The Human Edge Beyond Algorithms: Forecasting in Global Macro Shocks

Yuhan Ye

July 30, 2025
[Click here for the latest version]

Preliminary draft
Please do not circulate.

Abstract

This paper examines the enduring value of human judgment in an era of increasingly powerful AI. Focusing on over 200 macroeconomic shock episodes across 47 countries, I investigate when and where analysts retain an advantage over algorithmic models. I use machine learning trained on public data to construct benchmark forecasts for earnings expectations and decompose the gap between human and machine forecasts into soft information, bias and noise. The results show that soft information, such as contextual and non-public insights that are not captured in public data, significantly improves human forecast accuracy, especially at the onset of macroeconomic shocks. This advantage is particularly evident in emerging markets, where limited disclosure constrains the learning capacity of algorithms. As the shock progresses, however, the accuracy of human forecasts

Ye is with the Swiss Finance Institute (SFI) at Università della Svizzera italiana (USI, Lugano). I am deeply grateful to Laurent Frésard for his guidance and support throughout this project and my doctoral studies. I thank Daron Acemoglu, Tummim Cho, William Lin Cong, François Derrien, Andrew Ellul, Ruediger Fahlenbrach, Thierry Foucault, Francesco Franzoni, Gerald Hoberg, Augustin Landier, Martin Lettau, Nadya Malenko, Frederic Malherbe, Loriano Mancini, Ernst Maug, Thomas Andreas Maurer, Roni Michaely, Alessandro Peri, Gordon Phillips, Alberto Plazzi, Alberto Rossi, Paul Schneider, Tim de Silva, David Thesmar, Victoria Vanasco, Liliana Varela, Aurelio Vasquez, Petra Vokata, Alexander F. Wagner and Anthony Lee Zhang for helpful comments and suggestions. I also thank participants at various seminars and conferences for their feedback. Email: yuhan.ye@usi.ch. Preliminary draft. All errors are my own.

declines due to increasing bias and noise. These findings underscore the conditional value of human input and the informational limits of automation under uncertainty. The analysis also reveals substantial cross-country differences related to institutional transparency, contributing to our understanding of belief formation, systemic resilience, and the interaction between human and algorithmic decision-making in global financial markets.

Keywords: Soft Information, Macro Shocks, Belief Formation, Analyst Forecasts, Econometric Forecasts, Machine Learning

1. Introduction

The rise of machine learning in financial markets has redefined the boundaries between human judgment and algorithmic precision. While algorithms increasingly dominate routine forecasting tasks such as credit scoring and high-frequency trading Fuster et al. (2019), their performance during systemic macroeconomic shocks remains contested. A critical yet unresolved question emerges: **What residual value do human analysts provide when conventional models fail to navigate turbulent markets?**

As emphasized by the Lucas Critique (Lucas 1976), statistical relationships observed in historical data may not remain valid under changing economic conditions, unless expectations and behavioral adjustments are explicitly modeled. My paper takes this critique seriously by decomposing analyst forecasts into a machine-based baseline and human-specific adjustments, including soft information, behavioral bias, and noise. This allows us to separate truly predictable patterns from expectation-driven shifts, especially around macroeconomic shocks, where adaptive behavior becomes crucial.

To operationalize this idea, I develop a framework that explicitly decomposes analyst forecasts into two components: a machine-driven baseline derived from observable, stable predictors; and human-specific adjustments that embed behavioral features such as soft information, belief distortions, and idiosyncratic noise. This decomposition allows us to identify when and how humans retain forecasting advantages particularly in domains where algorithmic predictions may falter, such as during the onset of economic shocks or in data-sparse environments.

This approach directly addresses the Lucasian concern that fixed-rule statistical models may fail when expectations adjust endogenously. By contrasting machine and human forecast components, we uncover where human judgment still adds value, offering insight into the informational limits of prediction algorithms and the contexts in which behavioral inputs

become essential.

Ultimately, the goal is to better understand the *residual edge of human forecasting*: what informational or cognitive elements analysts bring to bear that are not yet captured by machine learning algorithms, and under what conditions those elements become essential for accuracy.

A key conceptual foundation for this paper comes from the "prediction machine" framework introduced by Agrawal, Gans, and Goldfarb (2022), which mentions a 2×2 matrix of uncertainty originally articulated by former U.S. Secretary of Defense Donald Rumsfeld. This matrix categorizes uncertainty based on whether relevant information is both identifiable and measurabledividing the world into known knowns, known unknowns, unknown knowns, and unknown unknowns.

	Known	**Unknown**
Known	Known Knowns	Known Unknowns
Unknown	Unknown Knowns	Unknown Unknowns

This classification provides a useful lens for evaluating the relative strengths of machine-based and human forecasts. Prediction models are most effective when operating on "known knowns," where relationships are stable and the data environment is rich and structured. They may also provide reasonable performance in "known unknown" settings, where risks are identifiable but difficult to quantify, though their reliability may decline when data are sparse or unstable.

Human forecasters can complement algorithms by incorporating "unknown knowns," such

as tacit knowledge, qualitative assessments, and contextual cues not captured in observable features. This category includes soft information derived from institutional knowledge, experience, or access to management commentary. In contrast, "unknown unknowns" represent unanticipated disruptions or regime shifts that neither humans nor machines can easily forecast.

This paper formalizes these insights by training machine learning models to capture the "known known" component of forecasts, and treating the analyst-specific deviation as a human adjustment. The residual component can be further decomposed into soft information, behavioral bias, and noise. This structure allows us to assess when human forecasts diverge from machine benchmarks and what those deviations reveal about the value of human judgment in uncertain environments.

This paper provides the first international analysis of cross-country variation in the relative forecasting performance of humans and machines, using data from 47 countries and more than 200 macroeconomic shocks. By leveraging variation along temporal, forecast horizon, and regional dimensions, I document systematic differences in when and where machines outperform humans, and vice versa. A central contribution of this study is to highlight the role of public data availability in shaping these dynamics: machine learning models tend to dominate in information-rich environments, whereas human analysts retain a comparative advantage where soft information helps bridge public information gaps.

In this paper, I quantify the role of soft information, defined as contextual and non-quantifiable insights embedded in human expectations, during global macro shocks. My analysis builds

on the structural decomposition framework of De Silva and Thesmar (2024), which separates the non-statistical components of human forecasts into bias, noise, and soft information. While their study emphasizes noise using U.S. data without distinguishing economic regimes, I extend this framework by leveraging a novel international dataset spanning 47 countries and over 200 macro shock events. This global perspective increases the coverage of soft information to include hard public data obscured by access barriers in smaller economies, legally soft information from regulatory nuances, and insider knowledge beyond public reach. By comparing human-machine performance gaps across macro shock and non-macro shock periods, I reveal how institutional heterogeneity and economic turbulence amplify the value of this expanded soft information, a dimension overlooked in single-country studies. These findings provide fresh insights into the comparative advantages of human analysts over machine learning models in financial forecasting, with significant implications for theory and practice.

I implement a two-step structural framework to isolate the unique value of human analysts, following established practices in the literature. In the first stage, I use supervised learning algorithms to predict both realized earnings and analyst forecasts using all contemporaneously available public information, thereby isolating the statistically replicable component. This approach builds on recent work by Van Binsbergen, Han, and Lopez-Lira (2020) and Cao et al. (2021), who demonstrate the effectiveness of machine learning in modeling earnings expectations and stock analyses. In the second stage, I decompose the residual non-statistical component into three latent factors using the structural estimation method proposed by De Silva and Thesmar (2024): soft information, systematic bias, and idiosyncratic noise. This approach contributes to the literature in three key ways. First, it provides a direct quantification of soft information's economic value, moving beyond reduced-form proxies such as textual sentiment scores Tetlock (2007). Second, it demonstrates that soft information's importance is state-dependent, peaking during macro shocks when algorithmic predictability collapses Gabaix and Koijen (2020). Finally, it documents how institutional factors such as regulatory transparency moderate the human-machine performance gap.

