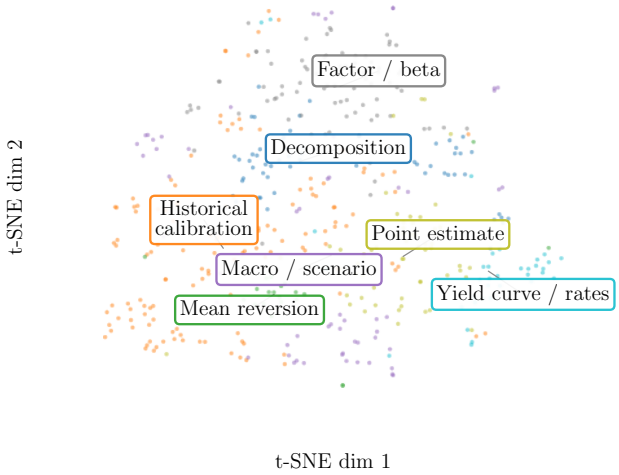
(d) U.S. High Yield



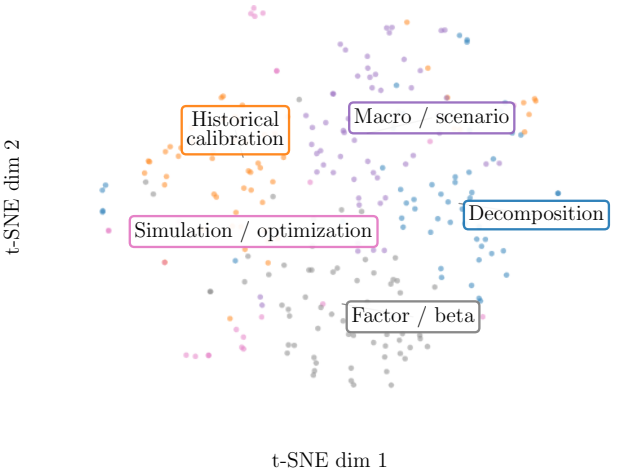
(e) Commodities



(f) Private Equity



(g) Hedge Funds



(h) Real Estate

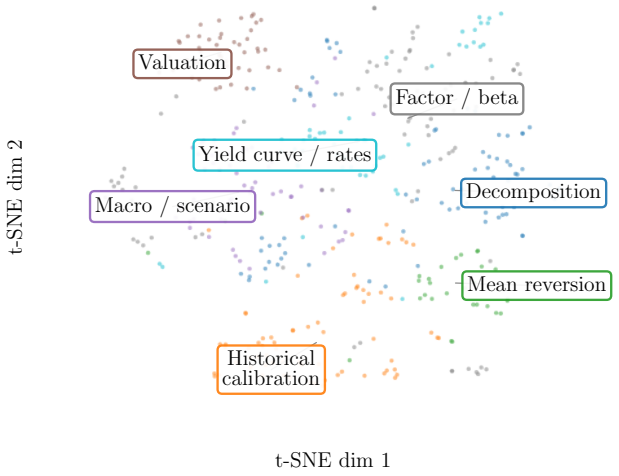


Table A1: Causal Network Metrics by Asset Manager. This table presents average number of nodes and edges per report for each asset manager in our Capital Market Assumptions causal networks dataset. Edges represent the total number of causal relationships identified, and Nodes represent the number of 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: Causal Network Complexity and Topic Attention Statistics. This table reports summary statistics for causal network measures 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 A3: Predictor variables. The 16 canonical Welch & Goyal (2008) predictors used in the total-return 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

