What Drives Very Long-Run Cash Flow Growth Expectations?*

Paul H. Décaire† and Marius Guenzel‡

March 20, 2026

Abstract

We study how forecasters form beliefs about very long-run firm growth using data on terminal growth rate (TGR) expectations, textual discussions, and demographics. TGR expectations are distinct from shorter-horizon beliefs, correlated with firms' long-run growth, and uncontaminated by expected returns. Compared to short-run expectations, persistent forecaster heterogeneity carries more weight, accounting for 69% of explained variation in TGR. Local extrapolation is strongest in distant and unfamiliar contexts. TGR heterogeneity reflects differential textual emphasis on macroeconomic, industry, and geographic topics, yielding a unifying insight: when outcomes are distant and fundamental anchors are weak, individual background and experience dominate in shaping beliefs.

JEL classification: D84, G17, G24, G41.

Keywords: Long-term expectations, behavioral finance, valuation, professional forecasters, analyst identities and backgrounds, local extrapolation, textual belief content.

*First Draft: November 22, 2023. This Version: March 20, 2026. For helpful comments and suggestions, we thank Aditya Chaudhry (discussant), Francesco D'Acunto (discussant), Ricardo Delao (discussant), Jakob Sorensen (discussant), Kaushik Vasudevan (discussant), Spencer Kwon, and Laura Lindsey, as well as seminar and conference participants at the W. P. Carey School of Business at Arizona State University, the 2024 Mid-Atlantic Research Conference in Finance, the 2024 Behavioral Finance Conference at Bocconi, the 2024 CESifo Venice Summer Institute Expectation Formation, the 2024 Yale Junior Finance Conference, the 2025 American Finance Association Annual Meeting, and the Spring 2025 NBER New Developments in Long-Term Asset Management Meeting. For their excellent research assistantship, we thank Eisa Alloughani, Sebastian Gortarez Ortiz, Tej Kulkarni, Leah Li, Aditeya Pattiwar, Andre Radic, Reed Simons, Cole Streitmatter, Mingyang Zhang, and Shuman Zhang. Décaire gratefully acknowledges financial support from the Center for Responsible Investing at Arizona State University. Guenzel gratefully acknowledges financial support from the Jacobs Levy Equity Management Center for Quantitative Financial Research.

†W. P. Carey School of Business, Arizona State University, email: paul.decaire@asu.edu.

‡The Wharton School, University of Pennsylvania, email: mguenzel@wharton.upenn.edu.

1 Introduction

Long-run growth expectations are central to asset valuations. Housing prices reflect the present value of cash flows extending hundreds of years into the future (Giglio, Maggiori, and Stroebel, 2015). Long-term interest rates embed variation in distant macroeconomic fundamentals (Bauer and Rudebusch, 2020; Gormsen and Lazarus, 2026). Equity valuations respond to revisions in expected earnings growth over multi-year horizons (Bordalo et al., 2024). For individual firms, in particular, the bulk of valuation reflects cash flows far beyond standard forecast windows: over 70 percent of firm value is tied to cash flows expected beyond ten years in the future (Decaire and Graham, 2024).

Yet despite the importance of distant cash flows, little is known about how market participants form expectations about firm growth at very long horizons. As Nagel (2021) emphasizes, we lack systematic evidence on investors' long-run expectations of firm cash flows, with most available beliefs data confined to shorter forecast horizons. Because firms' current dynamics are only loosely connected to distant outcomes, validation of long-horizon expectations is slow and imprecise. Consequently, long-term beliefs are weakly anchored to near-term fundamentals. This raises a key question: do expectations about firms' very long-run growth converge across forecasters, or do they reflect persistent differences in experience and interpretation?

In this paper, we provide direct evidence on professional forecasters' *terminal growth rate (TGR)* expectations, that is, the growth rate they expect firms to sustain *in perpetuity*. We show that TGR expectations contain information distinct from shorter-horizon forecasts, robustly correlate with firms' long-run growth, and are not confounded by investors' expected returns. A central result is that forecasters' individual characteristics account for 69% of the explained variation in TGR beliefs, compared to 33% for short-run expectations. Individual heterogeneity and local extrapolation are especially pronounced when analysts evaluate geographically and culturally distant firms, where analysts' anchors to observable fundamentals tend to be weaker. Differences in forecasters' topical emphasis in qualitative discussions regarding macroeconomic, industry, and geographic conditions—the fundamental drivers of very long-run growth—primarily drive quantitative belief variation. Taken together, these results indicate that when outcomes are distant and only weakly validated by realized data, convergence in beliefs is limited and individual heterogeneity emerges as a key determinant of expectations.

To conduct the study, we combine two datasets to obtain more than 41 thousand TGR forecasts, each associated with forecasters' qualitative text discussions of their beliefs, as well as detailed forecaster demographic and other background information. The first dataset, containing the TGR expectations and text discussions, is assembled directly from original equity reports by professional forecasters (sell-side analysts), using a combination of PDF parsing functions, optical character recognition, and AI-based document processing programs. It covers forecasts for 7,975 firms across

48 countries between 2000 and 2023, and 5,229 individual forecasters.

The second dataset provides comprehensive background information on forecasters. We use personal identifiable information extracted from the equity reports (full name, employer, and real time location) to merge the TGR data with the universe of individuals included on LinkedIn obtained through Revelio Labs. Doing so, we are able to collect detailed personal information, including gender, race, age, and educational background, for 2,555 forecasters located in 45 countries.

In the first part of the paper, we establish the systematic differences of long-term TGR expectations from shorter-term expectation series, document their significant real economic information content, and show that they are not contaminated by investor expected returns.

First, TGR expectations differ sharply from the long-term growth (LTG) series in IBES, the focal longer-run beliefs measure in prior research, which is explicitly designed to reflect a three- to five-year forecasting horizon. The two series are only weakly correlated ($\rho = 0.14$), highlighting a substantial disconnect in their information content. This disconnect mirrors patterns in realized growth, where 5- or 10-year growth rates are essentially uncorrelated with short-run growth rates. Additionally, LTG expectations exceed TGR by about 10 percentage points (pp) on average, indicating that LTG captures somewhat nearer-term expectations (Bordalo et al., 2024), rather than truly long-run horizons where expected firm growth aligns more closely with economy-wide growth rates.

Furthermore, TGR expectations are strongly and positively associated with firms' future realized earnings growth. Importantly, the strong positive association persists after controlling for LTG expectations and across forecasting horizons. We show that LTG expectations do correlate strongly with shorter-term realized earnings growth (1–5 years; coefficient of 0.37), but the estimated relation weakens substantially at longer horizons (3–13 years; coefficient of 0.05). By contrast, TGR expectations are consistently and strongly associated with earnings growth across long-run horizons (1–5, 3-8, and 3–13 years; coefficients of 0.56, 0.50, and 0.35, respectively). Economically, a one percentage point increase in average TGR expectations is associated with a 0.35 percentage point rise in 10-year earnings growth, seven times the magnitude of the IBES LTG coefficient. The dominance of TGR expectations in the long-horizon regressions underscores both the limitations of prior measures and the usefulness of the new TGR measure in capturing very long-run beliefs.

In addition to being informative about realized earnings growth, TGR expectations contain a firm life cycle component, further validating their economic substance. Average TGR beliefs decline monotonically with firm age, consistent with forecasters initially facing uncertainty about whether young firms will achieve long-term success or encounter early financial distress. As firms mature and uncertainty reduces, TGR beliefs increasingly align with macroeconomic and industry trends, while firm-specific factors, such as default risk, hold greater influence for younger firms.

Finally, we address the concern that long-run growth expectations may be contaminated by reverse-engineering from prices. If forecasters infer long-term growth expectations from observed prices but apply discount rates that differ from those used by market participants, their reported beliefs could partly embed those discount rates rather than reflect independent expectations (Nagel, 2024). Building on Bordalo et al. (2024, 2025),[1] who study LTG expectations, we extend the analysis to TGR expectations and are able to provide new, direct tests for belief contamination by expected returns due to reverse-engineering, as we observe *all* components of forecasters' valuation models. Reverse-engineering implies that errors in growth expectations must be offset by errors in subjective expected returns to satisfy the price equation, generating a mechanical link between the two. We find no such relationship empirically, at odds with the idea that reverse-engineering-driven contamination of TGR expectations is prevalent among forecasters.

In the second part of the paper, we examine the determinants of variation in long-run TGR expectations and how these determinants differ from those of shorter-run expectation series, thereby informing models of belief formation.

In a first step, we compare the explanatory power of cross-sectional drivers between TGR and short-term (one-year-ahead) growth expectations using variance decompositions with firm (i.e., the firm being forecast), individual, and brokerage house fixed effects. We document significant differences. For one, these fixed effects explain at least 61% of the TGR variation and up to 39% of short-run expectations. Additionally, their relative importance differs across horizons. For short-run expectations, firm fixed effects account for the largest share of the explained variation, reaching 66% in our main specification. Time-invariant firm factors matter for TGR beliefs, as well, though to a lesser extent, accounting for 27% of the explained variation.[2] Individual forecaster fixed effects account for 33% of the explained variation in short-run beliefs and 69% in TGR beliefs, indicating that while individual heterogeneity matters at short horizons—consistent with Antoniou, Frydman, and Kilic (2025)—it becomes even more important in the long run.[3]

Motivated by the prominent role of individual forecasters in shaping long-run beliefs, we then examine two key individual-level dynamics: (i) the role of personal characteristics, backgrounds, and experiences in explaining heterogeneity in long-term belief formation dynamics across forecasters, and (ii) the role of differences in the qualitative topic discussion in equity reports, reflecting differential allocation of attention, in shaping TGR belief dispersion across forecasters.

Related to (i), we first show that forecaster observable characteristics explain 17-31% of the

[1] These papers show that changes in LTG are linked to multi-year earnings growth and returns, with LTG predicting returns even when controlling for valuation ratios and proxies for time-varying expected returns by investors. While this evidence reduces the scope for risk-premium explanations consistent with the data, Nagel (2024) concludes that it "cannot unequivocally reject the concern that analysts' reported LTG inadvertently captures risk premia."

[2] This share reflects the role of persistent firm characteristics in shaping long-run beliefs, such as industry classification or firm age. In-sample, cross-firm variation in firm age outweighs within-firm variation and is therefore largely captured by firm fixed effects.

[3] Brokerage house fixed effects, perhaps surprisingly, contribute little to the explained variation at either horizon, implying that organizational norms and protocols exert limited influence on analyst expectations.

variation in forecaster fixed effects from our variance decomposition analysis, with key factors being age, race, and geographic location. This share is significantly larger than, for example, the 3-10% of variation in beliefs attributed to demographics in Giglio et al. (2021) for wealthy U.S. retail investors.[4] Several features of our setting likely account for the greater explanatory power, including having multiple observations per forecasters across years and firms, which facilitates the fixed effect estimation, a more diverse sample that includes forecasters from multiple countries and continents, and the persistence of long-term expectations over time, which renders the predominant share of the variation cross-sectional (Decaire and Graham, 2024).

Additionally, we examine the role of forecasters' local experiences in shaping long-term beliefs and find strong evidence of local extrapolation. Forecasters' TGR expectations for firms in *other* countries are strongly associated with recent GDP growth in their own country, controlling for the firm's country GDP growth. This pattern is not merely driven by regional shocks affecting clusters of countries since the effect is concentrated among distant countries. Nor can it be explained by special linkages between forecaster and firm countries, since the results are robust to including firm-country*forecaster-country fixed effects. Instead, consistent with extrapolation from locally experienced conditions, forecasters are particularly prone to extrapolating from their local context and conditions when forming beliefs for geographically distant firms that are economically disconnected, linguistically different, and culturally distinct.[5]

Related to (ii), we link analysts' TGR expectations, and heterogeneity therein, to differences in the thematic focus in the text portions of equity reports. This provides a rare setting where we observe both quantitative expectations and qualitative discussions of these estimates, allowing us to open the "black box" behind the belief formation process. Specifically, we estimate a latent Dirichlet allocation (LDA) model, and investigate how forecasters allocate topic focus along both the extensive and intensive margins.[6]

This qualitative-based analysis yields two key insights. The first is that variation in quantitative TGR expectations primarily reflects differences in how forecasters prioritize overlapping topics in their discussions, rather than the selection of entirely distinct topics. We establish the greater

[4] Additional prior studies also report low explanatory power of demographics for expectation variation, often focusing on more homogeneous samples (e.g., Malmendier and Nagel, 2011; Bailey et al., 2019; Ben-David et al., 2018; Kuchler and Zafar, 2019; Das, Kuhnen, and Nagel, 2020; D'Acunto et al., 2023).

[5] Corroborating these results on cultural frictions in assessing firm-relevant information and their impact on long-run belief formation, we also find evidence that, on average, forecasters are more optimistic about firms located in countries they perceive to be more trustworthy, as proxied by European cross-country survey data on trustworthiness. This finding, which is consistent with the notion that trust fosters optimism by shaping expectations about others' competence (Jones, 1996), is robust within firm-year, effectively comparing the TGR beliefs of two forecasters from countries with different levels of trust toward the same focal firm's country.

[6] The LDA approach is preferred for our purposes, as we are primarily interested in identifying thematic structures and variations in analysts' qualitative discussions. In related work, Bastianello, Decaire, and Guenzel (2025a) use large language models (LLMs) to extract specific reasoning patterns and underlying mental models from forecasters' qualitative discussions, linking these to subjective expectations, specifically price target forecasts.

importance of intensive- over extensive-margin topic differences through pairwise comparisons of forecasters evaluating the same firm at the same time. This approach collapses the high-dimensional text data space into a single, comparable measure for each forecaster pair, offering a new lens on how qualitative information shapes expectations. The second insight emerges from linking the qualitative topic focus patterns back to our extrapolation evidence. We find that greater geographic distance between forecasters is associated with stronger topic focus disagreement on the intensive margin. Moreover, variation in TGR beliefs is particularly driven by differences in topic intensity related to geographic, macroeconomic, and industry themes–domains that play a central role in shaping long-run beliefs–rather than firm-level operations or valuation. Together, the results point to local factors as a key driver of topic emphasis and, in turn, extrapolation.

Overall, the above results provide the first comprehensive analysis of the origins and drivers of professional forecasters' very long-run cash flow expectations. The findings inform macro-finance models by highlighting the dominant role of persistent individual heterogeneity in explaining long-run beliefs, where outcomes are distant and only weakly validated by realized data. A related key model ingredient for long-run beliefs is extrapolation from locally experienced economic conditions, especially when firms operate in geographically and culturally distant environments where anchors to observable fundamentals are weaker.

In the final part of the paper, we illustrate a simple asset-pricing application of our findings. While a full asset-pricing treatment is beyond the scope of the paper, we implement a proof-of-concept exercise showing that the differences in very long-run beliefs across forecasters that we document provide a new, quantitatively realistic channel for explaining valuation disagreements (i.e., valuation dispersion among forecasters evaluating the same firm at the same point in time), complementing other plausible sources such as information and learning differences. To make this point, we adapt the constant-gain learning model of Delao, Han, and Myers (2024) and calibrate an extension with small, persistent differences in forecasters' TGR beliefs. Quantitatively, even modest TGR dispersion can reproduce the same level of valuation disagreement as considerable learning differences, and does so even when heterogeneity in learning speeds is fully shut down.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 provides institutional background and introduces the data as well as methodology. Sections 4 and 5 present the results. Section 6 provides a simple calibration of the results to asset pricing as a proof of concept. Section 7 concludes.

2 Literature Review and Contributions

We make four main contributions to the literature. First, we add to the literature on investor beliefs by providing evidence on the (cross-sectional) determinants of truly long-run firm growth expectations, extending beyond near-term horizons of a few years. Research on long-run beliefs is lacking, despite the fact that expectations at long horizons are crucial for asset pricing and valuation. Mukhlynina and Nyborg (2020) and Decaire and Graham (2024) show that forecasters expect over 70% of discounted cash flow to be realized beyond a 5-10 year horizon. Relatedly, Nagel and Xu (2022) and Bordalo et al. (2024) find that long-term growth expectations, rather than short-term growth expectations, explain much of the stock market valuation cycles.[7] However, these studies rely on IBES LTG expectations (i.e., 3–5 year forecasts), consistent with a broader literature in behavioral finance and asset pricing that uses LTG as a proxy for the long run, going back to at least La Porta (1996). This paper uniquely contributes by providing a detailed analysis of market participants' beliefs, along with their heterogeneity, about firm growth at much longer, truly long-run horizons.

Second, we contribute to the rapidly growing literature on investor and analyst beliefs that draws its foundation from large datasets extracted from analyst equity reports. Decaire and Graham (2024) present numerous stylized patterns on how professionals perform valuation and provide an in-depth analysis of the underlying factors shaping how they estimate discount rates in practice. By contrast, we examine the heterogeneity in long-run expectations across forecasters and firms, leveraging variation in forecasters' backgrounds from matched LinkedIn data and the qualitative discussions in their reports to explain their quantitative beliefs. Other contemporaneous papers based on analysts' equity reports include Decaire, Sosyura, and Wittry (2024), who examine how financial professionals navigate parameter uncertainty in valuation and its connection to financial market outcomes, rather than analysts' long-run expectations, Bastianello, Decaire, and Guenzel (2025a), who extract the mental models underlying analysts' financial forecasts from equity reports (cf. footnote 6), and Bastianello, Decaire, and Guenzel (2025b), who study the forecast performance of different valuation models.

By analyzing how truly long-run, TGR expectations are shaped by variation in qualitative discussions across forecasters, we also contribute to the burgeoning literature in economics and finance that integrates quantitative with qualitative information, such as text data underlying quantitative forecasts, to shed light on economic agents' belief formation process (e.g., Kendall and Oprea (2024); Demgensky and Fritsche (2023); Bybee, Kelly, and Su (2023); Andre, Schirmer, and Wohlfart (2024); Charles and Kendall (2024); Charles and Sui (2024)).

Third, we contribute to the ongoing debate on whether cash flow expectations about firms are influenced by current prices. Reverse-engineering beliefs from prices could result in misattributing price variation to beliefs rather than to time-varying expected returns, as summarized in Nagel (2024). In the context of IBES LTG expectations, Chaudhry (2025) argues that instrumented non-news

[7] Nagel (2024) explains the differences in conclusions regarding the relative importance of short- and long-run cash flow expectations (e.g., Delao and Myers (2021)) by pointing to variations in the definition of earnings used and the degree of noise introduced by the inclusion or exclusion of special items.

fluctuations in stock prices affect analyst expectations, whereas Bordalo et al. (2024, 2025) conduct various tests to refute the notion of spurious correlations between LTG beliefs and prices, including adding controls for expected return proxies and price ratios. Our contribution lies in the unique advantage of our data, which allow us to observe all components of forecasters' belief formation models. This, in turn, enables us to derive and test novel predictions of the reverse-engineering hypothesis, namely, that subjective expected returns should exhibit offsetting errors with TGR beliefs to satisfy the price equation. Empirically, we find no support for this prediction, consistent with TGR expectations reflecting true subjective beliefs without price contamination.

Finally, we contribute to the literature on the macro-finance belief formation process of investors and financial market participants. In recent years, a robust body of work has successfully used surveys to measure subjective expectations of various economic agents across different settings, including beliefs about inflation and the stock market (e.g., Coibion and Gorodnichenko (2012, 2015), Giglio et al. (2021), Delao and Myers (2021, 2024), Weber et al. (2022), Delao, Han, and Myers (2024), D'Acunto and Weber (2024), and Cui, Delao, and Myers (2025)). Complementing this evidence, recent work shows that forming accurate long-horizon expectations requires agents to infer and apply complex predictive models, and that limitations in this process generate substantial heterogeneity in beliefs even in controlled settings (Kendall and Oprea, 2024). We extend this literature by incorporating self-reported belief data, along with textual explanations of those beliefs, from forecaster-generated documents (equity reports) in the field. By documenting the importance of individual heterogeneity in forecasters' very long-run firm growth expectations, we also contribute to prior work on the effect of personal characteristics, experiences, extrapolation, and trust on beliefs, learning, as well as economic outlooks and decisions (e.g., Malmendier and Nagel, 2011, 2016; Bailey et al., 2019; Ben-David et al., 2018; Kuchler and Zafar, 2019; Hartzmark, Hirshman, and Imas, 2021; D'Acunto et al., 2021; D'Acunto, Ghosh, and Rossi, 2022; D'Acunto, Xie, and Yao, 2022; D'Acunto et al., 2023). Relatedly, by focusing on biases in forecasts by individual analysts, we also contribute to the literature in behavioral corporate finance, particularly the less-developed strand that studies non-investors and non-managers (see Malmendier (2018) and Guenzel and Malmendier (2020) for recent surveys of the field).