The earnings forecasts datasets provide an excellent opportunity to analyze human judgment in three respects. First, professional equity analysts operate under strong reputational pressures (Hong and Kubik 2003a) and benefit from access to exclusive information channels (Groysberg, Lee, and Nanda 2011), which enables them to interpret qualitative signals beyond the reach of algorithms. Second, the global standardization of earnings forecasts across 47 countries offers a consistent measure to examine how institutional contexts shape macro shock interpretations. Third, because earnings expectations directly influence equity valuations through discounted cash flows (Campbell and Shiller 1997), revisions made during macroeconomic shocks reveal how analysts reframe disruptions that machines tend to misprice. When historical patterns break down, analysts reassess fundamentals using evolving macro shock narratives that algorithms fail to capture.

By integrating statistical methodologies such as machine learning algorithms into an extensive real-world dataset, I move beyond testing these models on simulated data or solely on U.S. markets. Because the United States represents a highly developed financial environment, my international analysis, which spans both developed and developing economies and covers efficient as well as semi-efficient markets, offers unique insights into variations in noise, bias, and non-statistical information. In particular, this approach expands the measurable range of soft information to include not only qualitative insights but also public data that is not readily accessible, legally soft disclosures, and insider information. These variations, in turn, shed light on the evolving nature of learning and prediction accuracy as economies transition from developing to developed status.

My analysis yields three central results. First, during the onset of macro shocks, human analysts significantly outperform machine forecasts in short term earnings predictions because they leverage soft information that algorithms cannot access. For example, when forecasting earnings 30 days ahead for a firm such as Tesla, analysts incorporate timely qualitative cues, whereas for long term forecasts, the human advantage diminishes as soft

information becomes scarce and human forecasts suffer from emotional bias and noise.

Second, my empirical findings indicate that at the beginning of a macro shock, the substantial human edge is driven by a strong soft information advantage. However, as the macro shock evolves, this edge gradually erodes because analysts' forecasts become increasingly affected by bias and noise. For instance, during the early stage of the 2020 pandemic, analysts predicted a swift resolution and outperformed algorithms, but as the macro shock persisted and sentiment turned pessimistic, machine forecasts eventually surpassed human predictions.

Third, the contribution of soft information to forecast accuracy varies markedly across countries. In developing economies, where data availability is lower and legal systems are less robust, soft information accounts for a substantially higher share of forecast accuracy gains, reflecting a greater reliance on non-statistical insights that conventional models fail to capture. This pattern aligns with institutional theories that predict weaker governance amplifies both human adaptability and cognitive fragility during disruptions. Moreover, the learning curve in developing countries is markedly steeper; at the onset of a macro shock, these markets rely more heavily on soft information, while bias and noise intensify more rapidly as the macro shock progresses, leading to greater fluctuations in human expectations compared to developed economies.

Overall, these findings deepen our understanding of belief formation by highlighting the critical role of soft information in navigating global macro shocks. This evidence challenges the inevitability of automation, emphasizing the need for forecasting models tailored to diverse institutional settings and economic regimes, where human judgment remains essential to interpret phenomena beyond algorithms' reach.

The structure of the paper is as follows: Section 1 introduces the research question, key findings, main contributions, and related literature. Section 2 describes the data sources, the

construction of forecasting targets and predictors, and the machine learning methodologies used in the analysis. Section 3 presents the main empirical results. It benchmarks human and machine forecasts, decomposes forecast errors into soft information, bias, and noise, and examines cross-country heterogeneity between developed and emerging markets. Section 4 discusses several robustness checks and model extensions, including the use of large language models (LLMs) to extract soft information from text data. Rather than serving as standalone prediction algorithms, LLMs are used to generate additional features that enhance machine-based forecasts, allowing for a richer integration of textual signals into the forecasting framework. Section 5 concludes the paper and outlines directions for future research. Supplementary results, derivations, and additional figures are reported in the Appendix.

Literature Review

This study contributes to the lit erature on expectation formation (e.g.,Sims (2003); Woodford (2003); Landier, Ma, and Thesmar (2017); De Silva and Thesmar (2024)) and the role of noise expectation in forecasting and human decision-making across various domains such as medicine, finance, hiring, and judicial decisions (Kahneman, Sibony, and Sunstein 2021). Subjective forecast noise has been extensively discussed in the literature on noisy information (e.g., Woodford 2003; Coibion and Gorodnichenko 2015) and behavioral economics (e.g., Khaw, Li, and Woodford 2019; Woodford 2020; Enke and Graeber 2020; Kahneman et al. 2021; Afrouzi et al. 2021).

My contribution to this literature is twofold. First, I provide evidence on the size and term structure of noise using analyst forecast data. Second, our methodology places no restrictions on the data-generating process. This approach is similar to that of Satopää, Salikhov, Tetlock, and Mellers (2020), who perform a bias–information–noise (BIN) decomposition and find that noise reduction is a consistent property of good subjective forecasters. It is also

complementary to the approach developed by Juodis and Kucinskas (2019), which exploits the factor structure in expectations implied by many models of belief formation. More broadly, our methodology relates to the work of Bianchi, Ludvigson, and Ma (2020) and Nagel (2021), who discuss how supervised learning is useful for studying subjective expectations data.

Furthermore, I extend Thesmar's methodology by testing this decomposition approach on global datasets, thereby connecting it to a broader range of contexts. Additionally, I engage with the literature that uses machine learning algorithms as benchmarks for studying human forecasts. This paper contributes to the literature of studying analyst forecasts, together with stattical forecasts as benchmark. Van Binsbergen, Han, and Lopez-Lira (2020) compares analyst forecasts with machine's forecasts, and define the difference as conditional bias in firms' earnings forecasts. The authors also show the term structure of this new measure. De Silva and Thesmar (2024) decompose the conditional bias into soft information, bias and noise of human compared to machine and analyze its term structure. This paper also generates a theoretical framework to account for the empirical facts. Ball and Ghysels (2018) uses mixed data sampling (MIDAS) regression methods to utilize high frequency data when constructing forecasts. Cao et al. (2021) studies human's capacity vs machine in forecasting stock prices.

Besides, this paper contributes to the literature on forecasting accuracy during macro shock. Our study also continues the line of research on the informativeness of financial forecasts and the literature on forecasts during crises, such as the work by Fouliard, Howell, and Rey (2021).

This paper also adds to the topic of informativeness of financial forecasts. Dessaint, Foucault, and Frésard (2021) study the horizon effect existing in the analyst forecasts with evidence from alternative data.

This paper resonate with the governance-based view of home bias proposed by Pinkowitz, Stulz, and Williamson (2001) and Dahlquist et al. (2003), extending its logic from the domain of asset ownership to that of forecast composition. Specifically, I show that analysts are more likely to generate value when they possess privileged access to local, non-public information. This result also connects to the literature on local analyst advantage, particularly Bae, Stulz, and Tan (2008), by structurally quantifying the contribution of soft information to forecasting performance.

In doing so, my paper contributes to both strands of literature by introducing direct, quantitative measures of local informational advantage in the context of financial forecasting. These measures allow for a systematic evaluation of when and how locally embedded analysts outperform, thereby offering new empirical content to theories of home bias and local expertise.

Prior literature extensively documents the prevalence of analyst optimism, whereby analysts tend to overestimate firm earnings and stock values due to various incentives. These include the motivation to maintain access to management (Lim 2001), the desire to secure trading commissions and retain clients (Cowen, Groysberg, and Healy 2006), as well as career concerns and institutional affiliations (Hong and Kubik 2003b; Mola and Guidolin 2009), among a broader literature on analyst behavior. This optimistic bias has been identified as a form of analyst activism influencing market expectations.