Table A4: Diebold–Mariano Test: CMA vs. Objective Model Forecasts. This table reports Diebold–Mariano test statistics comparing the predictive accuracy of asset managers' CMA forecasts, $F^b_{i,t,h}$, with the objective machine-learning benchmark, $G^b_{t,h}$, constructed in Section A.5. For each target b, forecast errors are computed relative to the annualized realized value of that target, $Y^b_{t,t+h}$, so that $e^b_{\mathrm{CMA},i,t,h} = Y^b_{t,t+h} - F^b_{i,t,h}$ and $e^b_{\mathrm{G},t,h} = Y^b_{t,t+h} - G^b_{t,h}$, and the statistic is based on the squared-error loss differential $d^b_{i,t,h} = (e^b_{\mathrm{CMA},i,t,h})^2 - (e^b_{\mathrm{G},t,h})^2$. A positive DM statistic indicates that the objective benchmark has lower squared forecast errors than the CMA forecasts; a negative statistic indicates the reverse. Panels (a) and (b) take the total U.S. equity return as the target ($b = ER$), using our CMA reports and the CMA forecasts of Dahlquist & Ibert (2024), respectively. Panel (c) takes as targets the four building blocks of the decomposition in (2): Valuation Change (VC), Real Growth (g_{real}), Income or Dividend Yield (DY), and Inflation (π). For each horizon $h \in \{3, 5, 10\}$ in Panels (a) and (b), and for each block in Panel (c), we report the number of paired observations (N), the root mean squared error (RMSE) of the CMA and objective-model forecasts, the DM statistic, and the p-value. The pooled rows of Panels (a) and (b), and every row of Panel (c), stack forecast pairs across horizons $h \in \{3, \ldots, 10\}$. Panel (c) is stacked rather than reported horizon by horizon because a single block-horizon cell leaves too few pairs to test, and its statistics carry the small-sample correction of Harvey et al. (1997). In every panel, a forecast pair enters only once its horizon has elapsed and the realized target is observed, which is why N falls with h; report years span 2011–2025. *, **, and *** denote 10%, 5%, and 1% significance levels for the null hypothesis of equal predictive accuracy, $\mathbb{E}[d^b_{i,t,h}] = 0$.

Panel (a): CMA reports (Our Sample)

Horizon	N	$\mathrm{RMSE}_{\mathrm{CMA}}$	$\mathrm{RMSE}_{\mathrm{Obj}}$	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	$\mathrm{RMSE}_{\mathrm{CMA}}$	$\mathrm{RMSE}_{\mathrm{Obj}}$	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

Panel (c): Equity building blocks (Our Sample)

Building block	N	$\mathrm{RMSE}_{\mathrm{CMA}}$	$\mathrm{RMSE}_{\mathrm{Obj}}$	DM Stat	p-value
Valuation change (VC)	15	4.92%	3.40%	2.58**	0.022
Real growth (g)	14	2.59%	2.05%	5.55***	<0.001
Income (DY)	19	0.88%	0.82%	1.62	0.122
Inflation (π)	18	1.47%	0.89%	3.41***	0.003