3 Institutional Details and Data

3.1 Institutional Details

Since we assemble the long-run cash flow expectations dataset from equity reports published by professional analysts (i.e., from the original documents), it is useful to first provide some background on the analytical practices in the equity research industry. Among the methods used to perform valuation and derive 12-month price target forecasts (the main quantitative output of equity reports), financial textbooks (Berk and DeMarzo, 2019; Brealey, Myers, and Allen, 2022) and business school curricula emphasize discounted cash flow models (DCF) and valuation multiples methods. Echoing these recommendations, these two approaches are the most common tools used by financial professionals to perform valuation (Graham and Harvey, 2001; Mukhlynina and Nyborg, 2020; Bastianello, Decaire, and Guenzel, 2025a,b), with one or both strategies included in close to 100% of equity reports housed by our data provider, Refinitiv (see Section 3.2.1). In the universe of equity reports on Refinitiv, 62% of reports use multiples to derive price target forecasts (32% for price-to-earnings ratio as the most common standalone valuation ratio), and 43% use DCF (Bastianello, Decaire, and Guenzel, 2025a,b). Our data collection specifically focuses on reports produced with DCF, for a key reason. In contrast to valuation multiples that only disclose a single metric (i.e., the multiple), DCFs require forecasters to report detailed modeling inputs, allowing us to cleanly extract and distinguish TGRs beliefs from the other parameters of the valuation model. Appendix Figure B.1 presents an example of a complete DCF model from an equity report, with the numbers used in the model redacted for copyright reasons. This granularity of the data guarantees that we never have to estimate or approximate any of the terminal growth rate expectations. Instead, our analysis is conducted entirely with data directly observed in the analyst reports. This contrasts with prior literature that relies on commercial databases such as IBES or Value Line, which provide only shorter-horizon expectations, forcing previous studies to use imprecise proxies for truly long-run beliefs in their analyses.[8,9]

3.2 Data

3.2.1 TGR Beliefs Data from Analyst Equity Reports

To construct the long-run cash flow expectations dataset, we initially downloaded 157,549 equity reports mentioning "DCF" from 42 major equity research firms via Refinitiv Eikon. We restricted the time window to reports published in the first quarter of each calendar year (January 1 to April 1) from 2000 to 2023, to ensure that belief data and other variables were systematically measured at a consistent time each year. When analysts published multiple reports on the same firm within the first quarter of a given year, we retained only the earliest publication to avoid duplicates within an analyst-firm-year pair. This procedure resulted in 41,070 reports that (1) contained the forecaster's

[8] As we discuss in Section 4.1, the "long-term" growth estimates reported in these databases (e.g., LTG in IBES) do not reflect analysts' truly long-term expectations, as they are an order of magnitude larger than actual TGR expectations.

[9] Of course, analysts do not randomly choose between multiples- and DCF-based methods. Prior work shows that DCF models and cash-flow–based valuation inputs are used more frequently for harder-to-value firms (Asquith, Mikhail, and Au, 2005; Hashim and Strong, 2018). Recent work using the near-universe of analyst reports shows that valuation methods and underlying reasoning are systematic, and that, even with tight controls, DCF-based valuations are on average less accurate than simpler approaches, though they outperform when used by sufficiently skilled analysts (Bastianello, Decaire, and Guenzel (2025a,b)). While our sample is thus selected along some dimensions, focusing on DCFs allows us to transparently measure long-run beliefs that are unobservable in multiples-based valuations.

TGR expectation and, crucially for our analysis of heterogeneity in long-run expectations across forecasters, (2) allowed us to unambiguously identify both the analyst's name and location at the time of report publication, using landline country indicators for the latter.

We then extract forecasters' TGR expectations (as well as additional relevant variables such as subjective discount rates) from each report in five steps. First, we pre-process each document using a series of Python programs, including PDF extraction functions and OCR packages, to identify all sections of text, tables, and figures that contain relevant quantitative information for our purposes. Second, we convert these different elements into short text snippets. Third, for each collected variable, we use the Open AI API to extract the numerical value from these snippets. Fourth, to ensure the highest level of accuracy, we export the text snippets and AI-extracted numerical values to Excel and manually verify each extracted value with the help of our research team. Finally, we trim the TGR expectations at the 0.5th and 99.5th percentiles to mitigate the influence of outliers and any residual transcription errors.

3.2.2 Analyst Backgrounds from Universe of LinkedIn Data

Next, we merge the TGR dataset from Section 3.2.1 with the universe of LinkedIn users obtained from Revelio Labs, to obtain additional (self-reported) personal characteristics of forecasters, such as their educational background and (approximate) age. This merge is feasible because the TGR dataset contains key identifiable information on forecasting analysts, including their full name and employer, enabling a clean and unambiguous match with the Revelio data.

In linking the datasets, we require precise name matches between individuals in the TGR and Revelio datasets, hand-verifying individuals to account for common name abbreviations. To further ensure the accuracy of our matched dataset, we also require that the LinkedIn-provided employment history aligns with the analyst's employer in the equity reports.[10] Additionally, the name-employer match must result in a unique LinkedIn within the dataset. That is, we exclude cases where multiple LinkedIn users with the same name and and the same employer are linked to the TGR dataset.

3.2.3 Summary Statistics

Table 1 provides summary statistics.

Firms and Coverage (Panel A). Our final sample consists of 41,070 equity reports with extracted information on TGR expectations, from analysts employed by 42 different brokerage houses, covering a total of 7,975 firms. The sample includes expectations data from 5,229 unique forecasting analysts, 2,555 (49%) of whom are successfully matched to LinkedIn. (For forecasters

[10] We account for common variations in employer names (e.g., Bank of America vs. BofA) as well as accented characters (e.g., Societe Generale vs. Société Générale).

who provide at least five data points in our sample, we take an extra step, including additional manual checking, to link forecasters to LinkedIn, allowing us to match 62% of individual observations in the sample.)

In additional unreported statistics, the average forecasted firm owns assets with a book value of USD 15.8 billion and has an investment rate of 5.55%. These numbers are similar to the IBES universe, where the average firm has assets with a book value of USD 15.2 billion, and an investment rate of 5.3%, suggesting that our sample is comprehensive, representative, and comparable to its commercial counterparts. Our sample includes two dozen NAICS industry sectors (2-digit), with the eight largest broad sectors accounting for 86% of the total coverage: 35% for manufacturing (NAICS 31-32-33), 19% for information (NAICS 51), 8% for professional services (NAICS 54), 7% for retail trades (NAICS 44-45), 4% for mining and oil & gas (NAICS 12), 4% for transportation (NAICS 48-49), 5% for utilities (NAICS 22), and 4% for finance and insurance (NAICS 52).

TGR Expectations (Panel B). The average subjective TGR belief in the sample is 2.2%, with the 25th-percentile estimate at 1.5% and the 75th-percentile estimate at 3.0%. By contrast, the average IBES LTG consensus forecast is 12.66%, i.e., an order of magnitude larger.[11] We defer a more in-depth discussion contrasting TGR and LTG expectations to Section 4.1.

Analyst and Country Statistics (Panel C). Regarding individual forecaster characteristics, just above 15% of analysts are female. The majority of analysts are Caucasian (63%) per Revelio's race prediction algorithm. The next largest group is Asian analysts (20%). 27% (21%) of analysts have a Master's (MBA) degree. There is considerable variation in analyst age and birth years, with the 25th-percentile analyst born in 1971 and the 75th-percentile analyst born in 1985.

The average analyst in the sample contributes eight reports and TGR forecasts. There are approximately 110 analysts on average who make forecasts for firms located in a given country, and about 140 analysts per brokerage house.

In unreported statistics, 37% of the reports are produced by analysts based in Europe, 30% in North America (27% from the US), 18% in Asia, 12% in Oceania, 2% in South America, and 1% in Africa. Additionally, 39% of firms have their headquarters located in Europe, 27% in North America (24% in the US), 19% in Asia, 12% in Oceania, 2% in South America, and 1% in Africa (also unreported). As a result of the broad geographic coverage in our sample, about 1 in 3 equity reports and TGR forecasts come from foreign forecasters, i.e., analysts who forecast the long-run growth of a firm headquartered in a different country than where they are located. Finally, there are approximately 860 equity reports and TGR forecasts for firms located in a given country.

Economic Data Series (Panels D and E). Panels D and E present descriptive statistics on various economic series for the countries in which forecasted firms and analysts are located. There

[11] We obtain LTG forecasts from Refinitiv (data item code TR.LTGMedian). We trim the IBES LTG variable at the 0.5th and 99.5th percentiles, consistent with the approach used for the TGR expectations.

is meaningful variation in economic conditions of both firm and analyst countries, both in the full sample (Panel D) and when restricting to the subsample consisting of foreign equity reports (Panel E). For example, the interquartile range (IQR) in firm-country real GDP and inflation in Panel D is 2.2pp and 1.9pp, respectively. Similarly, the IQR in analyst-country GDP and inflation in Panel E is 1.7pp and 1.4pp, respectively.

4 The Relevance of TGR Expectations

This section establishes the economic significance of TGR expectations, reflecting forecasters' truly long-term cash flow expectations. First, we demonstrate that TGR expectations provide distinct economic insights beyond those captured by other expectation series used in prior research. We show that TGR expectations exhibit limited overlap with the traditional long-term expectation benchmark, IBES LTG. Additionally, TGR expectations are strongly associated with ex-post realized long-term firm earnings growth, whereas IBES LTG (corresponding to analysts' 3- to 5-year expectations) better captures medium-term dynamics. Finally, we present novel evidence that analyst expectations (i.e., TGR expectations in our case) are not contaminated by investors' expected returns.

4.1 Properties of TGR and IBES LTG Expectations

As alluded to in Section 3.2.3, truly long-run TGR expectations and IBES LTG expectations only share limited similarities. Comparing the levels of both series reveals a systematic and persistent gap, averaging 10.45 percentage points throughout the sample period. This indicates that TGR expectations are substantially lower than IBES LTG expectations (Table 1; see also Panel A of Figure 1).[12]

Additionally, the correlation between TGR beliefs in our sample and IBES' consensus LTG expectation in a given firm-year is, while statistically significant, economically weak ($\rho = 0.14$; see Panel B of Table 1). Panel B of Figure 1 presents a scatter plot visualizing the relationship between TGR beliefs and IBES LTG expectations, further highlighting their weak and scattered association. It is difficult to visually discern a clear relationship between both expectation series, reflecting their considerable dispersion. The figure also shows that IBES LTG exhibits substantially greater volatility than TGR expectations. Indeed, the standard deviation of IBES LTG (13.9pp) is 11 times greater than that of TGR expectations rate (1.3 pp; see Table 1).

Altogether, the above evidence establishes that TGR expectations differ systematically from other expectation series that capture growth beyond a one- or two-year horizon, such as the IBES

[12] Relatedly, Appendix Figure B.2 shows that forecasters' long-term TGR stock growth expectations are systematically below historically realized long-term growth rates in every year of our sample, whereas LTG expectations examined in prior work are systematically above historical growth rates (Nagel and Xu, 2022; Bordalo et al., 2024).

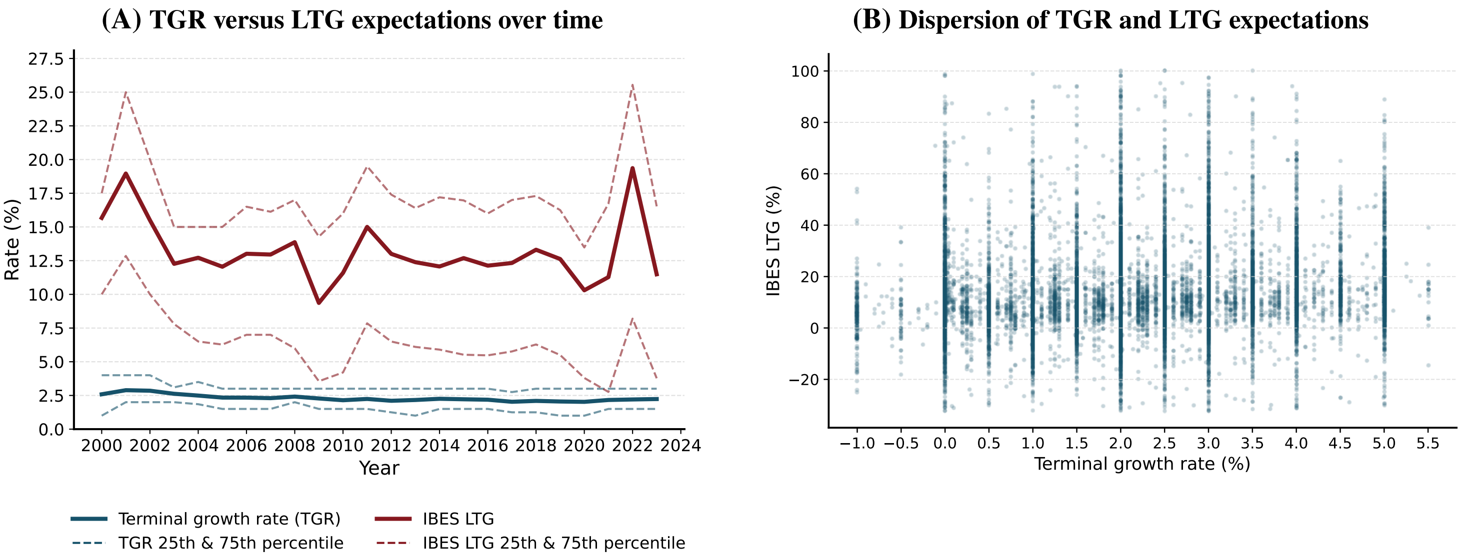


Figure 1: TGR versus IBES LTG expectations Panel A compares the levels and time trends of the TGR expectations in our sample and IBES' LTG expectations. The x-axis shows years and the y-axis is expressed in percent. Panel B shows the dispersion in TGR and LTG expectations. The x-axis represents analysts' TGR expectations in our sample, and the y-axis corresponds to the firm-year LTG consensus (median) in IBES, with each blue dot representing one TGR-LTG-consensus pair.

LTG series. As a direct implication of this, our understanding of how professional forecasters and market participants form long-term expectations about firms' growth potential remains limited.

4.2 TGR Expectations and Future Realized Firm Earnings Growth

To directly establish TGR beliefs as a meaningful measure of very long-run growth expectations, and further distinguish them from LTG expectations, we next analyze how the two series relate to firms' future (i.e., ex-post) realized long-term growth at various horizons.

Figure 2 illustrates the average association between TGR expectations and annualized ex-post realized earnings growth from year $t+3$ to $t+8$ over time.[13] The ($t+3$ to $t+8$)-window is our preferred horizon given that the modal DCF valuation model—which the TGR expectations in our sample are an input to—sets the explicit forecast horizon to three years and applies a constant terminal growth rate estimate to any subsequent cash flows afterward. Comparing the TGR beliefs with this ex-post measure of realized growth in the figure, we find that both series strongly comove. Both TGR beliefs and ex-post firm growth peaked in the early 2000s and then declined steadily, exhibiting a secular downward trend. We also note differences in the levels between TGR beliefs and realized growth rates (TGR expectations − ex-post realized growth = −3.2 pp).

[13] To clarify notation, we use $t+3$ to $t+8$ to refer to the period from the beginning of year $t+3$ through the beginning of year $t+8$. Similarly, we will later refer to the period from the beginning of year t to the beginning of year $t+5$ as (t to $t+5$), and so on.

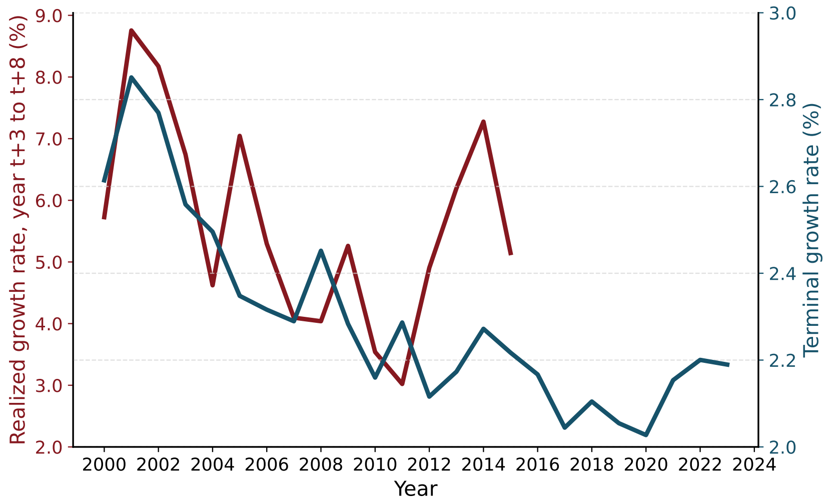


Figure 2: Terminal growth rate (TGR) expectations and realized growth. This figure plots average analysts' terminal growth rate (TGR) expectations alongside ex-post realized growth rates over the sample period from 2000 to 2023. Realized growth corresponds to the geometric average earnings growth realized from $t + 3$ to $t + 8$, averaged by year t. TGR expectations are averaged across analysts and firms by year.

In light of this, and to formalize the patterns in Figure 2, Panel A of Table 2 estimates linear regressions considering three different ways to measure ex-post long-run realized growth: the average geometric earnings growth for each firm in our sample over the periods t to $t+5$, $t+3$ to $t+8$ (as previously), and $t + 3$ to $t + 13$. The last window expands the horizon over which we measure realized growth to a 10-year period, trading off a reduced sample size for a more comprehensive measure of long-run firm growth.[14] Additionally, the formal regression analysis allows us to include firm-country*industry*year fixed effects (in the even columns), as well as the firm-year consensus LTG expectation variable from IBES, thus effectively estimating horse race regressions between TGR and LTG expectations at various horizons.

Focusing first on our preferred specifications in Columns (3) and (4), measuring realized earnings growth over the $t+3$ to $t+8$ horizon, we find that a one percentage point increase in TGR expectations is associated with a 0.50–0.70 percentage point increase in ex-post realized growth, depending on whether we include or exclude fixed effects. This strong, positive relationship between TGR beliefs and realized firm growth persists across the other examined horizons, as shown in Columns (1) to (2) and Columns (5) to (6). For instance, in the specifications with fixed effects, a one percentage point increase in TGR expectations is associated with a 0.56 percentage point increase in realized growth over the next five years, and with a 0.35 percentage point increase over the $t + 3$ to $t + 13$ horizon. The robustness and persistence of the positive relationship between average TGR expectations and long-run firm growth support the interpretation that TGR estimates are informative about firms'

[14] As Figure B.3 shows, firms' realized earnings growth rates tend to stabilize relatively quickly as they mature. Therefore, examining realized growth rates up to 13 years out captures a significant component of long-run firm growth.

growth trajectory in the (very) long run, establishing their economic substance and significance. Corroborating this conclusion, Table B.1 confirms that TGR expectations account for a significant share of ex-post realized growth in the aggregate, with an R^2 of 0.46.

We next focus on the comparison between TGR and LTG expectations. At IBES' stated forecast horizon for LTG beliefs of 3–5 years (Bordalo et al., 2024), we find a meaningful economic relationship between consensus LTG expectations and realized firm growth. A one percentage point increase in LTG expectations is associated with a 0.36–0.37 percentage point increase in average realized earnings growth over the $t + 1$ to $t + 5$ forecasting horizon. However, in contrast to TGR expectations, the relationship between LTG and realized firm growth weakens substantially and monotonically as we extend the forecasting horizon. At the $t + 3$ to $t + 8$ horizon, the association declines to just 0.10 percentage points, and at the $t+3$ to $t+13$ horizon, it further declines to just 0.05–0.06 percentage points. This pattern reveals a progressive attenuation of LTG forecast coefficients as the horizon lengthens, suggesting that long-term growth expectations become increasingly decoupled from subsequently realized earnings growth at more distant horizons. Consistent with this, in the aggregate time series regressions, LTG beliefs insignificantly predict realized growth, and the R^2 of improves only marginally to 0.49 (compared to 0.46 with only TGR expectations as a regressor; Table B.1).

In sum, both TGR and LTG are associated with long-run realized earnings growth, but their relationships evolve differently as the forecasting horizon extends. This confirms that the information content of each variable is concentrated on different time horizons of firm growth, with only TGR expectations effectively capturing long-run growth beyond the first few years.[15]

4.3 Reverse-Engineering of TGR Expectations from Prices and Return Contamination

Before we turn to studying the drivers of TGR expectations in Section 5, we first address one potential remaining concern: that analysts may reverse-engineer expectations (i.e., TGR expectations in our case) from prices—a possibility that has recently been the focus of active debate in the literature (c.f. Nagel, 2024). Our granular data allow us to make substantive progress on this issue and provide novel evidence against reverse-engineering.

Reverse-Engineering. The core idea behind reverse-engineering is that analysts do not report their own expectations in equity reports, but instead infer investors' beliefs from observed prices. That is, they report high expectations when prices are high, and low expectations when prices are low. Of course, this approach is not inherently problematic, if forecasters accurately extract investors' true long-term beliefs. In that case, it would be equivalent to forecasters independently forming

[15] This disconnect between shorter-run and truly long-run expectations is mirrored in realized growth patterns, where the correlation between current and historical growth rates steadily declines as the time gap between them increases (Figure B.4).

long-term beliefs using a valuation model aligned with that of investors. Under either scenario, extracted TGR estimates would convey meaningful insights into the long-run firm growth beliefs of key financial agents.

The underlying concern, instead, is that investors and forecasters may, unaccounted for by the econometrician, use different valuation models or inputs. Precisely, if forecasters report beliefs (e.g., TGR or other beliefs) inferred from current prices but use discount rates that differ from those used by investors, the resulting reported beliefs may be contaminated by investors' returns expectations. In the most extreme case—where forecasters apply a constant discount rate while investors use a time-varying rate with constant growth expectations—a high price today reflects low expected returns, but forecasters will mistakenly infer and report high growth beliefs, and vice versa in the case of a low price today.