Extending this understanding, I leverage international datasets to uncover a complementary pattern in noisy forecasting environments. In these markets, characterized by less transparent information and greater uncertainty, analysts exhibit systematic pessimism, consistently underestimating firm earnings and stock prices. This finding suggests that analyst biases are context-dependent and that pessimism may dominate when informational noise impedes accurate forecasting.

2. Data and Methodology

2.1. Data Sources and Coverage

2.1.1. International Coverage

The analysis leverages a novel international dataset designed to capture cross-country heterogeneity in human expectations during macroeconomic shocks. It spans 47 countries and regions, encompassing all constituents of the MSCI ACWI Index as of 2020. The sample covers approximately 85% of the global investable equity market (see Appendix Table A1), including 23 developed markets (e.g., United States, Japan, Germany) and 24 emerging markets (e.g., India, Brazil, Malaysia), ensuring broad representation across institutional and developmental contexts.

The dataset includes over 200 macroeconomic shocks from 1980 to 2025, classified into three categories:

- **Global Systemic Shocks**: Events with cross-border transmission mechanisms, such as the early 1990s recession, the 2008 Global Financial Crisis, and the COVID-19 pandemic[1].
- **Regional Shocks**: Geographically concentrated crises, including the 1997 Asian Financial Crisis and the 2010 European Debt Crisis.
- **Country-Specific Shocks**: Nationally confined disruptions, such as Argentina's 2001 sovereign default and Russia's 2014 currency crisis.

This taxonomy enables systematic examination of how shock type and geographic scope moderate the relative performance of human and machine forecasts. The sample begins in 1990 to align with the widespread adoption of digital data infrastructure and advances in computational forecasting. Shock selection criteria and the full event list are provided in

[1] Figure A1 in the Appendix illustrates selected global shocks for the United States using industrial production data.

Appendix Table A2.

2.1.2. Data Sources

This study integrates three types of publicly available data to construct a comprehensive cross-country forecasting framework:

- **Firm Fundamentals**: Over 200 firm-level variables are sourced from Compustat Global and Datastream, including balance sheet items, profitability metrics, and market prices. For stock price and market capitalization data, I use CRSP for firms in the United States and Canada, and Eikon Datastream for firms in other countries, ensuring consistent coverage across the global sample. These variables form the primary input for statistical forecast models. Comprehensive variable lists are provided in the appendix tables A3, A4, A5, and A6; see also the overview in Appendix Note E.1.

  I focus on commonly used Compustat variables to maximize data availability and minimize missing observations. This selection strategy follows the approach in Hansen and Thimsen (2021), who emphasize the importance of avoiding look-ahead bias by relying on contemporaneously observable inputs.

- **Analyst Forecasts**: I obtain analyst earnings forecasts from the I/B/E/S Global database, covering the period 1985 to 2023 and comprising more than 20 million observations across 47 countries. Fiscal-year-end earnings are used as the realized values to enhance cross-sectional comparability and address concerns related to seasonal reporting patterns (Kothari 2001). To limit the impact of extreme values, EPS figures are winsorized at the 5% level on an annual basis within each country.

- **Macroeconomic Indicators**: This study incorporates over 30 monthly country-level indicators from 1985 to 2023, sourced from Trading Economics and FactSet. The variables include GDP growth, inflation, unemployment, the Industrial Production Index, and the Consumer Price Index, among others. These indicators characterize the macroeconomic environment in which forecasts are formed and serve as key inputs for modeling earnings dynamics across countries. Their inclusion is motivated by the asset pricing literature emphasizing the role of macroeconomic conditions in shaping firm valuation and return predictability (Fama and French 1989; Chen, Roll, and Ross 1986). A complete list of variables is provided in the Appendix Table A7.

These three data sources are merged at the firm-country-date level to create a unified panel suitable for machine learning-based forecast modeling and structural decomposition.

To train machine learning models, I construct two versions of statistical inputs. The first relies solely on public financial and macroeconomic data, excluding any information about analyst forecasts. This version allows us to build fully independent benchmark forecasts that simulate algorithm-only predictions. The second version incorporates analyst forecasts as additional features, enabling us to explore potential complementarities between human and machine inputs in predictive performance.

2.1.3. Analyst Forecast Processing

The analyst forecasts from I/B/E/S provide a high-quality global panel of human judgment under uncertainty. Sell-side equity analysts are highly trained professionals with strong reputational incentives (Hong and Kubik 2003a), and they issue firm-level earnings forecasts across 47 countries, including 23 Developed and 24 Emerging Markets.

To ensure data quality and comparability across countries, I implement several preprocessing steps on the I/B/E/S analyst forecasts. First, EPS forecasts are winsorized at the 5% level annually within each country to mitigate the influence of outliers. Second, all forecast and earnings announcement dates are standardized using UTC timestamps. Finally, to ensure cross-country comparability, I standardize forecast dates, control for outliers, and classify forecasts into short-, mid-, and long-term horizons based on time to earnings release.

This structure allows us to systematically examine how analysts in different institutional environments, such as those with high or low information barriers, respond to the same macroeconomic shocks. This comparative design feature is absent in single-country studies.

2.2. Statistical Forecasting Framework

A key objective of this study is to compare the accuracy and composition of human and algorithmic forecasts during macroeconomic shocks. To establish a meaningful benchmark for analyst expectations, I implement a dual-model forecasting framework using supervised machine learning algorithms trained on public financial and macroeconomic data. This algorithmic forecasting framework is referred to as the ”machine analyst” throughout the analysis. This setup enables direct comparison between human and machine predictions across varying institutional, temporal, and informational contexts.

2.2.1. Benchmark Forecasting Models

Building on the approaches of Van Binsbergen, Han, and Lopez-Lira (2020), Cao et al. (2021), and De Silva and Thesmar (2024), I implement a dual-model framework in which machine learning forecasts are generated concurrently with human forecasts, using only publicly available information for the machine-based predictions. These models serve as

benchmarks for constructing forecast residuals used in the structural decomposition. A detailed description of the algorithms is provided in Appendix A.

- **Quasi-linear Models**: Lasso/Ridge/Elastic Net regression for sparse high-dimensional data:

$$\min_{\beta} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=1}^{p} \beta_j^2 \tag{1}$$

  where y_i is the target variable, x_{ij} are the input features, β_j are the model coefficients, and λ_1, λ_2 are regularization parameters controlling the strength of Lasso (L_1) and Ridge (L_2) penalties, respectively.

- **Non-linear Models**: such as Random Forests, Gradient-Boosted Trees predict outcomes by aggregating the predictions of multiple decision trees:

$$\hat{y} = \sum_{m=1}^{M} f_m(x), \quad f_m \in \mathcal{F} \tag{2}$$

  where each f_m represents an individual regression tree, and $\mathcal{F}$ is the space of all decision trees.

  Besides standard machine learning methods, I also employ Feedforward Neural Networks (FNN) to generate financial forecasts, allowing for nonlinear patterns and complex interactions among input features to be captured.

In addition to these classical supervised learning methods, I incorporate large language models (LLMs) to extract soft information from text data and generate supplementary features that enhance machine-based forecasts. This extension is discussed in detail in Section 4.

2.2.2. Forecast Horizons

Following Dessaint, Foucault, and Frésard (2021), forecast horizons are computed as the number of calendar days between the forecast date and the earnings release date:

$$\text{Horizon} = \text{Earnings Release Date} - \text{Forecast Date}$$

For example, if an analyst issues an earnings forecast for Tesla on January 1st, and the actual earnings are released on March 31st, the forecast horizon is 89 calendar days.

This continuous measure provides a more precise indication of the time remaining until the realization of the target variable, compared to the traditional use of the `FPI` (Forecast Period Indicator) in I/B/E/S datasets, which only reflects the fiscal period being forecasted. It better captures the actual temporal distance faced by forecasters, either human or machine, at the time of prediction.