Table A5: 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	Janney	Manager	2022–2026
Advent	Manager	2024–2024	Janus Henderson	Manager	2020–2025
AIMCo	Manager	2021–2023	JP Morgan	Manager	2009–2025
Allianz	Manager	2021–2025	KKR	Manager	2025–2025
American Century	Manager	2024–2025	Mackenzie	Manager	2019–2023
Amundi	Manager	2018–2025	Manulife	Manager	2020–2020
Angeles	Consultant	2018–2022	Meketa	Consultant	2020–2024
AON	Consultant	2012–2025	Merrill Lynch	Consultant	2018–2018
AQR	Manager	2013–2025	MFS	Manager	2024–2026
Baillie Gifford	Manager	2018–2018	Morgan Stanley	Consultant	2019–2025
Barclays	Manager	2021–2025	NEPC	Consultant	2016–2024
Benjamin F.E.	Manager	2024–2025	Neuberger	Manager	2022–2026
BlackRock	Manager	2019–2024	Ninety One	Manager	2023–2025
BMO	Manager	2026–2026	Northern Trust	Manager	2013–2025
BNP	Manager	2020–2025	Nuveen	Manager	2022–2025
BNY Mellon	Manager	2013–2025	OneAscent	Consultant	2024–2024
Callan	Consultant	2016–2025	PGIM	Manager	2018–2025
CalPERS	Manager	2021–2021	PIMCO	Manager	2022–2024
Candriam	Manager	2024–2025	PineBridge	Manager	2024–2025
Capital Group	Manager	2022–2025	RAM	Manager	2020–2024
CIBC	Manager	2018–2025	RBC	Manager	2022–2025
CIGAM	Manager	2024–2025	Robeco	Manager	2019–2024
Cliffwater	Consultant	2019–2026	Rockefeller	Consultant	2019–2020
Cohen Steers	Manager	2022–2024	RVK	Consultant	2021–2025
Columbia Thread.	Manager	2017–2026	Schroders	Manager	2017–2026
Credit Suisse	Manager	2020–2020	SEI	Consultant	2025–2025
Crescent Capital	Manager	2020–2024	Sellwood	Consultant	2012–2025
CWO	Manager	2019–2020	SIPA	Manager	2026–2026
Deutsche Bank	Manager	2023–2025	SJS	Consultant	2019–2020
DSA	Consultant	2011–2011	State Street	Manager	2016–2025
DWS	Manager	2017–2025	Syntrinsic	Consultant	2016–2025
Edward Jones	Manager	2020–2025	T. Rowe Price	Manager	2019–2025
EFG	Manager	2024–2024	TD Global	Manager	2022–2025
Ellwood	Consultant	2017–2017	UBS	Manager	2019–2023
Envestnet	Consultant	2013–2026	Vanguard	Manager	2014–2025
FI3	Manager	2022–2022	Verus	Consultant	2008–2025
Fidelity	Manager	2020–2025	Voya	Manager	2016–2025
Fiducient	Consultant	2021–2025	Wealthspire	Consultant	2023–2023
Franklin Temp.	Manager	2016–2025	Wellington	Manager	2021–2025
Graystone	Consultant	2021–2023	Wells Fargo	Manager	2023–2025
HSBC	Manager	2024–2025	Wilshire	Consultant	2021–2025
Invesco	Manager	2018–2025			

Table A6: Free-Text Modeling Assumption Clusters Across Asset Classes. This table defines the within-class clusters of the distinct modeling-technique phrases managers disclose, one sub-panel per asset class, plus a pooled Fixed Income panel. Clusters are labelled by their canonical methodological school and ordered by diffusion within each panel. "% reports" is the share of that panel's report-sections using the cluster at least once, and "Mentions" counts technique mentions. Sample period: 2008–2026.