Importantly, the most intuitive empirical test for reverse-engineering proves uninformative. Prior literature has documented a negative correlation between longer-term expectations and future realized returns (e.g., Bordalo et al. (2024)). In fact, we replicate this negative relationship for TGR expectations in Panel A of Table 3, whereas the existing literature has focused on LTG beliefs.[16] While these patterns are consistent with forecasters stating high beliefs when investor expected (and subsequently realized) returns are low, they are *equally* consistent with a world in which forecasters and investors use the same valuation model and inputs, but systematically overstate or understate firms' growth prospects in different periods. In that case, miscalibrated growth beliefs would generate high (or low) prices today and, correspondingly, low (or high) realized returns in the future.

Prior Evidence. Prior work has advanced two lines of argument to suggest that analyst beliefs do not spuriously capture investors' expected returns. The first shows that analyst expectations continue to predict realized returns even after controlling for proxies of time-varying required returns, narrowing the scope for misattribution or contamination of beliefs by return variation. Recent work by Bordalo et al. (2024), for instance, finds that LTG expectations retain their negative relationship with realized returns even after controlling for the surplus consumption and consumption-wealth ratios in Campbell and Cochrane (1999) and Lettau and Ludvigson (2001), respectively, as well as the SVIX^2 index in Martin (2017).

The second argument provides evidence that investors and forecasters rely on similar frameworks when forming beliefs about asset prices. A large body of work suggests that investors attribute price variation, at least in part, to changes in expected returns and expected earnings growth. Consistent with this, fundamentals such as dividends, earnings, and cash flows help predict returns in market

[16] The negative relation between TGR expectations and realized returns obtains both in specifications without fixed effects and in models including firm and firm HQ country×year fixed effects. The result holds when controlling for analysts' subjective discount rate, the effective input used in DCF models to compute the present value of cash flows, and when only using the components of the subjective cost of equity as an alternative predictor of realized returns (Decaire and Graham, 2024).

data (Golez, 2014; Sabbatucci, 2022; Hillenbrand and McCarthy, 2024; Pruitt, 2024). Mirroring this, analysts' subjective expected returns and expected earnings growth also inform changes in their subjective valuations (Decaire and Graham, 2024) *and* market prices (Chen, Da, and Zhao, 2013; Delao and Myers, 2021). A wide range of studies further documents that the capital asset pricing model (CAPM) is one of the most commonly used tools for estimating expected equity returns by *both* investors (Guercio and Tkac, 2002; Barber, Huang, and Odean, 2016; Berk and van Binsbergen, 2016) and forecasters (Brav, Lehavy, and Michaely, 2005; Andrei, Cujean, and Wilson, 2023; Decaire and Graham, 2024). Relatedly, Decaire and Graham (2024) show that forecasters who report TGR estimates within DCF frameworks apply time-varying discount rates—contrasting with the extreme possibility of reverse-engineering above in which analysts use a constant discount rate while investors use a time-varying one with fixed growth expectations. Finally, recent evidence shows that analysts' discount rate estimates track realized future returns on average (Decaire and Graham, 2024), implying that these inputs plausibly reflect the key features of investor return expectations as well.

New Empirical Tests. Nonetheless, the extent to which analysts engage in belief reverse-engineering remains debated (Nagel, 2024). To make further progress, we first lay out the valuation framework used by investors and analysts, which in turn allows us to formalize the reverse-engineering concern and derive new empirical tests.

We begin with the valuation model implicit in market prices. Let p_t denote the observed market price of the firm. Investor beliefs satisfy:

$$p_t - d_t = E \sum_{j=1}^{N} \rho^j \Delta d_{t+1+j} + \frac{\rho^N}{\rho} E\, TGR - \frac{1}{\rho} r_t, \tag{1}$$

where r_t denotes the discount rate implicit in market prices.

Analysts producing TGR estimates in equity reports rely on the same two-part short- and long-run valuation structure when forming price targets, but rely on subjective beliefs and discount rates:

$$p_t^a - d_t = E^* \sum_{j=1}^{N} \rho^j \Delta d_{t+1+j} + \frac{\rho^N}{\rho} E^* TGR - \frac{1}{\rho} r_t^a. \tag{2}$$

Here, p_t^a denotes the analyst's subjective valuation of the firm, and r_t^a is the analyst's subjective, time-varying discount rate.[17] The term $E^* \sum_{j=1}^{N} \rho^j \Delta d_{t+1+j}$ captures the analyst's forecast of discrete earnings growth over the next N periods, while from period $N+1$ onward valuation depends on the

[17] Decaire and Graham (2024) show that, in practice, analysts typically employ a single discount rate to discount all projected cash flows within a given DCF model. However, this discount rate is not fixed: it evolves as analysts revise and update their valuation models over time.

analyst's long-run growth expectation, E^*TGR, where the superscript * denotes subjective beliefs.

Working from equation 2, the key reverse-engineering concern is that analysts do not form beliefs independently, but instead start from the observed market price, p_t, implying $p_t^a = p_t$. Moreover, analysts may solve the investor pricing identity using a discount rate that differs from the discount rate implicit in market prices, $r_t^a \neq r_t$. Let $\epsilon \equiv r_t^a - r_t$ denote the discount-rate wedge induced by this misspecification. Because prices are taken as given, this wedge must be absorbed by analysts' growth beliefs.

Allowing this adjustment to be allocated across short- and long-run expectations, with fraction ϕ loading onto near-term growth beliefs and $1 - \phi$ onto terminal growth, the observed price can be written as:

$$p_t - d_t = \underbrace{\left(E\sum_{j=1}^{N}\rho^j \Delta d_{t+1+j} \; + \; \phi \cdot \frac{1}{\rho}\epsilon\right)}_{\text{implied short-run beliefs}} + \frac{\rho^N}{\rho}\underbrace{\left(E\,TGR \; + \; (1-\phi)\cdot\frac{1}{\rho^N}\epsilon\right)}_{=\,E^*TGR \text{ (implied long-run beliefs)}} - \underbrace{\frac{1}{\rho}\Big(r_t + \epsilon\Big)}_{=\,\frac{1}{\rho}r_t^a \text{ (analyst discount rate)}} . \tag{3}$$

Two scenarios are possible: (1) all contamination from using a different discount rate is contained in short-term growth expectations ($\phi = 1$), or (2) the contamination affects long-term TGR expectations as well ($\phi \neq 1$). While Case (1) does not pose a threat to our analysis of TGR expectations, Case (2) would be problematic.

Examining Case (2) more closely, ϵ implies a direct mechanical relationship between a given analyst's discount rate and their TGR estimate. That is, even after controlling for macroeconomic variables likely to affect both r_t^a and TGR beliefs (e.g., GDP, inflation, and interest rates), a residual comovement between the two variables should persist, driven mechanically by ϵ. This yields a new testable implication that we can implement by leveraging the granularity of our data: the absence of comovement between analyst's observed subjective discount rate and TGR beliefs (after controlling for macroeconomic factors likely to jointly drive variation in both variables) would provide evidence against a meaningful role for ϵ, and thus against contamination due to reverse-engineering.

Panel B of Table 3 implements this test, assessing whether the mechanical relation exists in the data by relating analysts' subjective discount rates (WACC) to their subjective TGR expectations:

$$E^*TGR_{i,j,t} = \beta r_{i,j,t}^a + \mu_{i,j,t}, \tag{4}$$

where $\beta = \frac{\text{cov}(E^*TGR_{i,j,t}, r_{i,j,t}^a)}{\text{var}(r_{i,j,t}^a)}$ reflects the total covariance between subjective TGR expectations and discount rates. This covariance may arise from genuine economic comovement (for example, through common exposure to macroeconomic conditions) or mechanically from contamination induced by the discount-rate wedge described in equation (3). To evaluate the reverse-engineering channel, we isolate the covariance component generated purely by the wedge, i.e., the portion driven by variation in $\epsilon_{i,j,t}$. This implies:

$$\text{cov}((1-\phi)\cdot\frac{1}{\rho^N}\epsilon_{i,j,t},\epsilon_{i,j,t}) = (1-\phi)\cdot\frac{1}{\rho^N}\,\text{var}(\epsilon_{i,j,t}). \tag{5}$$

Implementing equation 4 to estimate the mechanical component in equation 5, Column (1) of Panel B in Table 3 includes no fixed effects and yields a small, positive, and statistically significant coefficient. This estimate reflects the total covariance between TGR expectations and subjective discount rates, consistent with both mechanical contamination and common macroeconomic drivers (e.g., interest rates). Column (2) introduces country*year fixed effects to absorb broad macroeconomic conditions and analyst*firm fixed effects to focus on within-analyst-firm variation over time. This specification isolates the level at which reverse-engineering would occur (equation 3) and removes common macroeconomic variation that could otherwise generate non-mechanical covariance. In this specification the estimated coefficient is effectively zero and statistically insignificant at conventional levels (p-value > 0.10).[18,19] Finally, while the discount rate represents the relevant input in DCF valuation models, Columns (3) and (4) replicate the specifications of Columns (1) and (2), replacing the discount rate with the subjective cost of equity for completeness, and confirm our findings.

Effectively, the results in Panel B of Table 3 allow us to rule out problematic reverse-engineering in our setting, provided two conditions hold: (i) the discount-rate wedge $\epsilon_{i,j,t}$ varies over time within analyst–firm pairs, and (ii) the allocation parameter ϕ lies between zero and one.

Regarding condition (i), if analysts applied a fixed wedge $\epsilon_{i,j}$, there would be no within-analyst–firm variation in ϵ, so that a zero coefficient would be uninformative about the presence of reverse-engineering. We view this scenario as implausible. Reverse-engineering time-varying market prices requires time variation in the wedge; a constant adjustment would effectively pin down the discount rate up to a constant and offer no incremental benefit in matching observed prices as they evolve.

[18] The results in Panel B of Table 3 are robust to expressing TGR and r_t^a in logarithmic form (i.e., $\ln(1+TGR)$ and $\ln(1+r_t^a)$) instead of in percentage terms (i.e., TGR (%) and R_t^a (%), as shown in the table). This rules out both additive and multiplicative analyst errors ϵ.

[19] In principle, one could also proxy for the discount rate implicit in prices and test the correlation between the implied wedge and an analyst's TGR estimate directly; however, our fixed-effects strategy isolates the within-analyst–firm variation relevant for mechanical reverse-engineering while absorbing common macroeconomic discount-rate movements, and avoids reliance on potentially noisy proxies for their subjective discount rate implicit in prices. Indeed, Decaire and Graham (2024) provide a detailed discussion of the distinction between the subjective discount rate, the subjective cost of equity, and expected returns in the context of analyst data, and show that extracting expected returns from analyst price targets and current market prices fails to recover the true subjective cost of equity or their subjective discount rate. Rather, it yields a measure that conflates the subjective cost of equity with analysts' assessments of market mispricing.

Regarding condition (ii), if some analysts have $\phi < 1$ while others have $\phi > 1$, then $(1-\phi)$ would take both positive and negative values. In that case, the regression estimate of β, which captures the average of $(1 - \phi_{i,j,t})$, could be close to zero due to offsetting effects, even if reverse-engineering were operative. We view such systematic cancellation as unlikely, however. Values of ϕ above one imply that analysts over-allocate the discount-rate wedge to near-term forecasts and offset this through opposite distortions in long-run growth, a pattern that lacks a clear economic rationale and, importantly, is inconsistent with the very logic of concealing those *ad hoc* adjustments: it would require the analyst to introduce larger distortions in *both* short- and long-run expectations than are strictly necessary when using a ϕ smaller than one. Parsimonious manipulation, such as adjusting as little as economically feasible, is the more natural benchmark. Moreover, if ϕ is close to one, contamination would primarily affect near-term expectations rather than long-run growth beliefs.

Considering that equation 3 predicts a direct, mechanical, and thus statistically strong relationship between subjective TGR expectations and discount rate in the presence of return contamination in long-run beliefs, the evidence in Table 3 provides direct evidence against this possibility.[20] Our tests complement and extend the findings in Bordalo et al. (2024). Our contribution is that we observe all components of forecasters' valuation models and can therefore, for the first time, analyze the direct relationship between a given analyst's beliefs and expected returns, rather than relying on proxy measures used in the literature. Overall, we conclude that analysts' long-run TGR expectations are not mechanically inferred from market prices—which does not preclude analysts from learning meaningfully from prices—and next study the determinants of TGR expectations in Section 5.

5 The Drivers of Expectations: TGR vs. Shorter-Term Expectations

Having established that (i) TGR expectations exhibit distinct economic properties compared to other longer-term expectation series studied in prior work, such as IBES LTG (Section 4.1), (ii) TGR expectations correlate with realized firm growth at longer horizons than LTG (Section 4.2), and (iii) TGR expectations are unlikely to be confounded by return contamination due to belief reverse-engineering from prices (Section 4.3), we now investigate the drivers of TGR expectations. Specifically, we conduct a variance decomposition of TGR beliefs, compare their determinants to those of shorter-run expectations, examine how the heterogeneity in TGR beliefs relates to various underlying firm and forecaster characteristics, and "open the black box" behind TGR expectations by linking them to analysts' qualitative text discussions.

[20] Instead, the evidence is consistent with two alternative interpretations: (i) analysts engage in reverse-engineering, but allocate most or all of the discount-rate wedge to short-run beliefs, leaving their subjective long-term growth expectations largely unaffected (i.e., $\phi \to 1$); or (ii) reverse-engineering is not widespread to begin with and plays a negligible role in shaping analyst expectations. In either case, long-run growth beliefs are not meaningfully contaminated.

5.1 Variance Decomposition

Decomposing TGR Expectations. Table 4 decomposes the variation in TGR expectations across brokerage houses, analysts, and forecasted firms using an ANOVA-based sequential variance decomposition (in Columns (1) to (4) of both panels A and B). Panel A estimates the decomposition on the full sample, excluding firms, analysts, and brokerage houses with fewer than 2 observations, as the respective fixed effects would fully explain the associated variation in terminal growth rate expectations. Panel B restricts the sample to firms with at least ten observations for robustness. In both panels, we present specifications with at least 10, 25, and 50 observations per forecaster, to further address potential concerns about overfitting.

We consistently find that the three sets of fixed effects explain a substantial share— between 61% and slightly above 70%—of the variation in TGR expectations. Brokerage house fixed effects account for only a very small portion of the explained variation, whereas firm fixed effects and in particular analyst fixed effects drive most of the explained variation. For example, in the largest sample in Column (1) of Panel A, brokerage house fixed effects contribute 4% of the explained variation, whereas analyst and firm fixed effects contribute 69% and 27%, respectively. Similarly, in the narrowest sample in Column (4) of Panel B, we estimate contributions of 18%, 46%, and 36%, respectively. Throughout, analyst fixed effects consistently account for the largest share of the explained variation, suggesting that differences across forecasters play a particularly important role in shaping TGR expectations.

One limitation of the approach in Table 4 is that the ANOVA-based estimation is not fully indifferent to variations in the sequential order in which the fixed effects are included. As an alternative approach, we estimate standard OLS regressions in which we gradually include the fixed effects one at a time (Table B.2). Panels A and B replicate Column (1) of Panel A and Column (4) of Panel B from Table 4, respectively, representing the two extreme cases of that table. Columns (1) to (3) present the results when including each set of fixed effects individually, effectively estimating an *upper bound* on the amount of variation explained by the inclusion of the given fixed effects. Intuitively, the included fixed effects may be correlated with other (omitted) variables that also influence TGR expectations (see the related discussion in Ayyagari, Demirgüç-Kunt, and Maksimovic, 2008). The remaining columns add various sets of fixed effects jointly.

Using the alternative OLS approach, brokerage fixed effects alone explain at most 2% of the total variation in TGR expectations in the full sample in Panel A, while analyst fixed effects alone explain up to 47%, and firm fixed effects alone account for up to 45%. Additionally, brokerage and firm fixed effects together explain up to 47% of the variation, while adding analyst fixed effects increases the incremental R^2 by 19 percentage points, bringing the total R^2 to 66%. To further ensure that these results are not driven by overfitting due to the high degrees of freedom in the fixed effects, we also examine the adjusted R^2, which penalizes model complexity. The results remain consistent:

adding analyst fixed effects to brokerage house and firm fixed effect increases the adjusted R^2 by 17 percentage points, from 38% to 55%. The conclusions from the OLS results for the narrower sample in Panel B are similar.

Given the above results that analyst fixed effects account for a large share of the variation in TGR expectations, we also test the possibility that this result merely captures differences in the forecast horizon at which terminal growth begins. For the subsample in which we are able to observe the TGR horizon from analysts' reports, we replicate the R^2 analysis including horizon fixed effects. Table B.3 shows that horizon fixed effects contribute little additional explanatory power, while analyst fixed effects remain highly informative even conditional on brokerage house and firm fixed effects in this subsample. These findings indicate that dispersion in long-run growth expectations does not simply reflect differences in TGR horizons, but instead reflects stable analyst-specific heterogeneity.

Decomposing Shorter-Term Expectations. We next extend the ANOVA variance decomposition to short-term growth expectations—IBES 1-year EPS growth forecasts—to examine how the importance of cross-sectional drivers varies across horizons. The results, presented in Columns (5) to (8) of panels A and B in Table 4, reveal notable differences. Forecaster fixed effects account for a substantially larger share of the explained variation in TGR beliefs compared to short-run expectations, contributing 18 to 52 percentage points more for TGR. At the same time, individual heterogeneity remains economically present at short horizons. Similar to the results for TGR expectations, brokerage house fixed effects explain a negligible portion of the short-run belief variation, while firm fixed effects remain a key factor. These findings suggest that individual forecaster characteristics matter for belief formation at both horizons but become substantially more important for long-run expectations.

Overall, the variance decomposition results underscore the importance of both firm factors and individual forecaster characteristics in shaping forecasters' long-run cash flow expectations. These findings motivate a deeper analysis of firm and analyst characteristics to better understand the large and persistent heterogeneity in TGR expectations across firms and forecasters.

5.2 Firm Characteristics and Theorized Drivers of Long-Run Firm Growth

A natural approach to exploring how TGR expectations relate to firm characteristics is to link cross-firm heterogeneity in TGR expectations to factors associated with the firm life cycle, given that products tend to become more standardized and opportunities for innovation and growth diminish as firms age and industries mature (e.g., Klepper, 1996).

Firm Age. We begin by relating TGR expectations to firm age. As Figure 3 shows, forecasters' TGR expectations monotonically decline with firm age. The average TGR estimate in the sample is 2.5% for young firms (aged 10 years and less) and decreases to 0.5% for very old firms, reflecting diminishing growth opportunities as firms mature. The magnitudes in the figure reflect the relationship between TGR expectations and firm age cohort *after* controlling for firm and year fixed effects.

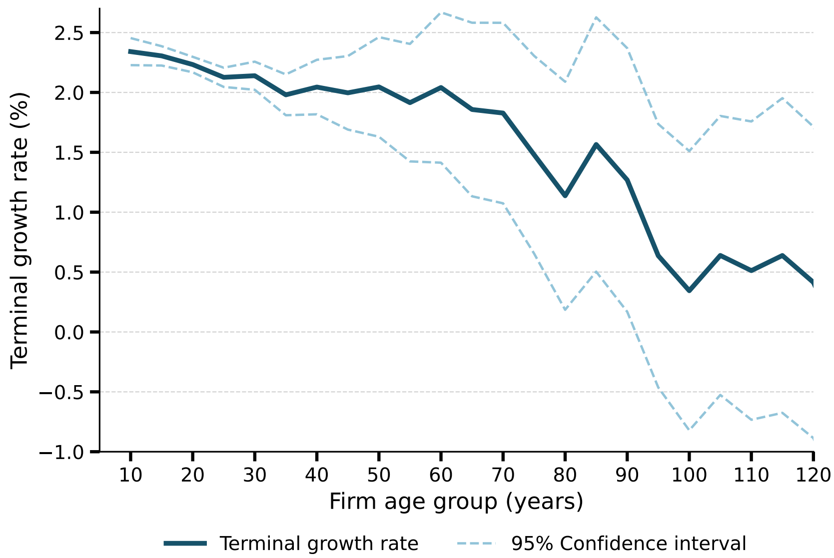


Figure 3: TGR expectations and firms' life cycle This figure plots TGR expectations over firms' life cycle. The solid blue line corresponds to age-group fixed effects obtained from the regression: $TGR_{i,j,t} = k + I_{\text{Age group}} + I_{\text{Firm}} + I_{\text{Year}} + \epsilon_{i,j,t}$, where we add back the regression constant (k) to better illustrate the terminal growth rate level. The pale blue lines correspond to 95^{th} confidence interval of the $I_{\text{Age group}}$ regression estimate, where we also add back the regression constant (k). Firm age-group captures firm age groups measured in 5-year increments such that, e.g., firm age-group = 10 corresponds to firms between 5 and 10 years old.

Industry, Country, and Firm-Specific Drivers. Motivated by this link between TGR expectations and firm age, we next examine when the explanatory power of firm characteristics is most pronounced and how the importance of firm features evolves over the firm life cycle. In Columns (1) and (2) of Table 5, we find that a firm's industry and headquarters' country explain only a small portion of TGR expectations for young firms. Specifically, these factors explain only 17% of the explained variation in very young firms aged five years or less. Even 20 to 25 years after a firm's IPO, industry and country explain only about 32% of TGR beliefs. In Columns (3) and (4), we find that in these young firms, firm fixed effects explain a large share of the variation in TGR expectations that is unaccounted for by analyst, brokerage house, industry, and firm country factors, consistent with the evidence in Section 5.1. Firm fixed effects account for up to 37% of unexplained variation in very recently IPOed firms and up to about 28% for firms aged 20 to 25 years.