Horizons are classified into three categories:

- **Short-term**: $<$ 1 year
- **Mid-term**: 1–2 years
- **Long-term**: $>$ 2 years

This categorization captures how analysts update beliefs across different macro shock phases while maintaining comparability with machine learning predictions.

As an illustrative example, Figure A2 presents the distribution of forecast horizons for analyst forecasts in the United Kingdom. The figure is based on over 700,000 analyst forecasts from the United Kingdom, with horizons measured in calendar months and rounded to the nearest whole number.

2.2.3. Structural Model

Building on De Silva and Thesmar (2024), I develop a structural decomposition of analyst forecast errors into three economically meaningful components: soft information, bias, and noise. I define the residual forecast error as the difference between analyst expectations and a benchmark prediction conditional on observable financial and macroeconomic variables X_t. The structural form is given by:

$$F_{ij} = x_i + z_i + b_{ij} + \eta_{ij} \tag{3}$$

where F_{ij} denotes the forecast by analyst j for firm i, $x_i = E[\pi_i \mid X_i]$ is the model-implied benchmark based on observable information, z_i captures firm-level soft information available to analysts but not to the model, b_{ij} represents analyst-specific bias, and η_{ij} is idiosyncratic noise.

This decomposition is grounded in the intuition that these components influence forecasts in distinct and empirically separable ways. For instance, if two analysts issue similar deviations from the benchmark for the same firm, the pattern likely reflects common soft signals. If one analyst consistently overpredicts regardless of fundamentals, this suggests bias. If residuals vary without relation to either fundamentals or outcomes, they are attributed to noise. Exploiting variation across analysts, horizons, and macroeconomic states enables identification of these latent components.

- **Soft Information**: Context-specific insights unavailable to algorithms. This includes:

  - *Unstructured public data* (e.g., narrative disclosures, local news),
  - *Institutional signals* (e.g., timing of earnings guidance or local enforcement norms),
  - *Private information* (e.g., management tone, site visits),
  - *Human judgment*, especially under uncertainty.

- **Bias**: Predictable deviations unrelated to information advantages. These often arise from behavioral tendencies (e.g., optimism or conservatism) or institutional incentives (e.g., affiliation pressures, career concerns).
- **Noise**: Unsystematic forecast variation caused by information frictions (Woodford 2003), inattentiveness, or bounded rationality (Landier, Ma, and Thesmar 2017).

I implement the decomposition in two stages. First, I estimate benchmark forecasts for both realized earnings and analyst expectations using supervised machine learning models trained on public financial and macroeconomic data. Second, I compute residuals and separate them into the three structural components. This empirical strategy builds on the identification design of De Silva and Thesmar (2024), originally applied to U.S. forecasts under normal economic conditions. I extend their approach to a global setting covering 47 countries and over 200 macroeconomic shocks, explicitly distinguishing between crisis and non-crisis periods. This setting introduces rich cross-sectional and temporal variation that supports identification even under institutional heterogeneity.

The structural framework is particularly suited to analyzing macroeconomic shock dynamics. Such shocks reduce the reliability of statistical models trained on historical patterns while amplifying the importance of contextual cues and subjective interpretation. In this environment, human forecasts may incorporate valuable soft information that escapes algorithmic detection. The international scope of the dataset enables decomposition of forecast errors into distinct components, even in the presence of varying regulatory and informational environments.

2.2.4. GMM Estimation Strategy

The parameters of interest are identified using Generalized Method of Moments (GMM), following the approach of De Silva and Thesmar (2024). The estimation is based on three moment conditions derived from residualized forecasts and outcomes:

- **Moment 1**: $\text{Cov}(\pi_i^*, F_{ij}^*) = \alpha\Theta$
- **Moment 2**: $\text{Var}(F_{ij}^*) = \alpha^2\Theta + \Sigma$
- **Moment 3**: $\text{Cov}(F_{ij}^*, F_{ik}^*) = \alpha^2\Theta$

where $\pi_i^* = \pi_i - \hat{E}[\pi_i \mid X_i]$ and $F_{ij}^* = F_{ij} - \hat{E}[F_{ij} \mid X_i]$ are residualized outcomes and forecasts after conditioning on public information.

Instruments Z_{it}, constructed from firm-level or macro variables (e.g., size bins, disclosure regimes), enter the moment conditions via interactions with residuals. Identification relies on the exclusion of these instruments from the noise component, and on the orthogonality between residual soft information and idiosyncratic error.

The model estimates four key quantities: Θ (variance of soft information), α (extent of analyst reliance on soft information), Σ (variance of noise), and Δ (public bias). The soft bias component is further computed as $\Delta_p = (1-\alpha)^2\Theta$. Estimation is performed by minimizing the weighted squared moments, with standard errors obtained via heteroskedasticity-robust sandwich formula.

For interpretability, results are reported separately by forecast horizon and sample split (e.g.,

cross-country, pre/post shock). Appendix C provides technical derivations and simulation validation.

2.3. Identification Strategy

The scope and design of the international panel provide a credible basis for identification along three key dimensions: institutional differences across countries, temporal variation induced by macroeconomic shocks, and heterogeneity across forecast horizons. This empirical framework allows for a quasi-experimental approach to isolating the drivers of forecasting behavior.

- **Cross-sectional variation**, by comparing countries exposed to similar macroeconomic shocks but differing in institutional and informational environments, such as regulatory transparency, capital controls, and disclosure regimes. These differences affect how information is processed and incorporated into forecasts.
- **Temporal variation**, by tracking forecast components before, during, and after macroeconomic shocks, as well as across different stages of the same shock, capturing dynamic adjustments over time.
- **Variation across forecast horizons**, by analyzing the term structure of forecast accuracy. Differences in short- versus long-term predictive performance offer insights into the relative strengths and weaknesses of human and machine forecasts under varying information complexity.

3. Forecasting Results: Human vs. Machine

Drawing on a global dataset covering 47 countries and more than 200 macroeconomic shocks, I analyze forecasting behavior across three empirical dimensions: cross-sectional variation, temporal evolution around shocks, and variation by forecast horizon. This setting allows the decomposition of forecast errors into components attributable to soft information, behavioral bias, and random noise under diverse economic and institutional conditions.

3.1. Universal Forecasting Patterns

3.1.1. Descriptive Statistics of Forecasting Variables

Table A8 presents summary statistics for a selected set of forecast horizons, including three quarterly horizons ($h = 0.25, 0.5, 0.75$ years) and three annual horizons ($h = 1, 2, 4$ years). We report the distribution of analyst consensus forecasts (F_{it}^h), realized earnings (π_{it+h}), and forecast errors ($F_{it}^h - \pi_{it+h})/P_{it}$, normalized by price. In addition, the table includes the number of analysts issuing forecasts (N_{it}^h) and firm size as measured by total assets.

Analyst forecasts are generally optimistic, with mean forecast errors tending to be positive across horizons. Dispersion increases with forecast horizon, consistent with rising uncertainty and decreasing information precision. The number of analysts per firm-time observation declines with horizon, while firm size remains relatively stable. These patterns highlight the increasing challenge of long-term forecasting and motivate our focus on decomposing forecast errors into structural components across time and horizon.

Table **??** reports the summary statistics of firm-level analyst forecasts and associated forecast errors for selected horizons ($h = 0.25, 0.5, 1, 2, 4$). Forecast errors are computed as the difference between analyst forecasts and realized outcomes. The results show that mean

forecast errors are generally close to zero, but the dispersion increases with the forecast horizon, reflecting growing uncertainty.

Table A9 reports the distribution of forecast errors at the analyst level, highlighting the dispersion across individual analyst expectations and the magnitude of disagreement relative to realized firm earnings. The statistics are computed separately by forecast horizon.