Cluster	Description	% reports	Mentions
Fixed Income			
Yield curve / rates	Government bond yields via curve dynamics and duration.	90.3	358
Decomposition	Building-block decomposition into income, roll-down, valuation change, and credit losses.	81.3	322
Macro / scenario	Policy-rate and inflation forecasts from macro fundamentals.	71.0	209
Historical calibration	Risk parameters calibrated on historical data, with regime adjustments.	65.8	239
Mean reversion	Yields, spreads, and valuations reverting to equilibrium.	63.9	119
Credit / spread	Credit-return decomposition into spread, default, and recovery.	52.3	125
U.S. Government Bond			
Yield curve / rates	Yield-curve shape, term premia, and slope dynamics.	84.0	89
Fixed Income Yield Building Block	Starting yield, rolldown, price change, and income components.	82.0	69
Macro / scenario	Macro scenario forecasts anchored to consensus and breakeven inflation.	80.0	71
Valuation Based Building Block	Carry, rolldown, and valuation-change building blocks.	66.0	91
Mean reversion	Policy rates via equilibrium frameworks and Taylor rules.	52.0	49
Other	Currency and hedging-cost adjustments for cross-border bonds.	44.0	28
Historical calibration	Volatility and correlation from historical and implied data.	38.0	33
Credit Spread Default Decomposition	Credit spreads, default losses, and downgrade losses.	30.0	22
U.S. Investment-Grade Corporate			
Yield curve / rates	Term-structure and forward-rate decomposition of yields.	89.5	82
Multi Asset Building Block	Building-block decomposition across multiple asset classes.	84.2	77
Credit Spread Decomposition	Spread income, default losses, and rating-transition modeling.	73.7	45
Historical calibration	Risk parameters and premia calibrated on historical data.	71.1	64
Mean reversion	Yields and spreads reverting to long-term equilibrium.	68.4	34
Macro / scenario	Policy rates and inflation from macro fundamentals.	65.8	44
U.S. High Yield			
Building Block Return Decomposition	Benchmark yields, credit spreads, and default/downgrade losses.	80.0	33
Macro / scenario	Forward macro scenarios across multiple horizons.	72.5	49
Yield curve / rates	Yield-curve dynamics, duration, and yield-to-maturity anchoring.	70.0	55
Mean reversion	Yields and spreads reverting to long-run equilibrium.	67.5	54
Credit Spread Decomposition Modeling	Credit-spread decomposition by rating, with mean reversion.	65.0	43
Historical calibration	Historical volatility, with de-smoothing adjustments.	47.5	30
Simulation / optimization	Mean-variance optimization on the efficient frontier.	45.0	33
Credit / spread	Default- and recovery-rate loss modeling.	42.5	28
Commodities			
Factor Based Building Block Forecasting	Factor-exposure forecasts within a building-block CMA framework.	76.3	53
Building Block Return Decomposition	Additive decomposition into income, growth, valuation, and roll return.	71.1	42
Yield curve / rates	Yield curves and forward rates for fixed income and inflation-linked returns.	68.4	49
Multi Stage Scenario Simulation	Multi-stage scenario simulation blending near-term and steady-state paths.	57.9	36
Historical calibration	Volatility and correlation estimated from historical return data.	55.3	44
Mean reversion	Long-term reversion of valuation and cyclical components to equilibrium.	42.1	16

continued on next page

Table A6, continued.

Cluster	Description	% reports	Mentions
Simulation / optimization	Mean-variance optimization on the efficient frontier.	42.1	27
Commodity Macro Factor Forecasting	Commodity returns from macro drivers (GDP, inflation) and index anchoring.	34.2	26
Private Equity			
Factor / beta	Mapping to public-market proxies via beta and factor decomposition.	78.2	88
Decomposition	Building-block decomposition into beta, alpha, and risk premia.	72.7	68
Historical Premium Calibration	Risk premia and returns calibrated on historical averages.	70.9	89
Macro / scenario	Scenario-based simulation of multi-asset outcome distributions.	65.5	60
Point estimate	Long-horizon (10-year) median or best-estimate return forecasts.	65.5	59
Forward Volatility Estimation	Volatility from historical and implied measures, adjusted for smoothing.	40.0	42
Mean reversion	Reversion of valuations, rates, and spreads to equilibrium.	36.4	24
Yield curve / rates	Yield curves, credit spreads, and default-recovery adjustments.	30.9	37
Hedge Funds			
Decomposition	Decomposition into cash, equity beta, fixed-income beta, and alpha.	91.7	68
Historical calibration	Risk parameters from historical data with recency (half-life) weighting.	72.2	61
Macro / scenario	Market-variable anchoring combined with forward macro forecasts.	72.2	60
Factor / beta	Multi-factor regression on public-market exposures.	66.7	69
Simulation / optimization	Stochastic regime-switching and mean-reversion models.	38.9	24
Real Estate			
Decomposition	Decomposition into income, growth, and valuation change.	92.5	100
Factor / beta	Mapping to risk factors and benchmarks via factor models.	75.5	90
Macro / scenario	Forward macro forecasts with inflation and currency adjustments.	69.8	60
Historical calibration	Volatility and correlation estimated from historical data.	66.0	78
Yield curve / rates	Yield curves, spreads, and default/recovery term structure.	50.9	54
Valuation	Cap rates, NOI yields, and discounted cash flows.	49.1	48
Mean reversion	Stochastic reversion to long-term equilibrium levels.	39.6	35