These patterns in young firms reverse as firms mature. As Column (2) of Table 5 shows, in old firms, a firm's industry and country explain a significantly larger share of the variation in TGR expectations—53% for 50-year-old firms, and an even greater magnitude for firms aged 70 years and above. Accordingly, for mature firms, the incremental explanatory power of firm fixed effects over industry and country declines significantly.[21]

This evidence highlights that the role and nature of firm-specific factors in explaining TGR heterogeneity evolve systematically over the firm life cycle. We interpret these findings through the lens of a firm life-cycle framework, in which forecasters contend with uncertainty about whether a new firm will achieve long-term success or face near-term bankruptcy.[22] As forecasters gain information about a firm's long-term viability, expectations for average long-run growth increasingly converge with broader country and industry trends.

As a concrete example of a firm-specific factor that is correlated with TGR, Panel B of Table 2 presents evidence on the role of long-term default risk. To assess how varying levels of default risk impact TGR expectations for firms operating in the same country and industry during the same year, our specification includes firm-country*industry*year fixed effects. We also control for firm age (measured as the natural logarithm of the firm's post-IPO age), which enters negatively in the estimation consistent with Figure 3. Using this specification, we estimate a negative link between TGR expectations and default risk, suggesting that analysts, on average, apply a 0.21 pp discount (i.e., a 10% reduction compared to the sample average) to their TGR beliefs about firms with an above-median default probability. This average discount appears reasonable considering that the average probability of default within the high default risk group (above the median) is 5.8%.

Finally, interpreting the downward trend in TGR expectations documented in Panel B of Figure 2 and Panel B of Table 2 through this life-cycle framework suggests that forecasters' favorable beliefs about the growth potential of young firms outweighs potential negative revisions. This pattern aligns with the broader tendency toward (over-)optimism in longer-term forecasts documented in Bordalo et al. (2024).

Dispersion in TGR Expectations Over Firms' Life Cycle. As a bridging step between analyzing firm and analyst characteristics, we examine how the within-industry-year *dispersion* of TGR expectations across forecasters varies over the firm life cycle. As Table B.4 shows, dispersion in TGR beliefs is highest in young firms. As firms mature, TGR belief dispersion across forecasters substantially declines, by as much as 88% for the most mature firms. These patterns hold across specifications with and without industry*year fixed effects as well as when controlling for firm size. Moreover, a Bonferroni-adjusted test for the joint equality of all the firm age coefficients is rejected at the 1% level in all specifications.

Overall, these patterns in TGR belief dispersion suggest that disagreement is greatest when

[21] While we report magnitudes for very old firms, we caution that the sample size for these cohorts is smaller—fewer than 100 observations—possibly raising concerns about overfitting. That said, the patterns observed for older firms align with the progression seen among younger cohorts, where sample sizes average several thousand observations per age group.

[22] Nearly 25% of public firms ultimately declare bankruptcy, with most failures occurring within 10 years of their IPO. This statistic is derived from North American Compustat data spanning 1950 to August 2023 and includes firms classified as bankrupt (code 02), liquidated (code 03), or active, excluding others based on the mnemonic DLRSN.

firm-specific anchors are limited.

5.3 Forecaster Characteristics and Backgrounds

Having established the importance of forecaster fixed effects in Section 5.1 and Table 4, we next examine the extent to which analyst observables help explain the persistent individual heterogeneity, leveraging our data on forecaster identities and LinkedIn-matched background information.

Panel A of Table 6 reports the total TGR variation explained by firm, brokerage house, and analyst fixed effects, similar to the information in Table 4, across various data cuts (excluding firms, analysts, and brokerage houses with fewer than two observations, as for these observations, as their respective fixed effects fully explain the associated variation in TGR expectations; excluding cases with fewer than five observations; examining analysts with foreign coverage only; and excluding domestic coverage). Panel B presents a sequential ANOVA decomposition of individual analyst fixed effects for each of these subsamples, estimated using a mover design—i.e., tracking analysts who switch brokerage houses within our sample—to separately identify analyst and employer fixed effects, following the labor economics literature (Abowd, Kramarz, and Margolis, 1999). Forecaster observable demographics jointly explain up to 31% of the variation in analyst fixed effects—a substantially larger share than found in prior studies on retail investors (cf. Giglio et al., 2021) and households (cf. Malmendier and Nagel, 2011; Bailey et al., 2019; Ben-David et al., 2018; Kuchler and Zafar, 2019; Das, Kuhnen, and Nagel, 2020; D'Acunto et al., 2023), which typically report that demographics account for at most 10% of belief variation. Demographic factors particularly associated with analyst fixed effects in our sample include birth year (age) and location (country of residence at the time of the report). These findings remain robust across the various estimated specifications.

Several factors likely contribute to the greater explanatory power of demographics in forecaster expectations in our setting than in prior work. First, our panel structure includes multiple observations per forecaster, with most analysts forecasting multiple firms over several years, allowing for more precise estimation of forecaster fixed effects. Second, our sample is substantially more diverse, spanning analysts from many different countries, which enables us to disentangle firm- and analyst-level country effects that are difficult to separate in more homogeneous samples (e.g., affluent U.S. retail investors affiliated with Vanguard (Giglio et al., 2021)). Finally, long-run expectations are highly persistent and exhibit limited time-series variation, making them primarily cross-sectional objects (Decaire and Graham, 2024), which likely amplifies the role of individual characteristics relative to the more volatile determinants of short-term expectations.

To ensure that the conclusions derived from the variance decomposition are robust, beyond their consistency across the various data filters in Table 6, we again conduct an additional test. Specifically, we incrementally add fixed effects to OLS regressions and evaluate the corresponding increases in R^2 (Table B.5). Panels A and B replicate Columns (1) and (6) of panel B in Table 6, respectively, representing the two most distinct specifications of the analysis. Columns (1) to (5) in Table B.5 sequentially introduce each fixed effect individually. Examining the R^2 values, birth cohort (R^2 = 0.09–0.12) and analyst geographic location (R^2 = 0.08–0.12) explain a substantial share of the variation in analyst fixed effects, whereas other demographic variables have less explanatory power. Moreover, adding analyst geographic location to a model that already includes gender, education, race, and birth cohort significantly improves explanatory power, with R^2 increasing by 0.07–0.16 percentage points, representing a 70–123% improvement (Columns (6) versus (7)). Overall, these OLS-based results corroborate the ANOVA-based findings in Table 6, with age and location effects emerging as particularly important.

5.4 Forecasters' Personal Economic Environment and Extrapolative Expectations

A natural next question is why forecasters with certain backgrounds, such as their geographic exposures, systematically hold higher or lower TGR expectations. This section examines how locally experienced economic conditions shape long-run expectations, including those for geographically distant firms. As a broader takeaway, recalling the, relatively speaking, smaller role of individual forecaster effects in short-term expectations (Table 4), we might expect long-run expectations—being more abstract and less tightly anchored to near-term fundamentals—to be more influenced by variation in forecasters' experiences.

Panel A of Table 7 shows how personally and locally experienced economic conditions, and specifically recent GDP growth in the forecaster's country of residence, influence forecasters' long-term firm growth expectations, exploiting the granularity of the data on the geographic locations of both firms and the analysts covering them. In Column (1) and (2), we find that recent GDP growth is strongly correlated with TGR expectations for firms. Notably, this relationship holds even after controlling for recent GDP growth in the firm's country.

In Column (3), we find that the effects of local experiences on TGR expectations are concentrated in forecasts for firms located in geographically distant regions. This pattern is inconsistent with potential concerns that the baseline relationship in columns (1) and (2) is driven by common shocks across neighboring countries. Instead, it supports the interpretation that local extrapolation reflects analysts' responses to differences in familiarity with firms in distant locations. To further address potential confounding mechanisms, all regressions include analyst-country*firm-country fixed effects, which control for specific linkages between analysts and firms. For instance, if analysts are systematically more likely to forecast firms with operations in their home country, these fixed effects ensure that local extrapolation is assessed within such dynamics. The inclusion of these fixed effects strengthens the interpretation of our findings as evidence that local economic conditions systematically shape analysts' beliefs.[23]

To further elucidate why and how analysts engage in local extrapolation for distant firms, we next examine three factors related to familiarity and cross-country frictions in Panel B of Table 7: (i) language similarity, (ii) economic ties, and (iii) cultural affinity. At a high level, we find that local extrapolation is more pronounced when forecasters face greater barriers along these dimensions, further suggesting that differences in expectations are particularly pronounced in more abstract and distant—geographically, economically, or culturally—contexts.

Language Similarity. Columns (1) and (2) of Panel B in Table 7 split the sample based on whether the dominant language of the firm's country matches that of the analyst's. Local extrapolation is significantly more prevalent when languages differ, in line with language barriers being associated with lower familiarity with firm- and country-specific information conveyed through, e.g., official disclosures, press releases, or local media. To validate the intuition behind this result, Figure B.5 illustrates variations in language similarity across both proximate and distant countries using four benchmark cases: South Africa, Argentina, Australia, and the United Arab Emirates. For all countries, we observe both proximate and distant regions sharing the same language, ensuring meaningful variation in the data.

Economic Ties. Next, Columns (3) and (4) split the sample at the median level of bilateral trade intensity between the firm's and the analyst's countries.[24] Forecasters are more prone to local extrapolation when economic ties are weak, consistent with the idea that stronger trade relations are associated with greater familiarity with the economic conditions and products of a foreign country. Figure B.6 further corroborates this interpretation by illustrating variation in trade intensity across the same four benchmark countries, where both geographically proximate and distant regions exhibit varying degrees of trade integration.

Cultural Affinity. Finally, Columns (5) and (6) examine the impact of cultural differences, measured using the degree of music consumption overlap between countries.[25] Local extrapolation is greater when cultural overlap per this measure is low, suggesting that familiarity with cultural preferences correlates with the greater integration and proximity to the context of the foreign country. Figure B.7 confirms sufficient variation in cultural exposure across both proximate and distant regions for the same benchmark countries as above.

[23] Interestingly, in Columns (4) to (6), we find no evidence that local inflation influences long-term expectations when recent inflation experiences are included in the regressions instead of recent local GDP growth. This result suggests that the extrapolation mechanism operates primarily through macroeconomic productivity factors rather than inflation. This conclusion aligns with prior evidence from Decaire and Graham (2024), which documents no significant relationship between TGR beliefs and inflation.

[24] Trade intensity is measured as the bilateral imports between two countries from 2000 to 2021 (OECD data), normalized by the total trade volume of the focal country, and averaged over the sample period.

[25] We construct an index based on Shazam's Top 200 charts for the week of December 17, 2024, normalized to range from 0 (no overlap) to 100 (perfect overlap). Higher overlap indicates greater familiarity with the cultural content of the firm's country.

Corroborating the importance of cultural familiarity in shaping expectations, Table B.6 shows that forecasters' TGR expectations are significantly higher for firms located in more trustworthy countries—as measured by *pairwise* trust between the analyst's and the firm's country—using data from the Eurobarometer survey. The relationship between trust and expectations remains highly stable when we include both analyst*year and firm*year fixed effects, thus tightening the estimation by relying on within-analyst-year and within-firm-year variation in trust between the forecaster and the forecasted firm.

Overall, the evidence in Table 7 extends the prior work on local extrapolation (e.g., Kuchler and Zafar, 2019) by documenting strong extrapolative effects across countries and economic environments, even among professional forecasters and for long-run economic variables. Our findings further contribute by highlighting how these effects vary with similarity and familiarity, with extrapolation being particularly pronounced in more abstract and less familiar contexts. These patterns suggest that lower firm (and cross-country) familiarity leads forecasters to rely more heavily on locally experienced economic conditions when forming abstract, long-run expectations.

5.5 Opening the Black Box of Quantitative Forecasts: Qualitative Textual Discussions in Analysts' Reports

To further shed light on the belief formation mechanisms underlying analysts' quantitative long-run expectations, we turn to the textual discussions included in equity reports that inform the quantitative beliefs, examining how analysts allocate attention in the qualitative sections of their reports across topics. We consider both the extensive margin (i.e., the decision to include a given topic) and the intensive margin (i.e., the depth of discussion devoted to covered topics), and explore how variation in topic focus contributes to heterogeneity in analysts' TGR beliefs.

5.5.1 Topic Modeling

To analyze these qualitative discussions systematically, we employ a Latent Dirichlet Allocation (LDA) topic model. This approach is well suited to our setting, as our primary objective is to identify the thematic structure of analysts' reports and to quantify forecasters' topic attention. The LDA framework allows us to map each report into a distribution over topics, capturing both which topics are discussed and the relative emphasis placed on each. We implement the LDA estimation in two steps: text pre-processing and statistical estimation.

Text Pre-Processing. For the topic analysis, we first extract all text (rather than the quantitative information as in Section 3.2.1) from each report using a series of Python routines, now excluding information contained in figures and tables, and excluding the legal boilerplate disclaimers at the end of reports. To address residual extraction issues, we include a refinement step using Meta's

LLM, Llama-2-70b.[26] We then apply a set of filters and transformations to prepare the text for topic modeling. Specifically, reports must contain at least 100 words to be included; verbs are removed; stopwords are dropped and remaining words are lemmatized using standard Python libraries; and terms are organized into unigrams and bigrams based on co-occurrence patterns across the corpus.[27] Finally, we retain only unigrams and bigrams that appear in at least 0.1% of reports.

LDA Model Estimation. Our estimation strategy is similar in spirit to the approach in Bybee et al. (2024), to which we refer readers for further details on the estimation technicalities. Intuitively, the LDA model represents each analyst report as a mixture of latent topics, where each topic corresponds to a distribution over words. This framework allows us to identify the set of themes discussed in a report and to quantify the relative attention allocated to each theme. Consequently, LDA provides a natural way to measure both the extensive margin of topic inclusion and the intensive margin of topic emphasis in analysts' qualitative discussions.

Estimating the LDA model requires selecting three hyperparameters: (i) the number of topics within reports (K), a parameter governing the distribution of topics within reports (α), and (iii) a parameter determining the distribution of terms within a topic (η). To determine the optimal number of topics, we estimate models with K ranging from 10 to 200 in increments of 10. Model selection is then based on three criteria: topic coherence, perplexity, and manual inspection of topic keywords. The coherence score measures semantic similarity within topics, while perplexity assesses in-sample fit. Manual inspection ensures that topics are economically meaningful and interpretable. We set $\alpha = 0.05$ and $\eta = 0.02$, values chosen to optimize perplexity and consistent with standard LDA practice, where both hyperparameters are typically initialized on the order of $1/K$.[28]

Figure B.8 plots coherence and perplexity across alternative values of K (Appendix Table B.7 reports topic keywords for the preferred specification). Most gains in both metrics occur up to 50 topics, after which improvements level off and the marginal topics beyond 50 increasingly fragment existing themes without capturing new economic content. Manual inspection confirms that the 50-topic model is parsimonious and yields distinct, largely non-overlapping topics, whereas models with a greater number of topics risk being more generic and exhibit greater overlap. Accordingly, we adopt the 50-topic model as our baseline and use the 80-topic specification to validate the robustness

[26] For example, when PDF text extraction removes spaces between words, LLMs can infer and restore the intended word boundaries.

[27] Words that frequently co-occur are more likely to be treated as bigrams rather than separate unigrams. For example, when "supply" and "chain" often appear consecutively, they are grouped as the bigram "supply chain."

[28] Our results are not sensitive to perturbations around these parameter values.

of our results.[29]

5.5.2 Topic Attention Allocation

Using the estimated LDA model from Section 5.5.1, we begin by examining the topics discussed in analysts' equity reports to characterize the basic properties of topic attention allocation. The average report concentrates on a small number of topics: the top seven topics account for 84% of total text content, on average. Additional topics receive little weight, with each accounting for about 0.4 percentage points of the total attention "budget."

Two non-mutually exclusive mechanisms could generate these patterns. First, a small number of topics may *objectively* capture a firm's prospects at a given point in time. Alternatively, analysts may *subjectively* focus on a limited set of topics when forming expectations. We distinguish between these explanations by comparing topic coverage across pairs of analysts covering the same firm at the same time.

Extensive Margin Differences. If objective factors primarily drive topic attention, we would expect substantial overlap in the topics discussed by different analysts. In contrast, if subjective weighting plays a larger role, topic coverage should diverge across analysts. To assess this, we compute the Jaccard similarity between the sets of topics covered by each analyst pair:

$$\text{Jaccard similarity} = \frac{|A \cap B|}{|A \cup B|}, \tag{6}$$

where A and B denote the sets of topics discussed by analysts A and B, respectively. The Jaccard similarity ranges from 0 (no overlap) to 1 (complete overlap). In our baseline analysis, we classify a topic as *covered* if it accounts for at least 1% of a report's qualitative discussion.

For the average analyst pair, 38% of topics covered by at least one analyst appear in both reports, with an interquartile range of 18 percentage points (29% to 47%). For illustration, a Jaccard score of 40% could arise if analyst A covers topics 1, 2, 3, and 4, while analyst B covers topics 1, 2, and 5.

Overall, these results indicate that while analysts agree on a core set of topics, substantial cross-analyst differences remain, consistent with person-specific subjective factors shaping the broader set of topics discussed.

Intensive Margin Differences. We next examine differences in topic emphasis at the intensive margin, i.e., how much weight analysts assign to each topic they cover. On average, analysts allocate

[29] The thematic granularity of our preferred LDA specification is consistent with alternative approaches that define economically interpretable topics ex ante and group them into broader categories (Bastianello, Decaire, and Guenzel, 2025a). Despite methodological differences, both approaches capture analysts' qualitative reasoning at a comparable level of abstraction, so the 50-topic LDA model provides a flexible, data-driven representation of the same underlying dimensions of discussion. Nonetheless, as described, we use the 80-topic model to validate the robustness of this section's main result in Table B.8.

15% of report content to each covered topic, with an interquartile range of 5 percentage points (10–20%), indicating meaningful variation in attention allocation.

To quantify intensive-margin similarity, we compare the vectors of topic weights for overlapping topics across analyst pairs covering the same firm at the same time, using Cosine similarity:

$$\text{Cosine similarity} = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}, \tag{7}$$

where $\mathbf{A}$ and $\mathbf{B}$ denote the vectors of topic weights assigned by analysts A and B, respectively. Cosine similarity ranges from -1 (opposite weighting) to 1 (identical weighting).

The average Cosine similarity across analyst pairs is 0.81, with an interquartile range of 0.73 to 0.93, corresponding to an average dissimilarity of roughly 0.19. While this level of similarity may appear high, Table B.9 shows that it is only slightly larger than the similarity observed between firms operating in different industries (0.78). This comparison implies only modest additional convergence in topic emphasis among analysts covering the same firm at the same time, relative to a cross-industry benchmark.

Taken together, these findings indicate that the LDA model captures economically meaningful heterogeneity in analysts' attention allocation, both in topic inclusion (extensive margin) and topic emphasis (intensive margin).[30]

5.5.3 Subjective Differences in Topic Attention and Variation in TGR Expectations

Having established systematic differences in analysts' topic focus at both the extensive and intensive margins, we now examine the extent to which these differences help explain heterogeneity in analysts' long-term TGR expectations. This analysis directly links the qualitative content of analyst reports to quantitative variation in beliefs, providing insight into the belief formation process.

Specifically, comparing analyst pairs, we ask whether differences in TGR expectations are driven by variation in the topics analysts choose to cover, or by differences in how they discuss the same topics. As before, this pairwise approach exploits a key feature of our setting–the frequent presence of multiple analysts forecasting the same firm at the same time. It also offers a practical way to address the dimensionality of LDA output by focusing on variation in topics that are not systematically covered across all analysts or firms within a subgroup, distinguishing our approach from alternatives such as Lasso-based methods (e.g., Bybee et al., 2024).

[30] These patterns are not driven by differences in report length. Both Jaccard and Cosine measures account for total report content. Restricting the sample to report pairs with less than a 1% difference in length yields similar results, with an average Jaccard similarity of 39% and a Cosine similarity of 83%.

We measure disagreement in TGR expectations between analyst pairs as:

$$TGR_{A,B,i,t} = (TGR_{A,i,t} - TGR_{B,i,t})^2, \tag{8}$$

where A and B denote the two analysts in the pair, i indexes the forecasted firm, and t denotes the forecast year. Higher values of $TGR_{A,B,i,t}$ indicate greater differences in analysts' TGR beliefs.[31]

Panel A of Table 8 presents the results. Columns (1) to (3) report estimates for the extensive margin using Jaccard similarity, while Columns (4) to (6) focus on the intensive margin using Cosine similarity. Across specifications, coefficients on both similarity measures are negative, indicating that lower alignment in topic coverage or emphasis is associated with greater disagreement in TGR expectations.

The strength of this relationship differs markedly across margins. While the extensive-margin Jaccard coefficients are negative, they are not statistically significant at conventional levels (p-values > 0.10). In contrast, the intensive-margin Cosine coefficients are large (around -0.50) and statistically significant across specifications. This significantly negative relationship holds in the raw data (Column (4)), with year and 2-digit NAICS fixed effects (Column (5)), and with year-by-industry fixed effects. Because both the dependent variable and topic similarity are defined as differences at the firm-year level, the analysis is inherently conducted within firm-year.

Finally, Column (7) includes both Jaccard and Cosine similarity measures simultaneously. The results reinforce the conclusion that intensive-margin differences in topic emphasis are particularly informative about cross-analyst differences in TGR expectations.

Overall, these findings indicate that heterogeneity in very long-term expectations is driven primarily by differences in how analysts emphasize shared topics rather than by differences in topic selection. More broadly, the results highlight that belief heterogeneity can persist even under common information environments, reflecting variation in emphasis rather than content.