This motivates the decomposition approach, which seeks to quantify the human-specific components, such as soft information, bias, and noise, embedded in analyst forecasts beyond what can be predicted by statistical models.

3.1.2. Consistent Forecast Patterns Across Models

To further evaluate the robustness of machine learning models in learning human forecasting behavior, we compare multiple algorithms' performance in predicting analysts' forecasts. Specifically, we define the *A-type forecast* as the machine's prediction of the analyst forecast, capturing the extent to which the machine can replicate human beliefs using public signals.

Figure A10 in the Appendix presents the mean squared error (MSE) of *A-type forecasts* across different horizons for two representative environments: structured (e.g., Germany) and noisy (e.g., Turkey). Across both countries, we observe a consistent pattern: MSE increases monotonically with the forecast horizon. This reflects the growing uncertainty in human forecasts as the horizon extends, and the greater difficulty for machines to replicate long-term analyst expectations accurately.

Importantly, both the elastic net and random forest models exhibit a similar directional trend in MSE growth, despite differences in absolute accuracy. This convergence in trend across algorithms highlights the shared underlying structure in how machines recognize patterns in analyst behavior. In particular, the similar slope and curvature of MSE lines across models indicate that pattern recognition, though varying in precision, follows a coherent logic across learning architectures.

Taken together, these results suggest that machines can internalize not only directional biases (as shown in Section 3.3.3) but also the structure of forecast uncertainty. The rising MSE profile implies that human forecasts become harder to predict as they become more speculative, and machines mirror this difficulty even when trained on rich public signals. This robustness across methods strengthens the evidence that human uncertainty is a learnable and measurable feature of forecast environments.

3.1.3. For Machines: Human Forecast Errors Are Most Predictable

Having shown earlier that machines struggle to predict analyst forecasts (A-type) and firm fundamentals (E-type) over long horizons, I now investigate a third, and perhaps more revealing learning task: can machines predict the analyst *forecast error*, that is, the difference between their forecast and the realized outcome?

The *AE-type forecast* is defined as the machine prediction of analyst forecast errors. Figure A11 and Figure A12 in the Appendix plot the MSE trends for E-type and AE-type forecasts, respectively, across forecast horizons and countries.

For E-type forecasts (Figure A11), MSE increases with horizon, as expected. This reflects the growing difficulty of forecasting actual earnings over longer time frames.

In contrast, AE-type forecasts (Figure A12) exhibit a *declining* MSE profile with horizon. That is, machines become better at predicting analyst forecast errors as the time horizon extends. This counterintuitive pattern reveals a surprising regularity: analysts become increasingly conservative at long horizons, and this behavior becomes more predictable to the machine.

Taken together, these results highlight a subtle but powerful insight: while machines struggle to predict analysts (A-type) and firms (E-type) directly, they can effectively learn how humans fail. That is, machines are especially good at predicting the *residual*—the behavioral bias in analyst forecasts—which turns out to be more systematic and learnable than the targets themselves. This underscores the promise of decompositional approaches to understanding human-machine interaction in forecasting.

This distinction highlights a key mechanism underlying my empirical findings. Predicting analyst forecasts (*F_ANALYST*) requires replicating raw human behavior, which is often noisy, biased, and highly context-dependent. In contrast, predicting analyst forecast errors (*F_ANALYST - ACTUAL*) enables the model to focus on systematic deviations from actual outcomes. When these deviations exhibit consistent patterns, such as persistent optimism or variation across forecast horizons, they become more amenable to learning.

In essence, machines find it difficult to replicate human forecasts, but are more successful at learning how humans tend to deviate from fundamentals. This insight underscores the value

of decompositional approaches for uncovering the structure of human forecasting errors.

I next examine whether this learnable residual structure varies systematically across markets with different levels of public information disclosure (see Section 3.3.4).

3.2. Shock-Time Dynamics

I next examine how forecasting behavior evolves across the life cycle of macroeconomic shocks. The relevance of soft information tends to be elevated during the early stages of a shock and declines during periods of stabilization.

As illustrated in Figure A4, human forecasts tend to outperform machine-based predictions in the early stages of financial crises, such as the onset of the dot-com bubble. This initial advantage likely reflects analysts' ability to rapidly incorporate qualitative signals, including emerging policy responses, supply chain disruptions, or shifts in market sentiment, which are not captured by structured historical data. In contrast, machine learning models, particularly those based on supervised learning algorithms such as Lasso or Random Forest, rely on patterns extracted from past data and typically produce estimates that reflect historically dominant regimes. As a result, they often interpret early-stage shocks as transient fluctuations rather than structural breaks, and therefore adapt more slowly.

As the crisis unfolds, however, the relative advantage of human forecasts declines. Machine models gradually adjust as new data accumulates and the underlying patterns shift. Meanwhile, human forecasters may become increasingly uncertain as the crisis persists beyond initial expectations. This uncertainty can lead to overreactions, inconsistent revisions, and increased forecast bias and noise. The narrowing, and at times reversal, of the performance gap between human and machine forecasts is consistent with the structural decomposition

of forecast errors and highlights the shifting balance of informational advantages over the course of financial shocks.

Similar patterns are observed in other crises, including the 2008 financial crisis and the collapse of the dot-com bubble. The persistence of analyst advantage varies with the nature of the shock.

The structural decomposition method complements traditional approaches that attribute forecast gaps solely to bias. By isolating the role of information, the model provides a more nuanced understanding of analyst behavior under conditions of macroeconomic turbulence.

3.3. Structured vs. Noisy Forecast Environments

3.3.1. Forecast Dynamics Cross Markets

Building on the temporal patterns documented earlier, I now compare how the relative performance of analyst and machine forecasts evolves across markets with different informational environments. This comparison highlights the extent to which market structure shapes the timing and persistence of forecasting advantages.

Appendix Figure **??** plots forecast errors for analysts and machines across structured and noisy markets during episodes of macroeconomic stress. A consistent pattern emerges across countries. In the early stages of a macroeconomic shock, analysts typically outperform machines, particularly in noisy environments. This initial advantage likely reflects analysts' ability to incorporate soft or firm-specific information that is unavailable to machines relying solely on public signals.

As the crisis progresses, however, the advantage of analysts diminishes. Forecast errors tend to rise most notably in noisy markets where analysts appear to revise expectations too frequently or too aggressively in response to evolving conditions. In contrast, machine forecasts remain more stable and gradually improve in relative accuracy as the informational content of public data increases and human forecasts become more erratic.

The timing of this reversal differs by market type. In structured environments, machines tend to catch up and outperform analysts relatively quickly, as the value of public information increases with resolution of uncertainty. In noisy markets, where public data remains sparse and soft information continues to play a role, the convergence is more gradual and sometimes incomplete.

Taken together, these results suggest that market-level information quality shapes not only forecast accuracy, but also the duration and stability of human-machine performance gaps. In the later section (Section 3.5.2), I discuss the underlying mechanisms, focusing on the role of soft information and information frictions in different markets.

3.3.2. Analyst Optimism vs. Pessimism

I examine the cross-country differences in analyst forecast bias by distinguishing between structured and noisy forecast environments. Consistent with the concept of analyst activism documented in the literature, I find that in structured markets, analysts tend to be overly optimistic, systematically forecasting earnings higher than the realized corporate profits. This pattern aligns with the notion that analysts in these markets actively shape investor expectations through optimistic forecasts.

In contrast, within noisy forecast environments, I observe a distinct pattern of pessimism. Here, analyst forecasts tend to underestimate actual firm earnings, revealing a downward bias. This divergence suggests that in less structured markets, analysts face greater uncertainty or noisier information, which leads to conservative or cautious earnings projections.