5.5.4 Understanding Subjective Differences in Topic Attention

In a final set of analyses leveraging analysts' qualitative discussions, we examine the determinants and implications of cross-analyst differences in intensive-margin topic focus. This also allows us to validate that the results reported in Panel A of Table 8 are driven by *economically intuitive* differences in how forecasters allocate attention across topics.[32]

[31] Since TGR is a rate, squared differences are comparable across firms without normalization. This contrasts with disagreement measures based on levels (e.g., earnings or price targets), which require normalization to account for differences in magnitude. Consistent with the approaches above, we winsorize the disagreement measure at the 0.5th and 99.5th percentiles.

[32] While the LDA model is well suited for identifying thematic structure in reports, as discussed, the resulting topics do not carry inherent semantic labels. Verifying the content of these topics is therefore essential to ensure that variation in topic emphasis reflects meaningful differences in attention rather than artifacts of unsupervised modeling.

To this end, we organize the topics extracted from the LDA exercise into five broad content categories: (i) macroeconomics, (ii) industry, (iii) geography, (iv) operations, and (v) valuation. Categories are assigned based on a review of the underlying word content of each LDA-derived topic.[33] We then reestimate Cosine similarity by grouping topics into categories tied to broader, less directly observable drivers of long-run expectations—macroeconomic, geographic, and industry-related—and those grounded in firm-level operations and valuation. This decomposition allows us to identify which types of content most strongly account for differences in long-run beliefs.

Panel B of Table 8 shows shows that the negative relationship between topic similarity and variation in TGR beliefs is driven by similarity in macroeconomic, geographic, and industry-related content, while similarity in operations and valuation topics is small and statistically insignificant. Consistent with our broader evidence on local extrapolation, these results indicate that differences in expectations are most strongly associated with variation in attention intensity devoted to broader, less directly anchored domains.

Finally, to further examine and validate the role of geographic location, we examine whether physical distance between forecasters influences how they emphasize topics at the intensive margin. Columns (1) to (3) of Panel C in Table 8 show that analysts located farther apart exhibit greater divergence in topic emphasis overall. Columns (4) to (6) focus specifically on geographic topics, documenting that greater geographic distance between analysts is associated with larger differences in how extensively they discuss geographic content.[34]

Taken together, these findings provide additional support for the role of local extrapolation in shaping belief heterogeneity, with variation in intensive-margin topic focus, linked to local conditions and experiences, serving as an underlying mechanism.

6 Application to Asset Pricing

In the final section of the paper, we provide a simple calibration to illustrate the asset-pricing implications of our results. Specifically, we build on the evidence above on cross-sectional differences in TGR expectations across forecasters and show that the differences observed empirically can generate substantial and realistic levels of valuation disagreement. While a full asset-pricing treatment that models equilibrium pricing in the presence of heterogeneous long-run expectations and traces their implications for returns is beyond our scope, the calibration highlights the quantitative potential of this channel.

Setup. We build on the framework of Delao, Han, and Myers (2024, DHM), who model how investors or forecasters learn about firms' future earnings growth over time. In their model, agents

[33] Appendix Table B.7 reports the complete list of LDA-derived topics and their assigned categories.

[34] For a concrete example, the results indicate that analysts based in, say, Argentina and Brazil are more likely to align in their discussions of a firm's geographic exposure than analysts based in Argentina and the United States.

form expectations by updating their beliefs each period based on new information about realized earnings. The process is referred to as "learning" because agents do not know the true long-run average growth rate and must infer it gradually as new data arrive. Each period, they revise their prior belief by a fixed proportion, which is the "learning gain," of the difference between what was realized and what had been expected. A higher gain implies faster learning (greater responsiveness to new data), while a lower gain implies slower learning (more persistence).

Formally, in the DHM framework the one-period-ahead expected growth rate follows a recursive updating rule:

$$E_t[g_{t+1}] = E_{t-1}[g_t] + \gamma\big(g_t^{\text{obs}} - E_{t-1}[g_t]\big), \tag{9}$$

where g_t^{obs} is the most recently observed earnings growth, and $0 < \gamma < 1$ is the learning gain.

To make the dynamics easier to interpret over longer horizons, we reparameterize the gain as a *persistence parameter* $\rho = 1 - \gamma$, where ρ measures how much of last period's belief carries forward into the next period. A higher ρ means that beliefs are more persistent (agents update slowly), while a lower ρ means that beliefs respond more strongly to new information. This reparameterization is useful because it allows us to express long-horizon expectations directly as a geometric decay toward a long-run anchor. For forecaster j, we write horizon-s expectations as

$$E_t^j[g_{t+s}] = g_{\infty,j} + \rho_j^{s-1}\big(E_t^j[g_{t+1}] - g_{\infty,j}\big), \qquad 0 < \rho_j < 1, \tag{10}$$

where $g_{\infty,j}$ is the forecaster-specific terminal growth anchor. Intuitively, this equation states that short-run beliefs $E_t^j[g_{t+1}]$ decay gradually toward a long-run anchor $g_{\infty,j}$ as the forecast horizon s increases. If all forecasters share the same $g_{\infty,j} = \bar{g}$, disagreement across them necessarily fades with horizon, as in DHM. Allowing $g_{\infty,j}$ to differ across forecasters instead generates persistent long-run disagreement.

Valuations are then obtained from a two-stage discounted cash flow model with a constant discount rate r:

$$P^j = \sum_{s=1}^{T^*} \frac{CF_0 \prod_{h=1}^{s}\big(1 + E_t^j[g_{t+h}]\big)}{(1+r)^s} + \frac{CF_0 \prod_{h=1}^{T^*}\big(1 + E_t^j[g_{t+h}]\big)}{(1+r)^{T^*}} \frac{1 + g_{\infty,j}}{r - g_{\infty,j}}, \quad r > g_{\infty,j}. \tag{11}$$

Calibration. We simulate $J = 1,000$ forecasters. Baseline parameters are $r = 10\%$, $T^* = 10$, near-term expected growth $\bar{g}_1 = 5\%$ (with a cross-forecaster standard deviation of 0.3 pp), and a persistence distribution $\rho_j \sim \text{Beta}(\mu_\rho = 0.6, \sigma_\rho = 0.20)$ defined on $(0, 1)$. We normalize initial cash flow to $CF_0 = 1$, which is without loss of generality because prices scale linearly in CF_0. In this baseline (DHM-style) case, all forecasters share a common long-run anchor $g_{\infty,j} = 2\%$. Disagreement arises from (i) heterogeneity in ρ_j (learning or persistence) and (ii) modest dispersion in near-term expectations $E_t^j[g_{t+1}]$ (the 0.3 pp noise around $\bar{g}_1$). We hold discount rates fixed across forecasters, so r does not contribute to disagreement here, and all forecasters share the same information set about realized earnings growth.

In the alternative calibration, we shut down heterogeneity in learning speeds (i.e., letting $\sigma'_\rho \to 0$), and instead introduce variation in long-run anchors, $g_{\infty,j} \sim \mathcal{N}(0.02, \sigma_\infty^2)$. We calibrate the standard deviation of long-run anchors to $\sigma_\infty = 0.3$ pp around the 2% mean.The resulting variation in long-run growth rates with this parameterization, shown in Panel A of Figure 4, is slightly smaller than that reported in Table 1, reflecting the fact that the summary statistic there also captures cross-firm variation. Throughout, we assume that differences in short-run expectations, long-run anchors, and learning speeds are uncorrelated across forecasters, which isolates the role of each channel. Allowing positive correlation between short- and long-run beliefs would only amplify the resulting valuation dispersion.

Results. Panel B of Figure 4 presents the calibration results for both scenarios, comparing the coefficients of variation. This measure captures the standard deviation of valuations relative to their mean and reflects the extent of disagreement across forecasters after normalizing for overall valuation levels. The figure shows that this relative dispersion is nearly identical across the two calibrations. In other words, modest differences in long-run growth anchors can generate as much relative variation in valuations as considerable heterogeneity in learning speeds, underscoring the quantitative significance of this channel.

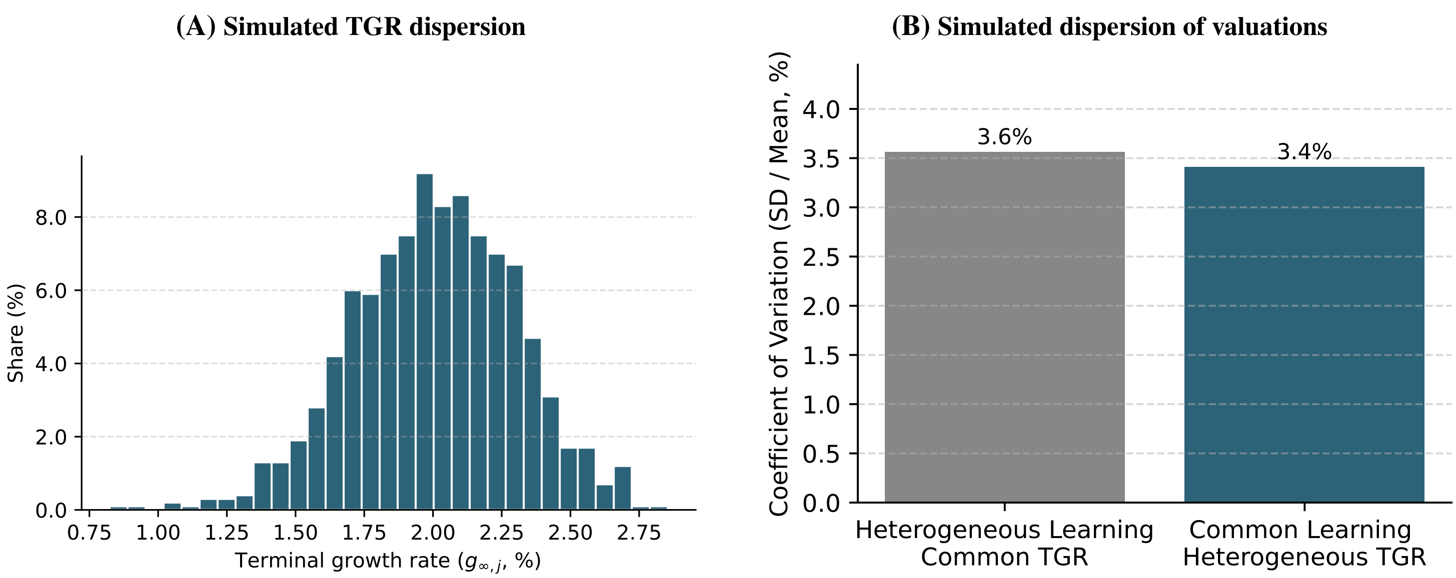


Figure 4: TGR dispersion and disagreement Panel A shows the simulated TGR dispersion in the described calibration with variation in long-run anchors, $g_{\infty,j} \sim \mathcal{N}(0.02, \sigma_\infty^2)$. Panel B plots the coefficient of variation in forecasts for the two calibrations, with heterogeneous learning and common long-run anchors, and with common learning and heterogeneous long-run anchors, respectively.

7 Conclusion

In this paper, we introduce a new large-scale dataset of terminal growth rate (TGR) expectations for individual firms provided by professional forecasters. The dataset combines analysts' very long-run quantitative forecasts with their qualitative discussions, as well as forecaster identities and backgrounds, allowing us to study the formation of long-horizon cash-flow expectations in a setting where such beliefs play a central role for valuation.

We document several new findings. TGR expectations capture economically meaningful variation distinct from shorter-horizon forecasts and are closely related to realized long-run firm growth, while showing no evidence of reverse-engineering from prices. At the same time, they exhibit substantial and persistent heterogeneity across forecasters—exceeding that observed for shorter-horizon beliefs—particularly for younger firms, where anchors to observable fundamentals are weaker. This heterogeneity is shaped in part by systematic extrapolation from locally experienced economic conditions, especially when forecasters evaluate geographically and culturally distant firms, and is closely linked to differences in analysts' textual emphasis on qualitative information. Finally, a simple asset-pricing application illustrates that even modest dispersion in very long-run beliefs can generate quantitatively realistic valuation disagreement.

Taken together, our findings inform macro-finance models and theories of belief formation. By directly observing very long-run expectations—an object that has remained largely unmeasured—we provide empirical evidence on how beliefs are formed when outcomes lie far in the future and validation from realized data is limited. A central implication is that when such validation is weak, persistent individual heterogeneity plays a larger role in shaping expectations. More broadly, the results highlight the importance of local extrapolation and differential interpretation of shared information in environments where anchors to observable fundamentals are limited. A further takeaway is the close connection between forecasters' quantitative beliefs and their qualitative discussions. We view these results as a step toward a richer understanding of long-run belief formation and its implications for asset prices and capital allocation, particularly given the central role of long-horizon expectations in firm valuation.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67:251–333.

Andre, P., P. Schirmer, and J. Wohlfart. 2024. Mental Models of the Stock Market. *Working paper* .

Andrei, D., J. Cujean, and M. Wilson. 2023. The Lost Capital Asset Pricing Model. *The Review of Economic Studies* 90.

Antoniou, C., C. Frydman, and M. Kilic. 2025. Beyond Firm Characteristics: Forecaster Heterogeneity Explains a Large Portion of Over and Underreaction. *Available at SSRN 5501818* .

Asquith, P., M. B. Mikhail, and A. S. Au. 2005. Information Content of Equity Analyst Reports. *Journal of Financial Economics* 75:245–282.

Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic. 2008. How well do institutional theories explain firms' perceptions of property rights? *The Review of Financial Studies* 21:1833–1871.

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

Barber, B. M., X. Huang, and T. Odean. 2016. Which Factors Matter to Investors? Evidence from Mutual Fund Flows. *The Review of Financial Studies* 29:2600–2642.

Bastianello, F., P. Decaire, and M. Guenzel. 2025a. Mental Models and Financial Forecasts. Tech. rep.

Bastianello, F., P. Decaire, and M. Guenzel. 2025b. Valuation Models. Tech. rep.

Bauer, M. D., and G. D. Rudebusch. 2020. Interest Rates under Falling Stars. *American Economic Review* 110:1316–54.

Ben-David, I., E. Fermand, C. M. Kuhnen, and G. Li. 2018. Expectations uncertainty and household economic behavior. Tech. rep., National Bureau of Economic Research.

Berk, J., and P. DeMarzo. 2019. *Corporate Finance*. New York.

Berk, J. B., and J. H. van Binsbergen. 2016. Assessing asset pricing models using revealed preference. *Journal of Financial Economics* 119:1–23.

Bharath, S., and T. Shumway. 2008. Forecasting Default with the Merton Distance-to-Default Model. *Review of Financial Studies* 21:1339–1369.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer. 2025. Finance without exotic risk. *Journal of Financial Economics* 173:104145.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer. 2024. Belief overreaction and stock market puzzles. *Journal of Political Economy* 132:1450–1484.

Brav, A., R. Lehavy, and R. Michaely. 2005. Using Expectations to Test Asset Pricing Models. *Financial Management* 34.

Brealey, R., S. Myers, and F. Allen. 2022. *Principles of Corporate Finance*. New York.

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

Bybee, L., B. Kelly, and Y. Su. 2023. Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text. *The Review of Financial Studies* 36:4759–4787.

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

Charles, C., and C. Kendall. 2024. Causal Narratives. *Working paper* .

Charles, C., and P. Sui. 2024. Marketwide Memory. *Available at SSRN* .

Chaudhry, A. 2025. The impact of prices on analyst cash flow expectations: Reconciling subjective beliefs data with rational discount rate variation. *Journal of Financial Economics* 171:104095.

Chen, L., Z. Da, and X. Zhao. 2013. What Drives Stock Price Movements? *The Review of Financial Studies* 26:841–876.

Coibion, O., and Y. Gorodnichenko. 2012. What can survey forecasts tell us about information rigidities? *Journal of Political Economy* 120:116–159.

Coibion, O., and Y. Gorodnichenko. 2015. Information Rigidity and the Expectations Formation Process: A Simple Framework and New Fact. *American Economic Review* 105:2644–2678.

Cui, T., R. Delao, and S. Myers. 2025. The Subjective Belief Factor. *Working Paper* .

D'Acunto, F., P. Ghosh, and A. G. Rossi. 2022. How costly are cultural biases? Evidence from FinTech .

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber. 2023. IQ, expectations, and choice. *The Review of Economic Studies* 90:2292–2325.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber. 2021. Exposure to grocery prices and inflation expectations. *Journal of Political Economy* 129:1615–1639.

D'Acunto, F., and M. Weber. 2024. Why survey-based subjective expectations are meaningful and important. *Annual Review of Economics* 16.

D'Acunto, F., J. Xie, and J. Yao. 2022. Trust and contracts: Empirical evidence. Tech. rep., LawFin Working Paper.

Das, S., C. M. Kuhnen, and S. Nagel. 2020. Socioeconomic status and macroeconomic expectations. *The Review of Financial Studies* 33:395–432.

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

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

Delao, R., X. Han, and S. Myers. 2024. The Cross-section of Subjective Expectations: Understanding Prices and

Anomalies. *Working Paper* .

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

Delao, R., and S. Myers. 2024. Which Subjective Expectations Explain Asset Prices? *The Review of Financial Studies* 37:1929–1978.

Demgensky, L., and U. Fritsche. 2023. Narratives on the causes of inflation in Germany: First results of a pilot study. Tech. rep.

Giglio, S., M. Maggiori, and J. Stroebel. 2015. Very long-run discount rates. *The Quarterly Journal of Economics* 130:1–53.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2021. Five Facts about Beliefs and Portfolios. *American Economic Review* 111:1481–1522.

Golez, B. 2014. Expected returns and dividend growth rates implied by derivative markets. *The Review of Financial Studies* 27:790–822.

Gormsen, N., and E. Lazarus. 2026. Interest Rates and Equity Valuations. Tech. rep., Working Paper.

Graham, J. R., and C. Harvey. 2001. The Theory and Practice of Corporate Finance: Evidence from the Field. *Journal of Financial Economics* 60:187–243.

Guenzel, M., and U. Malmendier. 2020. Behavioral Corporate Finance: The Life Cycle of a CEO Career. *Oxford Research Encyclopedia of Economics and Finance, September 2020* .

Guercio, D. D., and P. A. Tkac. 2002. The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds vs. Pension Funds. *The Journal of Financial and Quantitative Analysis* 37:523–557.

Hartzmark, S. M., S. D. Hirshman, and A. Imas. 2021. Ownership, Learning, and Beliefs*. *The Quarterly Journal of Economics* 136:1665–1717.

Hashim, N. A., and N. C. Strong. 2018. Do Analysts' Cash Flow Forecasts Improve Their Target Price Accuracy? *Contemporary Accounting Research* 35:1816–1842.

Hillenbrand, S., and O. McCarthy. 2024. Valuing Stocks With Earnings. *Working Paper* .

Jones, K. 1996. Trust as an Affective Attitude. *Ethics* 107:4–25.

Kendall, C., and R. Oprea. 2024. On the complexity of forming mental models. *Quantitative Economics* 15:175–211.

Klepper, S. 1996. Entry, exit, growth, and innovation over the product life cycle. *The American Economic Review* pp. 562–583.

Kuchler, T., and B. Zafar. 2019. Personal experiences and expectations about aggregate outcomes. *The Journal of Finance* 74:2491–2542.

Lettau, M., and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance* 56:815–849.

Malmendier, U. 2018. Behavioral Corporate Finance. In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 1, pp. 277–379. Elsevier.

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

Malmendier, U., and S. Nagel. 2016. Learning from inflation experiences. *The Quarterly Journal of Economics* 131:53–87.

Martin, I. 2017. What is the Expected Return on the Market? *The Quarterly Journal of Economics* 132:367–433.

Mukhlynina, L., and K. G. Nyborg. 2020. The choice of valuation techniques in practice: education versus profession. *Critical Finance Review* 9:201–265.

Nagel, S. 2021. Empirically Grounded Models of Subjective Beliefs, in: Perspectives on the future of asset pricing. *The Review of Financial Studies* 34:2126–2160.

Nagel, S. 2024. Stock Market Valuation: Explanations, Non-Explanations, and Some Open Questions. Comments Given at the NBER Summer Institute Asset Pricing Group Meeting (Expanded Version).

Nagel, S., and Z. Xu. 2022. Asset Pricing with Fading Memory. *Review of Financial Studies* 35:2190–2245.

Porta, R. L. 1996. Expectations and the Cross-Section of Stock Returns. *Journal of Finance* 51:1715–1742.

Pruitt, S. 2024. Dogs and Cats Living Together: A Defense of Cash-Flow Predictability. *Working Paper* .

Sabbatucci, R. 2022. Are dividends and stock returns predictable? New evidence using M&A cash flows. *Working Paper* .

Weber, M., F. D'Acunto, Y. Gorodnichenko, and O. Coibion. 2022. The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives* 36:157–184.

Table 1: Summary statistics This table reports summary statistics. The sample includes 41,070 equity reports and terminal growth rate (TGR) expectations for 7,975 firms, obtained from equity reports published between 2000 and 2023. Panel A describes the sample coverage. Panel B reports statistics on the TGR expectations. Panel C reports demographic and other characteristics measured at the individual analyst level (i.e., not at the analyst–year level), along with country-level summary statistics. Panels D and E present various economic data series. Variables definitions appear in Appendix A.

Panel A: Firms and Coverage	N					
No. Observations (TGR Expectations Sample)	41,070					
No. Firms	7,975					
No. Analysts	5,229					
No. Analysts (LinkedIn-matched)	2,555					
No. Brokerage houses	42					
Panel B: TGR Expectations	Mean	25^{th} Pct.	Median	75^{th} Pct.	Std. Dev.	N
Terminal growth rate (TGR) Expectation $_{i,j,t}$(%)	2.22	1.50	2.00	3.00	1.28	41,070
TGR$_{i,j,t}$(%) (IBES Overlap)	2.21	1.50	2.00	3.00	1.27	25,167
IBES LTG$_{i,j,t}$(%)	12.66	5.54	10.30	16.94	13.94	25,167
Correlation(IBES LTG, TGR)	0.14					
Panel C: Analyst and Country Statistics	Mean	25^{th} Pct.	Median	75^{th} Pct.	Std. Dev.	N
Female	0.16	0.00	0.00	0.00	0.36	2,555
Asian	0.20	0.00	0.00	0.00	0.40	2,555
Black	0.03	0.00	0.00	0.00	0.16	2,555
Caucasian	0.63	0.00	1.00	1.00	0.48	2,555
Hispanic	0.04	0.00	0.00	0.00	0.18	2,555
Master degree	0.27	0.00	0.00	1.00	0.44	2,555
MBA	0.21	0.00	0.00	0.00	0.41	2,555
Birth year	1977.51	1971.00	1977.00	1985.00	9.38	1,531
No. Reports by analyst	7.86	1.00	3.00	7.00	16.22	5,229
No. Analysts by country	108.94	1.00	5.00	35.00	225.00	48
No. Analysts by brokerage house	138.26	22.00	42.50	180.00	184.19	42
1(Foreign report)	0.31	0.00	0.00	1.00	0.46	40,770
No. Reports by country	855.60	2.00	18.00	248.50	1429.00	48
Panel D: Economic Data Series—Full Sample	Mean	25^{th} Pct.	Median	75^{th} Pct.	Std. Dev.	N
Firm country						
Real GDP growth $_{i,t}$ (%)	2.55	1.55	2.29	3.74	3.36	38,056
Inflation $_{i,t}$ (%)	2.56	1.26	2.00	3.16	2.50	37,933
10-year Treasury yield $_{i,t}$ (%)	3.15	1.68	2.91	4.10	2.31	39,648
Panel E: Economic Data Series—Foreign Reports	Mean	25^{th} Pct.	Median	75^{th} Pct.	Std. Dev.	N
Analyst country						
Real GDP growth $_{j,t}$ (%)	2.00	1.45	2.29	3.12	3.71	11,966
Inflation $_{j,t}$ (%)	2.47	1.45	2.16	2.85	1.92	11,955
10-year Treasury yield $_{j,t}$ (%)	2.98	1.42	2.85	4.13	2.05	12,497
Firm country						
Real GDP growth $_{i,t}$ (%)	2.11	0.96	2.21	3.82	3.58	11,767
Inflation $_{i,t}$ (%)	2.42	0.90	1.83	3.06	2.75	11,652
10-year Treasury yield $_{i,t}$ (%)	3.02	0.97	2.75	4.11	2.76	11,801

Table 2: Economic relevance of TGR expectations Panel A regresses ex-post realized firm earnings growth at different horizons on analysts' terminal growth rate expectations and IBES consensus LTG variable. The unit of observation is at the firm i, analyst j, and forecast year t level. The sample period is 2000–2023. The dependent variable is the geometric average of the annual earnings growth rates experienced by firm i over five- or ten-year periods at various horizons (t to $t+5$, $t+3$ to $t+8$, and $t+3$ to $t+13$, respectively). *Terminal growth rate*$_{i,j,t}$ is the analyst's terminal growth rate expectation, measured in percent. $LTG_{i,t}$ is the IBES consensus LTG belief in percent at time t for firm i. Panel B studies the relationship between TGR expectations and default risk, measured using the Merton Model adjustment from Bharath and Shumway (2008), as well as firm age. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Realized Earnings Growth	$(\Pi_t^{t+5}(1+g_{i,t})^{1/5}-1)$		$(\Pi_{t+3}^{t+8}(1+g_{i,t})^{1/5}-1)$		$(\Pi_{t+3}^{t+13}(1+g_{i,t})^{1/10}-1)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Terminal growth rate (%)$_{i,j,t}$	0.59***	0.56***	0.70***	0.50***	0.49***	0.35***
	(0.13)	(0.14)	(0.14)	(0.17)	(0.11)	(0.12)
IBES Median Consensus LTG (%)$_{i,t}$	0.36***	0.37***	0.10***	0.10***	0.06***	0.05*
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Firm Country*Industry*Year FE	No	Yes	No	Yes	No	Yes
Observations	15,242	13,825	12,582	11,355	7,112	6,367
F-statistics	154.09	80.06	22.03	9.81	13.52	6.54
R^2	0.10	0.41	0.01	0.36	0.01	0.43

Panel B: TGR, Default, and Firm Age	Terminal growth rate (%)$_{i,j,t}$				
	(1)	(2)	(3)	(4)	(5)
Default risk (10-year horizon)$_{i,t}$	-0.16***			-0.17***	
	(0.03)			(0.04)	
Default risk (20-year horizon)$_{i,t}$		-0.18***			-0.21***
		(0.03)			(0.04)
ln(Firm post-IPO age)$_{i,t}$			-0.14***	-0.16***	-0.16***
			(0.02)	(0.03)	(0.03)
Firm Country*Industry*Year FE	Yes	Yes	Yes	Yes	Yes
Observations	25,997	25,997	21,890	14,633	14,633
F-statistics	40.17	48.80	42.88	25.72	30.80
R^2	0.34	0.34	0.33	0.36	0.36

Table 3: TGR expectations, discount rates, and realized returns Panel A examines the relationship between one-year realized total equity returns, TGR expectations, and analysts' subjective discount-rate estimates. We measure 52-week realized total equity returns using data from Refinitiv Eikon. The subjective equity premium is constructed as $\beta \times ERP$, and the subjective risk-free rate corresponds to the rate analysts use when estimating their cost of equity. Panel B examines the relationship between TGR expectations and subjective discount-rate estimates. The subjective cost of equity is defined as $r_f + \beta \times ERP$, combining the subjective risk-free rate and the subjective equity risk premium. All subjective inputs are drawn directly from analysts' valuation reports. Variable definitions appear in Section Appendix A. All specifications are estimated by ordinary least squares. Standard errors, reported in parentheses, are heteroskedasticity-consistent and clustered at the firm level. Statistical significance is denoted as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Realized Returns	**Realized Total Equity Returns (%)$_{i,t+1}$**			
	(1)	(2)	(3)	(4)
TGR (%)$_{i,j,t}$	-0.72**	-1.52***	-0.95*	-2.20***
	(0.31)	(0.33)	(0.55)	(0.81)
Subjective Discount Rate (%)$_{i,j,t}$	1.97***	1.09***		
	(0.23)	(0.32)		
Subjective Risk Premium (%)$_{i,j,t}$			1.67***	1.66***
			(0.40)	(0.60)
Subjective Risk-Free Rate (%)$_{i,j,t}$			1.84***	-1.63**
			(0.55)	(0.77)
Firm FE	No	Yes	No	Yes
Firm Country × Year FE	No	Yes	No	Yes
Observations	24,907	22,852	5,454	4,163
F Statistics	38.13	12.44	7.83	7.02
R^2	0.01	0.47	0.01	0.60
Panel B: TGR and Discount Rate	**TGR$_{i,j,t}$**			
	(1)	(2)	(3)	(4)
Subjective Discount Rate (%)$_{i,j,t}$	0.12***	0.00		
	(0.01)	(0.01)		
Subjective Cost of Equity (%)$_{i,j,t}$			0.08***	-0.02
			(0.01)	(0.01)
Firm Country × Year FE	No	Yes	No	Yes
Firm × Analyst FE	No	Yes	No	Yes
Observations	39,239	23,693	8,399	4,035
F Statistics	449.50	0.01	92.98	2.53
R^2	0.03	0.80	0.02	0.86

Table 4: Variance decomposition of TGR and short-run expectations (ANOVA) This table presents a variance decomposition of analyst TGR and EPS growth expectations using ANOVA. We exclude firms, analysts, and brokerage houses with fewer than two observations, as their respective fixed effects fully explain the associated variation in TGR expectations. We also exclude brokerage houses that are not associated with equity analysts who switched employers during the sample period (2000–2023), as it is not possible to separate the effects of these brokerage houses from those of their analysts (Abowd, Kramarz, and Margolis, 1999). Panel A includes the full subsample meeting these criteria (with additional sample filters on the number of reports per analyst specified in the column headers), while Panel B further excludes firms with fewer than ten observations. We estimate sequential (Type I) sums of squares. In both panels, Columns (1) to (4) perform the variance decomposition for the TGR expectations, while Columns (5) to (8) apply it to the IBES 1-Year EPS growth forecasts. The IBES sample, covering the period 2000–2023, includes only firms also present in the TGR sample to allow for a direct comparison. The IBES 1-Year EPS growth forecast is based on the earliest forecast made by an analysts for a given fiscal year, provided that the forecast is issued at least nine months and at most one year before the target date. Analyst expected growth is measured as the EPS forecast divided by the most recent realized EPS of the firm at the time of the forecast, using IBES (ACTUAL) data, minus one. Variable definitions appear in Appendix A.

Panel A: Anova

	TGR				1-Year Expected EPS Growth Rate			
	Full sample	$\frac{Reports}{Analyst} \geq 10$	$\frac{Reports}{Analyst} \geq 25$	$\frac{Reports}{Analyst} \geq 50$	Full sample	$\frac{Reports}{Analyst} \geq 10$	$\frac{Reports}{Analyst} \geq 25$	$\frac{Reports}{Analyst} \geq 50$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Brokerage house FE								
Share of total R^2 (%)	4	4	6	14	0	2	2	3
Degrees of Freedom	39	38	35	27	324	219	102	44
Analyst FE								
Share of total R^2 (%)	69	62	58	41	33	24	7	1
Degrees of Freedom	3,509	1,028	395	103	3,476	736	149	30
Firm FE								
Share of total R^2 (%)	27	33	36	45	66	74	91	96
Degrees of Freedom	5,065	4,538	3,574	2,198	2,567	2,198	1,444	726
Total R^2	0.66	0.66	0.68	0.72	0.36	0.20	0.34	0.39
Observations	36,544	27,398	18,591	9,043	25,600	15,090	6,302	2,609

Panel B: Anova—Obs per firm ≥ 10

	TGR				1-Year Expected EPS Growth Rate			
	Full sample	$\frac{Reports}{Analyst} \geq 10$	$\frac{Reports}{Analyst} \geq 25$	$\frac{Reports}{Analyst} \geq 50$	Full sample	$\frac{Reports}{Analyst} \geq 10$	$\frac{Reports}{Analyst} \geq 25$	$\frac{Reports}{Analyst} \geq 50$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Brokerage house FE								
Share of total R^2 (%)	5	6	8	18	2	2	3	15
Degrees of Freedom	38	37	34	27	308	210	100	44
Analyst FE								
Share of total R^2 (%)	76	70	64	46	58	38	12	1
Degrees of Freedom	2,695	954	383	102	3,063	728	148	30
Firm FE								
Share of total R^2 (%)	19	23	28	36	41	60	85	84
Degrees of Freedom	1,172	1,162	1,126	849	1,182	1,162	907	490
Total R^2	0.62	0.61	0.63	0.67	0.22	0.17	0.25	0.20
Observations	21,833	17,006	11,808	5,614	19,365	11,952	5,056	2,096

Table 5: TGR expectations over firms' life cycle This table analyzes how firm life cycle factors contribute to explaining TGR expectations. The unit of observation is at the firm *i*, analyst *j*, and forecast year *t* level. The sample period is 2000–2023. Column (1) reports the R^2 obtained from estimating the Case 1 Stage 2 regression displayed below for firms within the age range specified in the first column. Column (2) presents the share of the Case 1 Stage 2 regression R^2 attributable to *Firm country* and *Industry* factors, derived through a sequential ANOVA decomposition of the Case 1 Stage 2 specification. Column (3) reports the R^2 obtained from estimating the Case 2 Stage 2 regression for firms within the age range specified in the first column. Column (4) provides a conservative back-of-the-envelope calculation to estimate the share of TGR expectations in Case 1 Stage 1 explained by the firm factors, after accounting for firm country and industry. This estimate is calculated by computing $(1 - R^2_{Column\ 1}) \times R^2_{Columne\ 3}$. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares.

	Case 1		**Case 2**	
Stage 1:	None		$TGR_{i,j,t} = I_{\text{Firm country}} + I_{\text{Industry}} + I_{\text{Analyst}} + I_{\text{Brokerage}} + \epsilon_{i,j,t}$	
Stage 2:	$TGR_{i,j,t} = I_{\text{Firm country}} + I_{\text{Industry}} + I_{\text{Analyst}} + I_{\text{Brokerage}} + \epsilon_{i,j,t}$		$\hat{\epsilon}_{i,j,t} = I_{\text{Firm}} + \upsilon_{i,j,t}$	
	Stage 2 R^2 (1)	**Share of R^2 Driven by Country + Industry** (2)	**Stage 2 R^2** (3)	**Share of R^2 Driven by Firm Residual Factors** (4)
0 <= Firm age < 5	0.67	17%	0.37	12%
5 <= Firm age < 10	0.68	18%	0.34	11%
10 <= Firm age < 15	0.64	18%	0.29	10%
15 <= Firm age < 20	0.68	26%	0.30	10%
20 <= Firm age < 25	0.69	32%	0.28	9%
25 <= Firm age < 30	0.77	38%	0.27	6%
30 <= Firm age < 35	0.74	37%	0.17	5%
35 <= Firm age < 40	0.79	30%	0.25	5%
40 <= Firm age < 45	0.75	33%	0.08	2%
45 <= Firm age < 50	0.90	67%	-0.00	-0%
50 <= Firm age < 55	0.88	53%	0.13	2%
55 <= Firm age < 60	0.78	82%	0.05	1%
60 <= Firm age < 65	0.91	64%	0.00	0%
65 <= Firm age < 70	0.91	80%	0.02	0%
70 <= Firm age < 75	0.86	88%	0.11	2%
75 <= Firm age < 80	0.38	45%	0.13	8%
80 <= Firm age < 85	0.88	59%	0.00	0%
85 <= Firm age < 90	0.89	29%	0.01	0%
90 <= Firm age < 95	0.97	76%	0.00	0%

Table 6: Analyst fixed effects decomposition This table presents the combined explanatory power of various fixed effects for TGR expectations in Panel A and presents a variance decomposition of analyst fixed effects in TGR expectations in Panel B. We exclude firms, analysts, and brokerage houses with fewer than two observations (left three columns), as for these observations, as their respective fixed effects fully explain the associated variation in TGR expectations. We also present robustness excluding observations with fewer than five observations (right three columns). We also exclude brokerage houses that are not associated with equity analysts that switched between employers during the sample period (2000–2023), as it is possible not separate the effects of these brokerage houses and from those of their analysts (Abowd, Kramarz, and Margolis, 1999). Columns (1) and (4) use the entire respective subsample, Columns (2) and (5) restrict the analysis to firms with both domestic and foreign coverage, and Columns (3) and (6) focus on the foreign coverage of firms included in Column (2) or (5), respectively. Panel A shows the combined explanatory power firm, brokerage house, and analyst fixed effects, with the unit of observation at firm *i*, analyst *j*, and forecast year *t* level. Panel B conducts a variance decomposition of analyst fixed effects into five categories using ANOVA: gender (binary variable equal to 1 if the analyst is female and 0 otherwise), race (categorical variable indicating White, Asian, Hispanic, and other races), education (categorical variable for holding a Master's degree, an MBA, or neither), and birthyear (birth year indicators). We estimate sequential (Type I) sums of squares. Variable definitions appear in Appendix A. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Fixed effects	**Terminal growth rate$_{i,j,t}$**					
Sample filters	At least 2 observations per firm, brokerage house, and analyst			At least 5 observations per firm, brokerage house, and analyst		
	Full Sample	Has foreign coverage	Exclude domestic coverage	Full Sample	Has foreign coverage	Exclude domestic coverage
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	2.17***	2.09***	2.11***	2.14***	2.06***	2.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Brokerage FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,626	5,606	3,611	10,141	4,727	3,126
R^2	0.69	0.70	0.74	0.67	0.69	0.73

Panel B: Anova decomposition	**Analyst FE**					
Sample filters	At least 2 observations per firm, brokerage house, and analyst			At least 5 observations per firm, brokerage house, and analyst		
	Full Sample	Has foreign coverage	Exclude domestic coverage	Full Sample	Has foreign coverage	Exclude domestic coverage
	(1)	(2)	(3)	(4)	(5)	(6)
Gender FE						
Share of total R^2 (%)	0	0	1	1	0	1
Degrees of Freedom	1	1	1	1	1	1
Education FE						
Share of total R^2 (%)	1	2	1	3	1	2
Degrees of Freedom	2	2	2	2	2	2
Race FE						
Share of total R^2 (%)	4	3	2	2	1	1
Degrees of Freedom	3	3	3	3	3	3
Birthyear FE						
Share of total R^2 (%)	53	37	35	25	36	43
Degrees of Freedom	46	43	43	43	42	41
Analyst country FE						
Share of total R^2 (%)	43	58	61	69	62	53
Degrees of Freedom	35	36	28	32	33	26
Total R^2	0.17	0.21	0.29	0.27	0.31	0.29
Observations	844	551	361	587	411	289

Table 7: Analyst–firm distance, local economic conditions, and extrapolation This table examines the effect of analysts' local economic conditions on their TGR expectations. The unit of observation is at the firm i, analyst j, and forecast year t level. The sample period is 2000–2023. The dependent variable, $Terminal\ growth\ rate_{i,j,t}$, is equal to the terminal growth rate expectation by the analyst, measured in percent. The 10-year average real GDP growth rates for both the firm's and the analyst's country are obtained from the World Bank and calculated as 10-year trailing averages. Similarly, the 10-year average inflation rates for both the firm's and the analyst's country are constructed using World Bank. as 10-year trailing averages. I^{far} is a binary variable equal to 1 if the geographic distance between the analyst's country and the firm's country is above the sample median, and 0 otherwise. Distance between countries is measured as the shortest distance between their capitals. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Local Extrapolation	Terminal growth rate$_{i,j,t}$					
	(1)	(2)	(3)	(4)	(5)	(6)
10-year avg. real GDP growth (Firm)$_{i,t}$	0.13***	0.09***	0.09***			
	(0.02)	(0.02)	(0.02)			
10-year avg. real GDP growth (Analyst)$_{j,t}$		0.09***	-0.03			
		(0.03)	(0.05)			
10-year avg. real GDP growth (Analyst)$_{j,t} * I^{far}_{i,j,t}$			0.14***			
			(0.05)			
10-year avg. inflation rate (Firm)$_{i,t}$				-0.00	-0.00	-0.00
				(0.00)	(0.00)	(0.00)
10-year avg. inflation rate (Analyst)$_{j,t}$					-0.00	-0.00
					(0.00)	(0.00)
10-year avg. inflation rate (Analyst)$_{j,t} * I^{far}_{i,j,t}$						0.02
						(0.03)
Analyst country*firm country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,658	11,642	11,642	11,558	11,542	11,542
F Statistics	45.05	29.62	20.76	0.28	0.14	0.20
R^2	0.15	0.15	0.15	0.14	0.14	0.14

Panel B: Drivers of Local Extrapolation	Terminal growth rate$_{i,j,t}$					
	Language		Trade relation		Cultural affinity	
	Different	Same	< median	≥ median	< median	≥ median
	(1)	(2)	(3)	(4)	(5)	(6)
10-year avg. real GDP growth (Analyst)$_{j,t} * I^{far}_{i,j,t}$	0.13**	0.08	0.32***	0.09	0.18***	0.01
	(0.05)	(0.21)	(0.09)	(0.06)	(0.07)	(0.09)
Analyst country*firm country FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	10-year avg. real GDP growth (Firm)$_{i,t}$, 10-year avg. real GDP growth (Analyst)$_{j,t}$					
Observations	9,717	1,925	5,774	5,827	5,569	5,562
F Statistics	20.83	4.84	12.30	12.51	14.23	11.48
R^2	0.15	0.15	0.17	0.13	0.18	0.12

Table 8: Variation in TGR Expectations and Differences in Qualitative Discussion of Topics This table examines the relationship between differences in textual discussions in equity reports and disagreement in TGR expectations between pairs of analysts. The unit of observation is at the analyst a, analyst b, firm i, and forecast year t level. The sample period is 2000–2023. The dependent variable in Panels A and B, $Disagreement_{a,b,i,t}$, is defined as the squared difference between analyst a's and analyst b's TGR beliefs: $(TGR_{a,i,t} - TGR_{b,i,t})^2)$. The key independent variables are the Jaccard and Cosine scores, which measure the similarity levels between the two analysts' textual discussions along the extensive and intensive margins, respectively. Panel C relates the Cosine similarity scores to analyst distance. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the analyst-pair level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Panel A: Differences in Qualitative Discussion	Disagreement$_{a,b,i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Jaccard $_{a,b,i,t}$	-0.17	-0.17	-0.11				-0.13
	(0.25)	(0.25)	(0.25)				(0.25)
Cosine $_{a,b,i,t}$				-0.53**	-0.47**	-0.42*	-0.42**
				(0.22)	(0.22)	(0.22)	(0.21)
Year FE	No	Yes	No	No	Yes	No	No
Industry FE	No	Yes	No	No	Yes	No	No
Industry*Year FE	No	No	Yes	No	No	Yes	Yes
Observations	23,466	23,433	23,400	23,466	23,433	23,400	23,400
F Statistics	0.49	0.50	0.21	5.69	4.60	3.77	2.11
R^2	0.00	0.02	0.07	0.00	0.02	0.07	0.07

Panel B: Differences by Content Type	Disagreement$_{a,b,i,t}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Cosine $^{\text{Distant Drivers}}_{a,b,i,t}$	-0.63**	-0.56**	-0.62**	-0.69***	-0.62**	-0.67***
	(0.25)	(0.25)	(0.25)	(0.25)	(0.25)	(0.25)
Cosine $^{\text{Operations \& Valuation}}_{a,b,i,t}$				0.05	0.09	0.10
				(0.18)	(0.17)	(0.17)
Year FE	No	Yes	No	No	Yes	No
Industry FE	No	Yes	No	No	Yes	No
Industry*Year FE	No	No	Yes	No	No	Yes
Observations	22,927	22,895	22,861	22,426	22,394	22,360
F Statistics	6.49	5.16	6.22	3.90	3.22	3.78
R^2	0.00	0.02	0.07	0.00	0.02	0.07

Panel C: Distance Between Analysts	Cosine$_{a,b,i,t}$			Cosine $^{\text{Geography}}_{a,b,i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathrm{I}^{far}_{a,b,t}$	-0.02***	-0.02***	-0.02***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year FE	No	Yes	No	No	Yes	No
Industry FE	No	Yes	No	No	Yes	No
Industry*Year FE	No	No	Yes	No	No	Yes
Observations	24,898	24,893	24,859	18,276	18,271	18,235
F Statistics	12.04	17.05	17.74	31.66	28.23	21.22
R^2	0.00	0.02	0.07	0.01	0.02	0.06

Online Appendix

What Drives Very Long-Run Cash Flow Growth Expectations?