Figure A7 and A8 illustrates the term structure of annual and quarterly analyst forecast errors in two representative countries. Germany, a structured market, exhibits a consistent downward bias in analyst forecasts relative to actual earnings. In contrast, Turkey, characterized by a noisier forecast environment, shows an upward bias where analyst forecasts systematically exceed realized earnings. This cross-country evidence supports the hypothesis that the informational environment shapes the direction of analyst forecast bias.

Overall, these findings provide new insights into how the forecasting environment influences analyst optimism and pessimism, extending the literature on analyst activism by emphasizing that such activism is context dependent and varies across countries.

3.3.3. Machines Learn Human Forecast Bias

To assess the extent to which machines can learn systematic bias in human forecasts, I compare forecast error patterns across countries with contrasting institutional environments. Specifically, I examine countries representing *structured environments* (e.g., Germany), characterized by transparent financial reporting, stable macroeconomic policy, and low inflation volatility, versus those representing *noisy forecast environments* (e.g., Turkey), where analysts face greater informational frictions and macroeconomic uncertainty.

Figure A9 in the Appendix plots the term structure of forecast errors, both actual and

machine-predicted, across different horizons for these representative environments. In structured settings like Germany, analyst forecast errors display a steadily increasing trend with horizon, suggesting growing optimism. The machine effectively learns this directional pattern but consistently produces more muted forecast errors, indicating a conservative adjustment that avoids extreme human beliefs.

In contrast, in noisy environments such as Turkey, analysts display persistent pessimism across all horizons. Once again, the machine detects and reproduces the directional bias, yet its predictions are systematically closer to zero. This reinforces the finding that machines absorb the *direction* of human bias while tempering its *magnitude*.

3.3.4. Forecast Errors Are More Learnable in Structured Environments

As previewed in Section 3.1.3, I now examine whether the learnability of analyst forecast errors varies systematically across markets with different levels of public information disclosure.

Appendix Figure A12 shows that the MSE of AE-type forecasts declines more smoothly and consistently in structured environments. This pattern suggests that analyst forecast errors in these settings are more systematic and therefore more learnable. Machines replicate these patterns well, indicating that analyst behavior is driven by stable, repeatable biases rather than ad hoc adjustments.

By contrast, in noisy environments, the MSE trend is less stable and more volatile across horizons. Analyst forecast errors appear to be shaped by subjective views, firm-specific

narratives, or transitory shocks—factors that are often unobservable to the machine. As a result, model performance deteriorates.

Appendix Figure A11 reinforces this contrast in the context of E-type forecasts. As expected, MSE increases with horizon in both market types. However, the increase is more erratic and pronounced in noisy environments, reflecting the difficulty of forecasting fundamentals when public data is sparse or unreliable. In structured settings, the MSE slope is flatter and more regular, consistent with more stable earnings dynamics.

Taken together, these results underscore that forecast accuracy depends not only on data availability, but also on the nature of analyst behavior. In structured environments, analysts appear to follow consistent decision rules that can be learned and replicated by machines. In contrast, where behavior is shaped by discretion, intuition, or local information, prediction becomes less feasible.

This interpretation aligns with the weaker AE-type performance observed in emerging markets. When analyst errors reflect non-structural or unobservable components, machine learning models have limited capacity to generalize. Cross-country variation in model performance thus reflects deeper differences in the predictability of analyst behavior.

Finally, I note that across environments, machine-generated AE forecasts remain consistently closer to zero. This conservative tendency likely reflects the model's inability to replicate the full extent of human biases, but also its strength in avoiding overreaction. In this sense, machine learning may serve as a stabilizing force in noisy forecasting contexts.

3.3.5. Institutional Environments and Forecasting Potential

Forecast accuracy and composition vary systematically across countries facing similar macroeconomic shocks. A key explanatory factor is the availability of public data, shaped by institutional features such as regulatory quality, disclosure standards, and capital account openness. These institutions determine how easily algorithms can learn from data and adapt to changing conditions. For example, in countries like Germany with high data transparency, machine learning models can more effectively extract signals and improve forecast performance. In contrast, in countries like Turkey, where reliable public information is more limited, algorithms face constraints in both forecasting potential and learning speed.

This variation provides a foundation for decomposing forecast errors into components such as soft information, bias, and noise. In transparent markets, analysts respond more directly to latent signals, as confirmed by higher values of α in structural estimation. In less transparent settings, forecast errors contain more unstructured elements. Figure illustrates this cross-country dispersion in soft information reliance.

3.3.6. When Algorithms Learn and When They Struggle

Structured forecast environments are characterized by stable institutions and abundant data, allowing analysts and algorithms to form expectations based on repeatable patterns. In these settings, forecast errors are more systematic and reproducible, enabling strong model performance.

By contrast, noisy environments lack consistent disclosure and are subject to idiosyncratic

shocks and discretionary forecasting behavior. Here, analysts depend more on private judgment, and forecast errors appear less structured. As a result, machine learning models struggle to identify stable predictive signals.

Importantly, this asymmetry extends to prediction targets. Machine learning models are more effective at predicting forecast errors—systematic deviations from outcomes—than they are at replicating full analyst forecasts, which embed behavioral biases and inaccessible information.

Recognizing these differences is critical. Structured environments offer fertile ground for model-based forecasting, while noisier environments limit the scope for algorithmic learning and reinforce the value of human expertise.

3.4. Forecast Horizon and Term Structure

3.4.1. General Trends in the Term Structure

My findings from the international datasets are consistent with the general trends documented in De Silva and Thesmar (2024). Forecast performance varies systematically across horizons: human predictions tend to be more accurate in the short term, while model-based forecasts become relatively more effective at longer horizons.

For horizons within one year, analyst forecasts generally achieve higher accuracy than machine forecasts. This pattern reflects the analysts' ability to incorporate contextual knowledge and qualitative insights that are difficult to encode into statistical features.

Beyond two years, algorithmic forecasts close the performance gap. Human forecasts at long horizons exhibit increasing variance and systematic deviation from outcomes, particularly in sectors characterized by technological uncertainty.

3.4.2. Country Differences

However, I also find meaningful cross-country differences. In countries where access to information is more restricted, such as Turkey, the performance of machine learning models appears to be more limited compared to countries with more transparent and comprehensive public data, such as Germany. I interpret this as evidence that limited data availability constrains the effectiveness of model-based forecasting in information-scarce environments.

3.4.3. Machine Insensitivity to Horizon

As shown in Figure A5, forecast accuracy exhibits distinct dynamics across time horizons for human analysts and machine learning models. The blue line depicts the term structure of forecast errors for human analysts, measured as the deviation between their forecasts and the realized earnings of firms. I observe a clear upward trend in human forecast errors as the time horizon increases, indicating that analysts tend to make larger mistakes the further into the future they attempt to forecast.

In contrast, the orange line reflects the forecast error produced by the random forest model. Unlike the human benchmark, the model's error does not exhibit a strong upward or downward trend over the forecast horizon. This pattern suggests that the machine learning model is relatively insensitive to the forecast horizon, likely because it consistently applies patterns learned from historical data without adjusting expectations based on horizon-specific intuitions or macro narratives.

Together, these results underscore the differing nature of forecast formation between human and machine. While human forecasters may benefit from near-term qualitative signals, their performance deteriorates at longer horizons, potentially due to overconfidence, narrative extrapolation, or underweighting uncertainty. The model, by contrast, maintains stable performance across horizons, though it may also fail to capture important forward-looking dynamics at short horizons that are not present in the training data.

This contrast in sensitivity may also reflect differences in how soft information, bias, and noise evolve across forecast horizons for each approach. Motivated by this observation, I proceed with a structural decomposition of the forecasts in the following section.

These results suggest a form of forecast specialization. Human forecasters offer greater value in the interpretation of short-term developments, especially during uncertain periods, while machine learning models are better equipped to extrapolate long-run trends from structured data.