Paul H. Décaire — ASU
Marius Guenzel — Wharton

Not For Publication

Appendix A Variable Definitions

Table A.1: Variable Definitions

Subscript t indicates the forecast year, i indicates a firm, and j indicates a forecaster.

Variable	Definition
10-year average inflation rate (Analyst)$_{j,t}$	The 10-year rolling average of the analyst's country inflation rate, obtained from the World Bank.
10-year average inflation rate (Firm)$_{i,t}$	The 10-year rolling average of the firm's headquarters' country inflation rate, obtained from the World Bank.
10-year average real GDP growth (Analyst)$_{j,t}$	The 10-year rolling average of the analyst's country real GDP growth rate, obtained from the World Bank.
10-year average real GDP growth (Firm)$_{i,t}$	The 10-year rolling average of the firm's headquarters' country real GDP growth rate, obtained from the World Bank.
Default risk (10-year horizon)$_{i,t}$	A binary variable equal to 1 if the firm probability of default is above the sample median, and 0 otherwise. The probability of default is measured following the naive Merton's distance-to-default model discussed in Bharath and Shumway (2008), using a horizon of 10 years (T=10).
Default risk (20-year horizon)$_{i,t}$	A binary variable equal to 1 if the firm probability of default is above the sample median, and 0 otherwise. The probability of default is measured following the naive Merton's distance-to-default model discussed in Bharath and Shumway (2008), using a horizon of 20 years (T=20).
Disagreement$_{a,b,i,t}$	The difference between analyst A and analyst B TGR estimates squared: $(TGR_{a,i,t} - TGR_{b,i,t})^2$.
Firm age$_{i,t}$	The firm post-IPO year age.
$I^{far}_{i,j,t}$	A binary variable equal to 1 if the distance between the analyst's country capital and the firm's country capital is above the sample median, and 0 otherwise. The pairwise distance is measured as the shortest distance between both cities using their centroid GPS coordinates.
LTG (IBES)$_{i,t}$	The consensus measure (median) of the LTG variable provided by the Institutional Brokers' Estimate System (IBES), for the firm and the beginning of the calendar year.
DCF models, measured from the equity reports.	
TGR Industry Dispersion	The difference between analysts' terminal growth rate and the average terminal growth rate measured for each age group-industry-year squared: $(TGR_{i,j,t} - \overline{TGR}^{\text{Age Group}}_{k,t})^2$

Appendix B Supplementary Figures and Tables

Appendix Figures

DCF Model - Aixtron Figures in EUR m	2011e	2012e	2013e	2014e	2015e	2016e	2017e	2018e	2019e	2020e	2021e	2022e	2023e	2024e
Sales														
Change														
EBIT														
EBIT-Margin														
Tax rate														
NOPAT														
Depreciation														
in % of Sales														
Change in Liquidity from														
- Working Capital														
- Capex														
Capex in % of Sales														
Other														
Free Cash Flow (WACC-Model)														

Model parameter			
Debt ratio		Beta	
Costs of Debt		WACC	
Market return			
Risk free rate		Terminal Growth	

Valuation (mln)			
Present values 2024e			
Terminal Value			
Liabilities			
Liquidity		No. of shares (mln)	
Equity Value		**Value per share (EUR)**	

Figure B.1: Example of complete DCF equity report This figure shows a representative example of a discounted cash flow (DCF) valuation model table. This figure is taken from an equity report for Aixtron (Ticker = AIXGn), published by Warburg Research on October 27, 2011. For copyright reasons, all information in the table is redacted.

Figure B.2: TGR expectations versus historical returns This figure compares the time trends of historically realized earnings growth at the $t-8$ to $t-3$ horizon with analysts' terminal growth rate expectations at time t. Realized earnings growth corresponds to the geometric average realized between $t-8$ and $t-3$, averaged by year (t). Terminal growth rate is the average terminal growth rate expectation by year. The x-axis denotes years. The y-axis corresponds to rates, in percent (%).

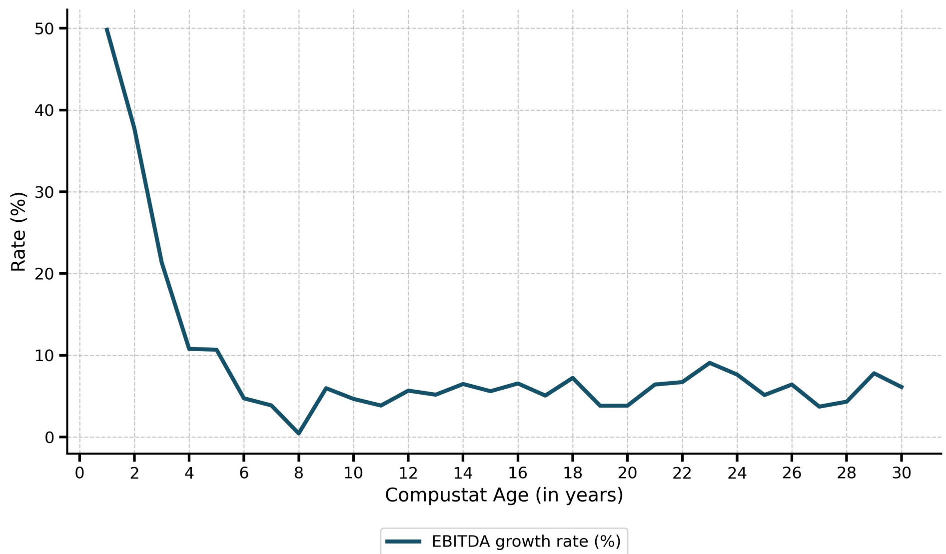


Figure B.3: Annual earnings growth rate and firm age This figure plots realized annual earnings growth as a function of firm age. The x-axis is firm age (based on the IPO date in Compustat). The y-axis is the average annual earnings growth rate across firms. The sample includes all firms included in Compustat for which the data item OIBDP (Operating earnings before depreciation) is available.

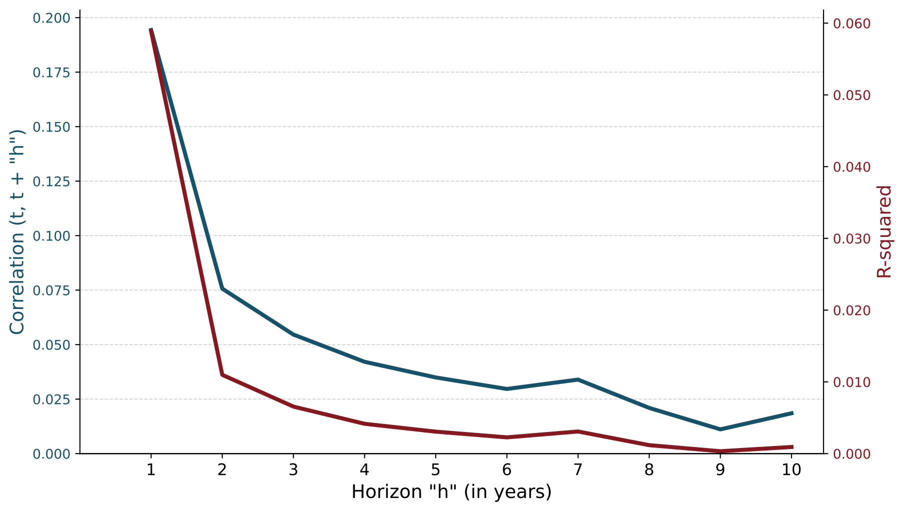


Figure B.4: Correlation between realized growth rates at different horizons This figure plots the average correlation between the realized earnings growth rate at time t and realized earnings growth rate at time t+h. Note both t and t+h refer to realized growth rates observed after t+h. The blue lines shows the estimated correlation coefficients for different values of h. The red line shows the R^2 when regressing earnings growth at time t+h on earnings growth at time t for different values of h.

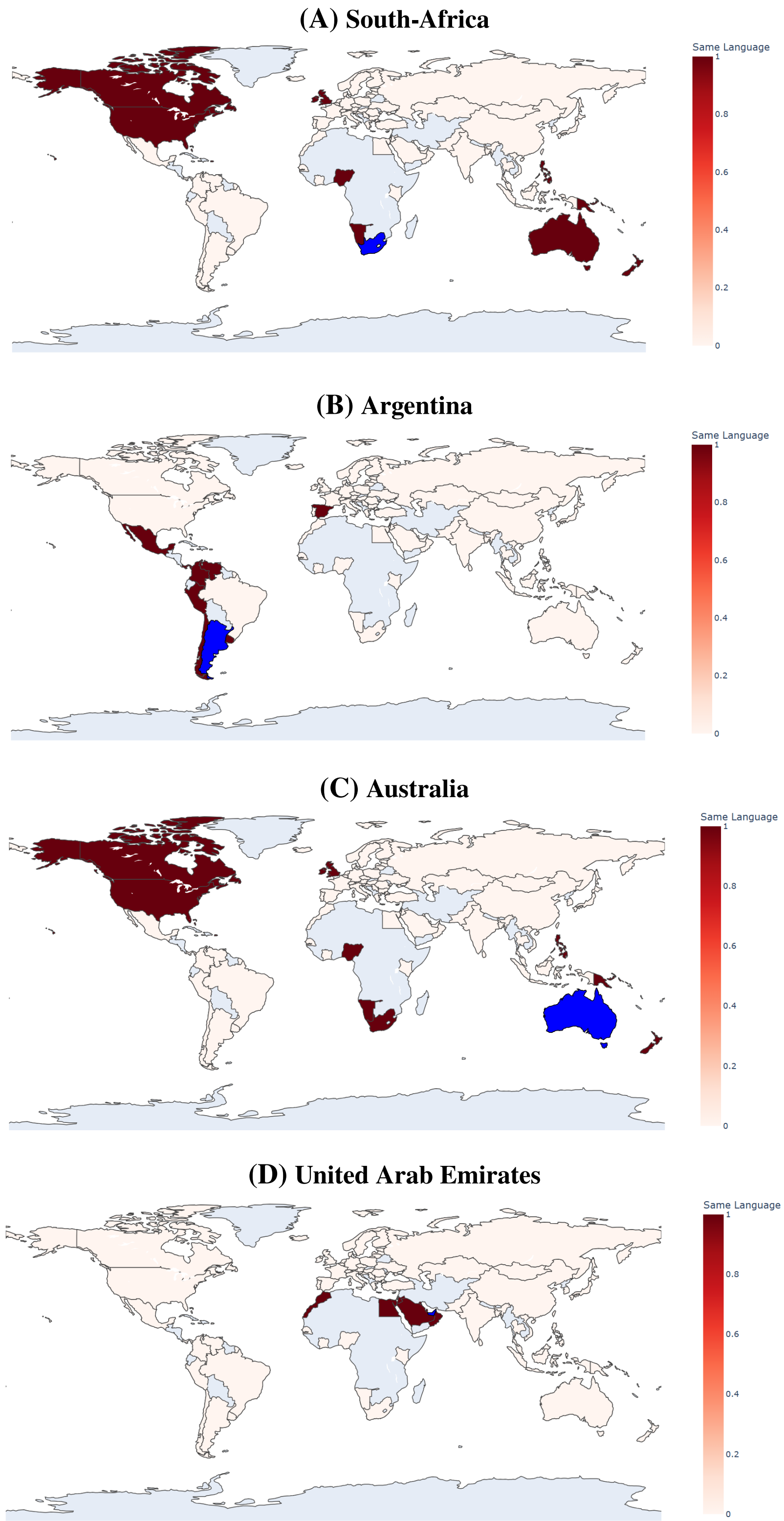
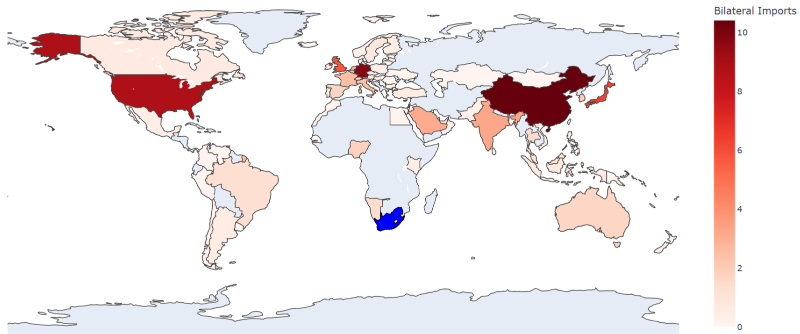
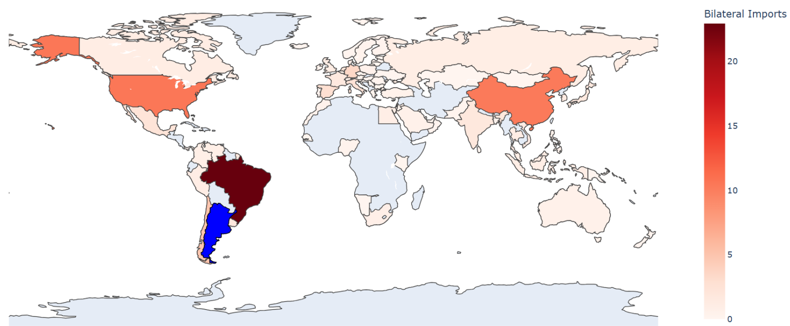


Figure B.5: Language similarity between countries This figure depicts language similarity between a focal country and other countries for four focal countries: (i) South Africa, (ii) Argentina, (iii) Australia, and (iv) the United Arab Emirates. We define a country's language based on its official language. For countries with multiple official languages, we select the most spoken one. However, if English is among the official languages, we prioritize English. For countries without an official language (e.g., the United States prior to March 1, 2025), we use the most widely spoken language. The language similarity measure is assigned a value of 1 if the focal country and a given other country share the same defined language and 0 otherwise.

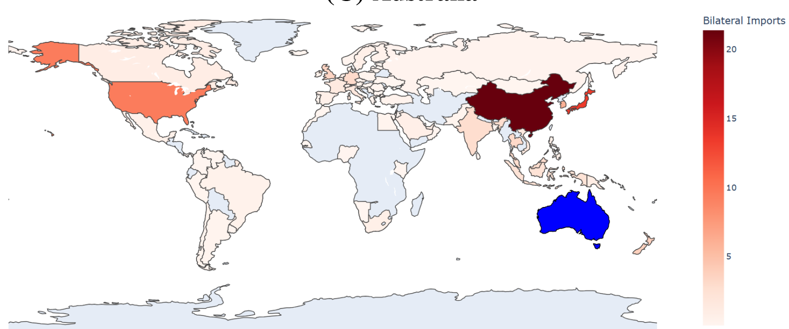
(A) South-Africa



(B) Argentina



(C) Australia



(D) United Arab Emirates

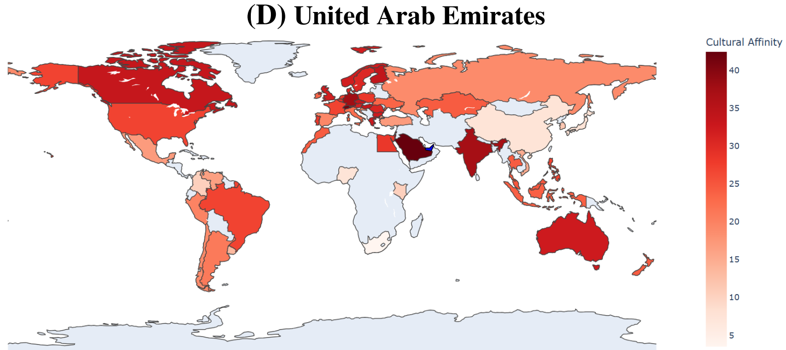


Figure B.6: Economies ties between countries This figure depicts economic ties between a focal country (in blue) and other countries, represented as the share of total bilateral imports. Four focal countries are considered: (i) South Africa, (ii) Argentina, (iii) Australia, and (iv) the United Arab Emirates. For each focal country–partner country pair, total bilateral imports are defined as the sum of the focal country's imports from and exports to the other country. To assess the relative importance of each trade relationship, we divide the country-pair's total bilateral imports by the focal country's total bilateral imports across all of its trade partners. The data, sourced from the OECD, cover the period 2000–2021. For each country pair, we compute the average share over this entire period. A value of 0 indicates no trading relationship on average, while a value of 1 signifies that the focal country traded exclusively with that single trade partner.

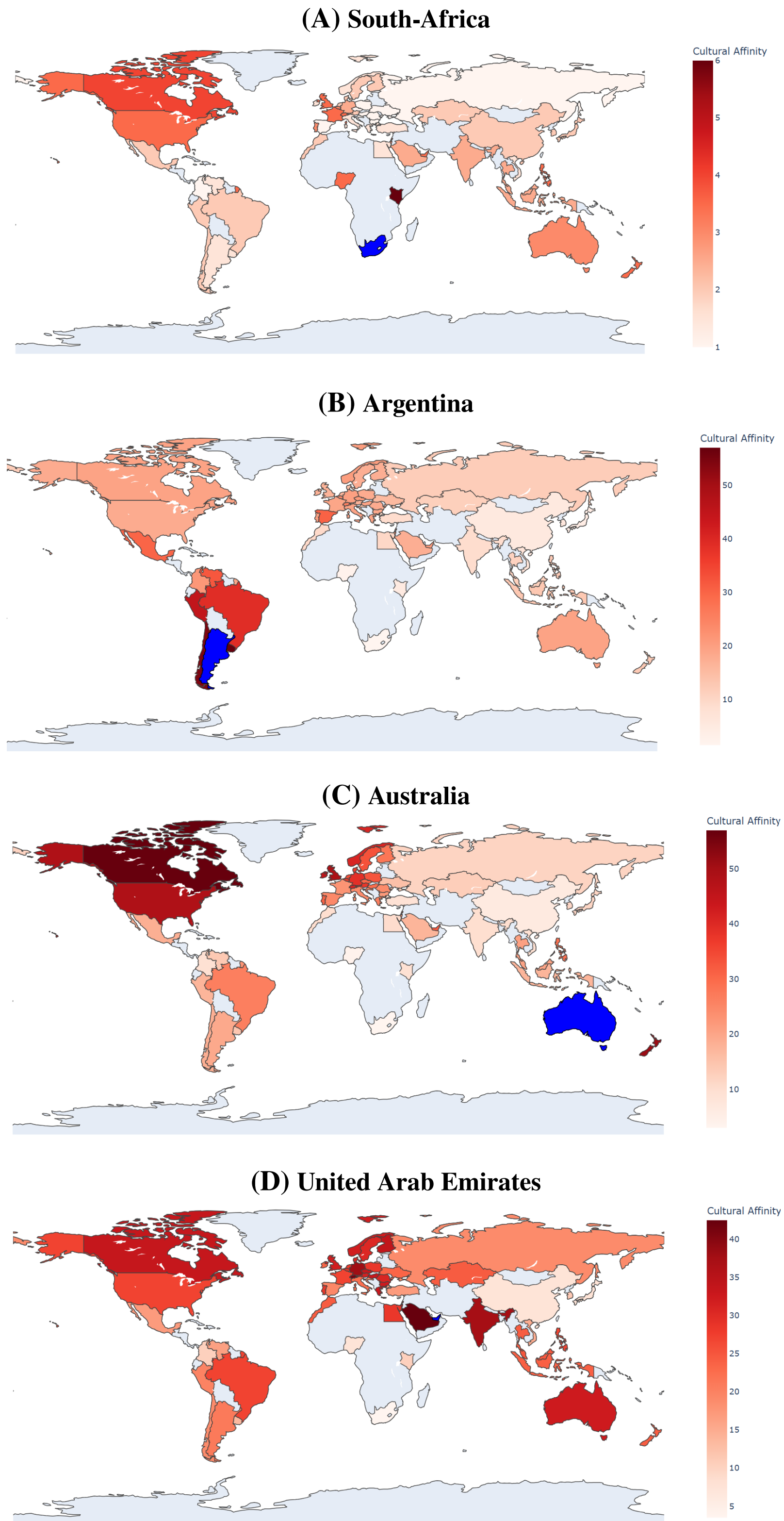


Figure B.7: Cultural affinity between countries This figure depicts cultural affinity between a focal country (in blue) and other countries, measured by the degree of overlap in music consumption. Four focal countries are considered: (i) South Africa, (ii) Argentina, (iii) Australia, and (iv) the United Arab Emirates. Music consumption overlap is measured using each country's Shazam Top 200 Chart for the week of December 17, 2024. The resulting index is normalized to range from 0 to 100, where 0 indicates no overlap in musical content and 100 indicates a perfect overlap.