3.5. Structural Decomposition

3.5.1. Model Overview

To estimate the composition of forecast errors, I implement a structural model following the approach proposed by De Silva and Thesmar (2024). The model expresses analyst j's forecast for firm i at time t as:

$$F_{ijt} = E_{it}^{ML} + z_{it} + b_{ijt} + \eta_{ijt} \tag{4}$$

where E_{it}^{ML} is the machine learning prediction of the realized outcome, z_{it} denotes unobserved soft information, b_{ijt} captures systematic bias, and η_{ijt} represents idiosyncratic forecast noise.

I estimate the model using the Generalized Method of Moments, relying on forecast and realization residuals. The parameters of interest include:

- α: the responsiveness of analyst forecasts to soft information,
- θ: the variance of the soft information component,
- Σ: the variance of forecast noise across analysts,
- Δ: the average deviation between analyst and machine forecasts,
- Δ_p: the bias component resulting from under- or over-reaction to soft signals, defined as $(1-\alpha)^2 \cdot \theta$.

These parameters are estimated separately by country group and forecast horizon. Appendix C provides further details on identification assumptions and variable definitions.

3.5.2. Decomposition Results

In this section, I present the decomposition of human adjustments into three key components: soft information, bias, and noise. These components are derived from my structural estimation, performed using the Generalized Method of Moments (GMM), a robust econometric technique. This decomposition sheds light on the differences between human and machine

forecasts and explains why human prediction accuracy tends to decrease with longer forecast horizons, a trend observed in earlier sections.

As illustrated in Figure A6 in the Appendix, the contribution of these components varies significantly across forecast horizons. For short-term horizons, human forecasts often outperform machine learning models. I attribute this superior performance primarily to humans' richer access to and effective integration of soft information, which encapsulates unique private information.

However, as the forecast horizon extends, the dynamics shift. Figure A6 clearly demonstrates that the contribution of soft information to human forecasts is decreasing with the increasing forecast horizon. Concurrently, behavioral bias and noise play an increasingly prominent role, showing an increasing contribution to human forecasts over longer horizons. This accumulation of bias and noise ultimately leads to the observed decline in the relative accuracy of human forecasts compared to machine learning models at extended horizons.

4. Robustness and Extensions

4.1. Extension: Large Language Models for Soft Information Extraction

As an extension to the benchmark machine learning framework, I incorporate large language models (LLMs) to evaluate their ability to capture soft information from firm-level textual data. Rather than using LLMs as independent forecasters, which may introduce lookahead bias due to training on the full textual corpus, I employ them to extract sentiment-based signals such as tone, narrative focus, and linguistic uncertainty. These signals are derived

from forward-looking documents including earnings call transcripts, management guidance, and macroeconomic news summaries.

The extracted indicators are then used as additional inputs within the machine learning models, resulting in what I refer to as LLM-augmented forecasts. This design enables a structured comparison across three types of forecasters: human analysts, traditional machine learning models based solely on structured data, and machine learners supplemented with soft information extracted from unstructured textual sources.

This approach contributes to a more nuanced understanding of how qualitative information influences financial forecasting. Although LLMs cannot replicate the full depth of human intuition or access to private channels, they provide a scalable and transparent method for incorporating public soft information into forecasting models. Their integration clarifies the respective strengths of human forecasters, conventional algorithms, and language-based tools in settings where soft signals are expected to be valuable.

4.2. Disclosure Quality and Machine Forecast Accuracy

To explore how disclosure environments shape the effectiveness of machine forecasts, I regress forecast accuracy on cross-country differences in disclosure quality. The central hypothesis is that higher-quality public information environments improve the relative performance of machine learning models, which rely primarily on structured and widely available data.

The main explanatory variable is the Accounting Standards Index introduced by Porta et al. (1998), which has been widely used in the literature on law, finance, and information environments. This index captures the comprehensiveness and quality of financial accounting standards at the country level.

An alternative measure of accounting disclosure quality is proposed by Lang, Lins, and Maffett (2012), which is widely regarded in the accounting literature as a reliable proxy for the quality of financial transparency across countries.

The regression includes several layers of controls. At the country level, I control for gross national product per capita and other macroeconomic indicators that may confound the relationship between disclosure quality and forecasting performance. At the firm level, I include controls for financial fundamentals from Compustat and market-based variables from Datastream. To account for potential variation in data availability, I also control for the size of the IBES forecast panel and the rate of missing forecasts.

All regressions include fixed effects for country, legal origin, industry, and calendar season to absorb systematic variation unrelated to disclosure practices. As part of ongoing robustness checks, I also explore alternative measures of data infrastructure and transparency. In particular, I consider the ODIN Pillar Scores from the World Bank's Open Data Inventory, which provide country-level indicators on data use, services, products, sources, and infrastructure.

This empirical design allows for a rigorous assessment of whether and where machine-based forecasts benefit from stronger public information environments.

5. Conclusion

This paper provides an in-depth analysis of how human expectations respond to macroeconomic shocks by decomposing belief formation into context information, bias, and noise. Through a comprehensive meta-statistical analysis of international datasets from 47 countries and over 200 macroeconomic events, I have illustrated the critical role of these components in shaping human expectations during periods of significant structural change, such as financial crises.

Our findings highlight the importance of context information, particularly during the early stages of a macro shock, when human forecasters hold a significant advantage due to their access to soft information and real-time contextual cues. This advantage gradually declines as the shock progresses, and the influence of bias and noise becomes more pronounced. This pattern is especially evident in emerging markets, where the learning process is slower and the fluctuations in expectations are more extreme compared to developed economies.

By examining forecasts across different horizons, I find that short-term predictions benefit substantially from private information, granting humans an advantage over purely statistical methods. In contrast, long-term forecasts are more vulnerable to behavioral distortions and noise, which makes machine forecasts more stable and reliable over extended periods.

The availability of public data is also essential in shaping forecast accuracy, particularly in emerging markets where limited transparency amplifies the importance of both soft information and the ability to filter out noise. These findings underscore the limitations of static machine learning forecasts and emphasize the value of combining algorithmic models with human judgment, especially under uncertainty or structural change.

The application of machine learning algorithms on a large and diverse international dataset allows this study to move beyond conventional approaches that rely on simulations or data from a few highly developed markets. The cross-country analysis reveals meaningful variations in the roles of context information, bias, and noise under different institutional and informational conditions.

In conclusion, this research affirms that human forecasts remain valuable. Soft information plays a critical role, especially at short horizons and in the early phases of crises. Forecast errors vary systematically across countries, forecast horizons, and stages of macroeconomic shocks, shaped by differences in information environments and behavioral responses. These results offer new insights into belief formation and support the development of forecasting systems that combine the complementary strengths of humans and machines.

Limitations and Future Directions.. While our structural model provides a useful decomposition of forecast errors, it assumes that the informational variance parameters (such as θ) remain constant over time. In practice, the quality and availability of soft information may evolve as a macro shock unfolds or as market institutions adapt. Future work could extend the model by allowing θ to vary over time or across institutional contexts, capturing more dynamic patterns of learning and adjustment in belief formation.

References

Acemoglu, Daron. 2021. *Redesigning Ai.*: MIT Press.

Acemoglu, Daron, and Pascual Restrepo. 2018. "Artificial intelligence, automation, and work." In *The economics of artificial intelligence: An agenda*, 197–236: University of Chicago Press.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction machines: the simple economics of artificial intelligence.*: Harvard Business Press.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2022. *Prediction machines, updated and expanded: The simple economics of artificial intelligence.*: Harvard Business Press.

Autor, David H. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30.

Bae, Kee-Hong, René M Stulz, and Hongping Tan. 2008. "Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts." *Journal of financial economics* 88 (3): 581–606.