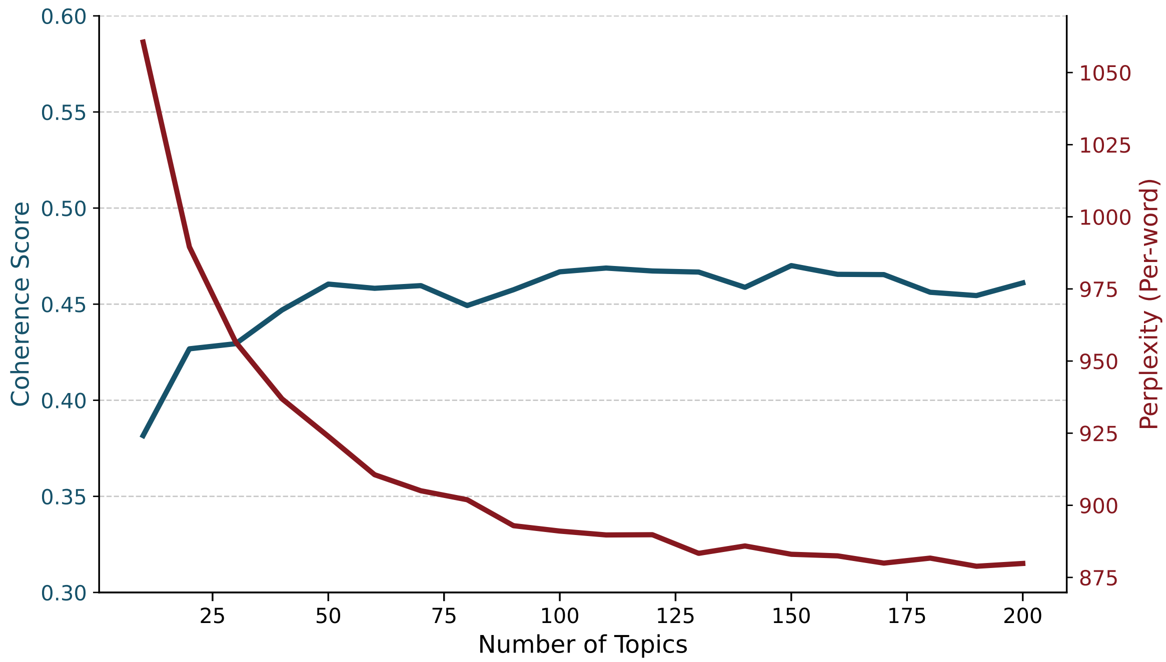


Figure B.8: Coherence and perplexity scores This figure plots the coherence and in-sample perplexity scores for different numbers of topics used in the LDA model estimation.

Appendix Tables

Table B.1: Aggregate informativeness of TGR expectations This table replicates the results presented in Table 2, but uses aggregated measures of TGR expectations and earnings growth (i.e., aggregating across firms in each year). The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable	Average $(\Pi_{t+3}^{t+8}(1+g_{j,t})^{(1/5)}-1)$	
	(1)	(2)
Average terminal growth rate (%)$_t$	4.96***	6.30***
	(1.19)	(1.73)
Average LTG (%)$_t$		-0.18
		(0.17)
Observations	14	14
F Statistics	17.29	7.84
R^2	0.46	0.49

Table B.2: Variance decomposition of TGR expectations (OLS) This table presents the explained variation in OLS regressions of TGR expectations on brokerage house, analyst, and firm fixed effects. Panel A replicates the specification in Column (1) of Panel A from Table 4 (i.e., full sample analysis, excluding firms, analysts, and brokerage houses with fewer than 2 observations in total, as their respective fixed effects fully account for the variation in TGR expectations). Panel B excludes brokerage houses with fewer than 2 observations, firms with fewer than 10 observations, and analysts with fewer than 50 observations, replicating the specification in Column (4) of Panel B in Table 4. Both panels also exclude brokerage houses that are not associated with equity analysts who switched employers during the sample period (2000–2023), as it is not possible to separate the effects of these brokerage houses from those of their analysts (Abowd, Kramarz, and Margolis, 1999). Variable definitions appear in Appendix A.

Panel A: Column 1 of Panel A in Table 4	Terminal growth rate$_{i,j,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Brokerage house FE	Yes	No	No	Yes	Yes	No	Yes
Analyst FE	No	Yes	No	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	Yes	Yes	Yes
Observations	36,544	36,357	36,367	36,357	36,367	36,094	36,094
R^2	0.02	0.47	0.45	0.48	0.47	0.65	0.66
Adjusted R^2	0.02	0.42	0.37	0.42	0.38	0.55	0.55
Panel B: Column 4 of Panel B in Table 4	Terminal growth rate$_{i,j,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Brokerage house FE	Yes	No	No	Yes	Yes	No	Yes
Analyst FE	No	Yes	No	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	Yes	Yes	Yes
Observations	5,614	5,614	5,486	5,614	5,486	5,486	5,486
R^2	0.12	0.41	0.56	0.43	0.59	0.65	0.66
Adjusted R^2	0.12	0.40	0.49	0.41	0.52	0.59	0.60

Table B.3: Accounting for TGR horizons This table presents the explained variation in OLS regressions of TGR expectations on brokerage house, analyst, firm, and TGR horizon fixed effects, for the subsample for which the information on TGR horizon is available. The TGR horizon denotes the forecast year at which terminal growth begins in the analyst's reported valuation model. Variable definitions appear in Appendix A.

Dependent Variable	Terminal growth rate $_{i,j,t}$				
	(1)	(2)	(3)	(4)	(5)
Brokerage house FE	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes
Horizon FE	No	No	Yes	No	Yes
Analyst FE	No	No	No	Yes	Yes
Observations	4,953	4,953	4,953	4,953	4,953
R^2	0.07	0.63	0.64	0.77	0.78

Table B.4: Dispersion in TGR expectations This table presents the results of the dispersion of TGR expectations over firms' life cycle. The unit of observation is at the firm i, analyst j, and forecast year t level. The sample period is 2000–2023. The dependent variable is the squared difference between an analyst's TGR belief and the industry (k) consensus for the firm's age group in forecast year t. The independent variables are indicator variables equal to one if the firm's age falls within the specified range and zero otherwise. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dispersion in TGR Expectations	$(TGR_{i,j,t} - \overline{TGR}^{\text{Age Group}}_{k,t})^2$		
	(1)	(2)	(3)
5 <= Firm age < $10_{i,j,t}$	-0.02	-0.02	-0.04
	(0.05)	(0.05)	(0.05)
10 <= Firm age < $15_{i,j,t}$	-0.03	-0.03	-0.06
	(0.06)	(0.06)	(0.06)
15 <= Firm age < $20_{i,j,t}$	-0.14**	-0.13**	-0.17***
	(0.06)	(0.06)	(0.07)
20 <= Firm age < $25_{i,j,t}$	-0.11	-0.11	-0.17**
	(0.07)	(0.07)	(0.07)
25 <= Firm age < $30_{i,j,t}$	-0.14*	-0.13	-0.18**
	(0.08)	(0.08)	(0.08)
30 <= Firm age < $35_{i,j,t}$	-0.43***	-0.45***	-0.57***
	(0.08)	(0.08)	(0.08)
35 <= Firm age < $40_{i,j,t}$	-0.54***	-0.55***	-0.73***
	(0.08)	(0.08)	(0.08)
40 <= Firm age < $45_{i,j,t}$	-0.61***	-0.60***	-0.76***
	(0.12)	(0.11)	(0.12)
45 <= Firm age < $50_{i,j,t}$	-0.86***	-0.82***	-0.98***
	(0.09)	(0.10)	(0.10)
50 <= Firm age < $55_{i,j,t}$	-0.79***	-0.77***	-0.86***
	(0.12)	(0.12)	(0.13)
55 <= Firm age < $60_{i,j,t}$	-0.99***	-0.96***	-1.14***
	(0.09)	(0.10)	(0.12)
60 <= Firm age < $65_{i,j,t}$	-0.59***	-0.61***	-0.80***
	(0.18)	(0.17)	(0.19)
65 <= Firm age < $70_{i,j,t}$	-1.10***	-1.05***	-1.19***
	(0.11)	(0.10)	(0.12)
70 <= Firm age < $75_{i,j,t}$	-0.64***	-0.71***	-0.93***
	(0.19)	(0.19)	(0.20)
75 <= Firm age < $80_{i,j,t}$	-0.57***	-0.62***	-0.46
	(0.20)	(0.23)	(0.32)
80 <= Firm age < $85_{i,j,t}$	-0.92***	-1.00***	-1.13***
	(0.19)	(0.29)	(0.34)
85 <= Firm age < $90_{i,j,t}$	-1.06***	-0.90***	-0.96***
	(0.16)	(0.18)	(0.20)
90 <= Firm age < $95_{i,j,t}$	-0.72***	-0.81***	-0.93***
	(0.25)	(0.23)	(0.23)
95 <= Firm age < $100_{i,j,t}$	-1.26***	-1.23***	-1.34***
	(0.06)	(0.08)	(0.11)
Firm $\text{Size}_{i,t}$			0.06***
			(0.01)
$\text{TGR}_{i,j,t}$			0.04
			(0.03)
Constant	1.43***	1.43***	-0.01
	(0.04)	(0.04)	(0.25)
Year*Industry FE	No	Yes	Yes
H_0 = All Age coefficients are equal (Bonferroni p-value)	0.00	0.00	0.00
Observations	24,046	24,020	21,105
F Statistics	44.26	23.29	18.83
R^2	0.01	0.06	0.07

Table B.5: Decomposition of analyst fixed effects (OLS) This table presents the explained variation in OLS regressions of the analyst fixed effects estimated in Table 6 on analyst characteristics. Panel A excludes firms, analysts, and brokerage houses with fewer than 2 observations in total, as their respective fixed effects fully explain the associated variation in terminal growth rates. Panel B presents robustness excluding cases with fewer than five observations. Both panels also exclude brokerage houses that are not associated with equity analysts who switched employers during the sample period (2000–2023), as it is not possible to separate the effects of these brokerage houses from those of their analysts (Abowd, Kramarz, and Margolis, 1999). The unit of observation is at firm i, analyst j, and forecast year t level. Gender is a binary variable equal to 1 if the analyst is female and 0 otherwise. Race is a categorical variable indicating White, Asian, Hispanic, and other races. Education is a categorical variable indicating a master's degree, MBA, or other degrees. Birth cohort is a series of birth year indicators. Analyst country is a categorical variable indicating an analyst's country at the time of producing the report. Variable definitions appear in Appendix A.

Panel A: Column 1 of Panel B in 6	Analyst FE$_j$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender FE	Yes	No	No	No	No	Yes	Yes
Education FE	No	Yes	No	No	No	Yes	Yes
Race FE	No	No	Yes	No	No	Yes	Yes
Birth cohort FE	No	No	No	Yes	No	Yes	Yes
Analyst Country FE	No	No	No	No	Yes	No	Yes
Observations	844	844	844	844	844	844	844
R^2	0.00	0.00	0.01	0.09	0.08	0.10	0.17
Panel B: Column 6 of Panel B in 6	Analyst FE$_j$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender FE	Yes	No	No	No	No	Yes	Yes
Education FE	No	Yes	No	No	No	Yes	Yes
Race FE	No	No	Yes	No	No	Yes	Yes
Birth cohort FE	No	No	No	Yes	No	Yes	Yes
Analyst Country FE	No	No	No	No	Yes	No	Yes
Observations	289	289	289	289	289	289	289
R^2	0.00	0.01	0.00	0.12	0.12	0.13	0.29

Table B.6: Trust toward other countries and TGR expectations This table presents the relationship between analysts' trust toward other cultures and their TGR expectations about firm located in these countries. The unit of observation is at the firm i, analyst j, and forecast year t level. The sample period is 2000–2023. The independent variable, $Trust_{c,k}$, is equal to the value of the measure of "trust" averaged across all the participants in the data from the Eurobarometer survey "ZA Study Number" 2935, which was collected in 1997. The sample used in the regressions covers all firm HQ-country analyst-country pairs that are included in the survey. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the firm level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable	$\text{TGR}_{i,j,t}$	
	(1)	(2)
$\text{Trust Index}_{c,k}$	0.42***	0.42***
	(0.11)	(0.15)
Analyst * Year FE	No	Yes
Firm * Year FE	No	Yes
Observations	13,194	3,559
F Statistics	13.84	7.91
R^2	0.00	0.89

Table B.7: Representative words from LDA topics This table reports the key words representing LDA topics. Words are selected based on reversed frequency weights, where a word's weight in a given topic is scaled by its relative frequency in the texts, similar to the strategy in Bybee, Kelly, and Su (2023).

Category	Topic Name	Word List
Geography	Asia	rmb, thailand, hong_kong, korea, china, overseas, singapore, india, chinese, domestic
Geography	Europe	spanish, italian, spain, italy, france, french, russian, germany, german, russia
Geography	Oceania	australian, new_zealand, australia, half, earning, forecast, period, earn, result
Geography	South-America	mexican, brazilian, argentina, brazil, chile, mexico, latam, inflation, country
Industry	Health Care	hospital, implant, medicare, pharmacy, diagnostic, medical, care, healthcare, reimbursement, procedure
Industry	Oil and Gas (Transformation)	refining, petrochemical, refinery, feedstock, crude_oil, gasoline, crude, bbl, downstream, oil
Industry	Oil and Gas (Production)	rig, boe, drilling, exploration, field, discovery, reserve, natural_gas, oil_gas, oil
Industry	Manufacturing	equipment, industrial, manufacturing, manufacturer, chemical, system, supplier, component, technology, application
Industry	Entertainment	programming, broadcast, newspaper, television, film, music, radio, audience, entertainment, advertising
Industry	Resources	ethanol, pulp, farmer, sugar, crop, grain, cigarette, corn, seed, fertilizer
Industry	Automobile	battery, dealer, car, truck, vehicle, auto, engine, automotive, man, unit
Industry	Banking	loan, lending, bank, mortgage, banking, insurance, deposit, fund, bond, roe
Industry	Carrier	satellite, tower, fiber, cable, video, carrier, network, voice, speed, broadband
Industry	Communication	arpu, wireline, mobile, subscriber, telecom, churn, broadband, wireless, sub, operator
Industry	Construction	non_residential, toll_road, cement, housing, railway, concession, construction, road, contractor, building
Industry	Consumer goods (Clothing)	apparel, fashion, luxury, brand, wholesale, sport, distributor, channel, inventory, retail
Industry	Consumer goods (Food)	beverage, spirit, tobacco, beer, wine, category, innovation, brand, food, consumer
Industry	Energy	renewable, grid, mwh, hydro, solar, electricity, power_plant, wind, nuclear, utility
Industry	Gaming	gaming, search, game, online, commerce, website, internet, user, platform, marketing
Industry	IT	cloud, software, sap, vendor, solution, enterprise, subscription, application, license, platform
Industry	Mining (Metal Other)	gold, mineral, zinc, south_african, copper, ore, nickel, mining, south_africa, grade
Industry	Mining (Steel)	potash, aluminium, steel, iron_ore, aluminum, coal, tonne, ton, import, producer
Industry	Pharmaceutical (R&D)	generic, pharma, pharmaceutical, patent, drug, fda, royalty, launch, franchise, pipeline
Industry	Pharmaceutical (Clinical trial)	phase_trial, dose, efficacy, therapy, cancer, disease, patient, trial, clinical, study
Industry	Real Estate	reit, tenant, occupancy, property, hotel, rent, rental, real_estate, land, nav
Industry	Retail (Other)	supermarket, store, restaurant, department_store, retailer, format, food, retail, chain, comp
Industry	Semi-Conductor	dram, chip, semiconductor, panel, handset, shipment, tech, device, design, technology
Industry	Shipping	container, vessel, parcel, port, mail, ship, rail, tanker, freight, shipping
Industry	Transportation	airline, aircraft, cruise, passenger, airport, aviation, flight, defense, travel, toll
Macroeconomics	Legal Risk	court, litigation, settlement, legal, card, claim, merchant, affiliate, payment, information
Macroeconomics	Economics Cycles	economy, gdp, index, economic, cycle, recovery, global, weak, sector, decline
Macroeconomics	Regulation	reform, proposal, regulator, regulation, regulatory, government, tax, review, state, policy

Continued on next page.

Continued from previous page.

Category	Topic Name	Word List
Operations	Outsourcing	outsourcing, client, service, contract, employee, provider, revenue, business, segment, acquisition
Operations	Demand	volume, pricing, raw_material, demand, capacity, supply, pressure, increase, recovery, cost
Operations	Expenditure and Disbursement	dps, dividend, debt, buyback, capex, dividend_yield, cash, fcf, balance_sheet, ebitda
Operations	Management	officer, education, student, school, director, executive, president, esg, cfo, ceo
Operations	Merge and Acquisition	merger, bid, synergy, deal, stake, transaction, shareholder, offer, acquisition
Operations	Organization	disposal, restructuring, division, group, trading, european, profit
Operations	Project Management	waste, project, water, backlog, delay, pipeline, facility, contract, financing, funding
Valuation	Valuation facts	fact, different, reason, important, big, time, long, case, point, issue
Valuation	Accounting Norms and Model Selection	ambiguity_local, modelware_separation, accounting_regulation, value_distortion, financing_complete, fifo_basis, inventory_lifo, eps_adjust, modelware_proprietary, interest_objectivity
Valuation	Free Cash Flow	approximately, roughly, free_cash, flow, facility, cash_flow, company, capital, distribution, acquisition
Valuation	Operating Perfomance	sek, organic, ebit, ebita, currency, sale, europe, margin, effect, growth
Valuation	Profitability	profit, expense, income, loss, operating, gross, charge, net, gain, tax
Valuation	Firm growth	covid, headwind, bps, incremental, leverage, fcf, expansion, organic, upside, margin
Valuation	Recommendation	outperform, neutral, catalyst, recommendation, stock, momentum, buy, rating, earning, earn
Valuation	Guidance	eur, usd, fcf, ebit, slightly, guidance, positive, estimate, good, development
Valuation	Consensus	guidance, street, consensus, estimate, eps, prior, revenue, slightly, expectation, ahead
Valuation	Relative Valuation	multiple, fair_value, historical, peer, analysis, average, discount, relative, scenario, value
Valuation	Risk	beta, wacc, risk, terminal, downside, free, dcf, investment_thesis, hold, unchanged

Table B.8: TGR disagreement and textual discussions (Alternative topic model with 80 topics) This table presents the relationship between differences in textual discussion and disagreement in TGR expectations between pairs of analysts, using an alternative LDA model with 80 topics. The unit of observation is at the analyst a, analyst b, firm i, and forecast year t level. The sample period is 2000–2023. The dependent variable, $Disagreement_{a,b,i,t}$, is defined as the squared difference in TGR estimates between analyst a's and analyst b's TGR beliefs: $(TGR_{a,i,t} - TGR_{b,i,t})^2$. The key independent variables of interest are the Jaccard and Cosine similarity scores, which capture the similarity of the analysts' textual discussion on the extensive and intensive margins, respectively. Variable definitions appear in Appendix A. The regressions are estimated using ordinary least squares, and the standard errors (in parentheses) are heteroskedastic-consistent and clustered at the analyst-pair level. Significance levels are shown as follows: * = 10%, ** = 5%, *** = 1%.

Dependent variable	Disagreement$_{a,b,i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Jaccard $^{80}_{a,b,i,t}$	-0.19	-0.33	-0.30				-0.31
	(0.28)	(0.28)	(0.28)				(0.28)
Cosine $^{80}_{a,b,i,t}$				-0.53***	-0.49**	-0.37*	-0.37*
				(0.20)	(0.20)	(0.20)	(0.20)
Year FE	No	Yes	No	No	Yes	No	No
Industry FE	No	Yes	No	No	Yes	No	No
Industry*Year FE	No	No	Yes	No	No	Yes	Yes
Observations	23,466	23,433	23,400	23,463	23,430	23,397	23,397
F Statistics	0.46	1.43	1.19	6.70	6.01	3.56	2.51
R^2	0.00	0.02	0.07	0.00	0.02	0.07	0.07

Table B.9: Cosine similarity across different industries This table reports the Cosine similarity scores for different pairs of firms. The unit of observation is at the analyst a, analyst b, firm i, and forecast year t level. The sample period is 2000–2023.

	Cosine similarity$_{a,b,i,t}$	
	Different First NAICS Digit (1)	Same Firm (2)
Mean	0.78	0.81
Median	0.82	0.85
Standard Deviation	0.20	0.15
25th Percentile	0.64	0.73
75th Percentile	0.95	0.93
Number of Observations	272,359,714	162,160