Ball, Ryan T, and Eric Ghysels. 2018. "Automated earnings forecasts: Beat analysts or combine and conquer?" *Management Science* 64 (10): 4936–4952.

Barrot, Jean-Noël, and Julien Sauvagnat. 2016. "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks." *The Quarterly Journal of Economics* 131 (3): 1543–1592.

Bonelli, Maxime. 2022. "The Adoption of Artificial Intelligence by Venture Capitalists."

Born, Benjamin, and Johannes Pfeifer. 2009. "Policy Risk and the Business Cycle." *Journal of Monetary Economics* 56 (4): 623–633.

Campbell, John Y., and Robert J. Shiller. 1997. "Valuation Ratios and the Long-Run Stock Market Outlook." *The Journal of Portfolio Management* 24 (2): 11–26.

Cao, Jian, and Mark Kohlbeck. 2011. "Analyst quality, optimistic bias, and reactions to major news." *Journal of Accounting, Auditing & Finance* 26 (3): 502–526.

Cao, Sean, Wei Jiang, Junbo L Wang, and Baozhong Yang. 2021. "From man vs. machine to man+ machine: The art and ai of stock analyses."Technical report, National Bureau of Economic Research.

Chen, Nai-Fu, Richard Roll, and Stephen A Ross. 1986. "Economic forces and the stock market." *Journal of business*: 383–403.

Cowen, Amanda, Boris Groysberg, and Paul Healy. 2006. "Which types of analyst firms are more optimistic?" *Journal of Accounting and Economics* 41 (1-2): 119–146.

Dahlquist, Magnus, Lee Pinkowitz, René M Stulz, and Rohan Williamson. 2003. "Corporate governance and the home bias." *Journal of financial and quantitative analysis* 38 (1): 87–110.

De Silva, Tim, and David Thesmar. 2021. "Noise in expectations: Evidence from analyst forecasts."Technical report, National Bureau of Economic Research.

De Silva, Tim, and David Thesmar. 2024. "Noise in expectations: Evidence from analyst forecasts." *The Review of Financial Studies* 37 (5): 1494–1537.

Dessaint, Olivier, Thierry Foucault, and Laurent Frésard. 2021. "Does alternative data improve financial forecasting? the horizon effect."

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2007. "The Law and Economics of Self-Dealing." *Journal of Financial Economics* 88 (3): 430–465.

Fama, Eugene F, and Kenneth R French. 1989. "Business conditions and expected returns on stocks and bonds." *Journal of financial economics* 25 (1): 23–49.

Forbes, Kristin J., and Francis E. Warnock. 2014. "Capital Flow Waves: Surges, Stops, Flight, and Retrenchment." *Journal of International Economics* 92 (2): 235–251.

Fouliard, Jeremy, Michael Howell, and Hélène Rey. 2021. "Answering the queen: Machine learning and financial crises."Technical report, National Bureau of Economic Research.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery. 2019. "The Role of Technology in Mortgage Lending." *The Review of Financial Studies* 32 (5): 1854–1899.

Gabaix, Xavier, and Ralph S. J. Koijen. 2020. "In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis." *NBER Working Paper* (27367).

Groysberg, Boris, Lex Lee, and Ashish Nanda. 2011. "Does Individual Performance Affect Entrepreneurial Mobility? Evidence from the Financial Analysis Market." *Journal of Financial Economics* 100 (3): 581–595.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. "Empirical Asset Pricing via Machine Learning." *The Review of Financial Studies* 33 (5): 2223–2273.

Hansen, Jorge W, and Christoffer Thimsen. 2021. "Forecasting corporate earnings with machine learning." *It takes considerable knowledge just to realize the extent of your own ignorance.*: 57.

Hong, Harrison, and Jeffrey D. Kubik. 2003a. "Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts." *The Journal of Finance* 58 (1): 313–351.

Hong, Harrison, and Jeffrey D Kubik. 2003b. "Analyzing the analysts: Career concerns and biased earnings forecasts." *The Journal of Finance* 58 (1): 313–351.

Kothari, SP. 2001. "Capital markets research in accounting." *Journal of accounting and economics* 31 (1-3): 105–231.

La Porta, Rafael, Florencio Lopez-de Silanes, and Andrei Shleifer. 2006. "What Works in Securities Laws?" *The Journal of Finance* 61 (1): 1–32.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert W. Vishny. 1998. "Law and Finance." *Journal of Political Economy* 106 (6): 1113–1155.

Landier, Augustin, Yueran Ma, and David Thesmar. 2017. "New experimental evidence on expectations formation."

Lang, Mark, Karl V Lins, and Mark Maffett. 2012. "Transparency, liquidity, and valuation: International evidence on when transparency matters most." *Journal of Accounting Research* 50 (3): 729–774.

Lim, Terence. 2001. "Rationality and analysts' forecast bias." *The journal of Finance* 56 (1): 369–385.

Lucas, Robert E. 1976. "Econometric policy evaluation: A critique." *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

Mola, Simona, and Massimo Guidolin. 2009. "Affiliated mutual funds and analyst optimism." *Journal of Financial Economics* 93 (1): 108–137.

Obizhaeva, Anna A, and Yajun Wang. 2019. "Trading in crowded markets." *Available at SSRN 31527*.

Pinkowitz, Lee, René M Stulz, and Rohan Williamson. 2001. "Corporate governance and the home bias."

Porta, Rafael La, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert W Vishny. 1998. "Law and finance." *Journal of political economy* 106 (6): 1113–1155.

Price, S. McKay, James S. Doran, David R. Peterson, and Brett A. Bliss. 2012. "Earnings Conference Calls and Stock Returns: The Incremental Informativeness of Textual Tone." *Journal of Banking & Finance* 36 (4): 992–1011.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*.: Princeton University Press.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics* 50 (3): 665–690.

Susan Athey. 2023. "The Brookings Institution Webinar: ChatGPT and the Future of Work." https://www.brookings.edu/wp-content/uploads/2023/03/es_20230315_chatgpt_work_transcript.pdf. Accessed on 15 March 2023.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance* 62 (3): 1139–1168.

Van Binsbergen, Jules H, Xiao Han, and Alejandro Lopez-Lira. 2020. “Man vs. machine learning: The term structure of earnings expectations and conditional biases.”Technical report, National Bureau of Economic Research.

Woodford, Michael. 2003. “nImperfect Common Knowledge and the Effects of Monetary Policy. oIn Philippe Aghion, Roman Frydman, Joseph E. Stiglitz and Michael Wood *ford*, *eds*., *Knowledge*, *Information*, *and Expectations in Modern Macroeconomics* : *In Honor of Edmund S. Phelps*."

Appendix A. Machine Learning Techniques

This section provides a more detailed description of the supervised learning techniques I explore for forecasting firm earnings. I begin with the class of penalized linear estimators, followed by tree-based methods.

A.1. Quasi-Linear Models

Quasi-linear models are a class of supervised learning algorithms that combine linear regression with regularization techniques. These models aim to balance the trade-off between model complexity and prediction accuracy. The following three quasi-linear models are commonly used in financial analysis:

A.1.1. Lasso

Lasso (Least Absolute Shrinkage and Selection Operator) is a penalized linear estimator that adds an L1 penalty term to the mean squared error (MSE) loss function. The objective function for Lasso is defined as:

$$L(\beta, \alpha_1, \alpha_2) = \sum \left[(EPS - X'\beta)^2 \right] + \alpha_1 \|\beta\|_1 + \alpha_2 \|\beta\|_2^2,$$

where β represents the coefficient vector, α_1 and α_2 are the penalty parameters controlling the amount of regularization.

A.1.2. Ridge

Ridge is another penalized linear estimator that introduces an L2 penalty term to the MSE loss function. The objective function for Ridge is given by:

$$L(\beta, \alpha_1, \alpha_2) = \sum \left[(EPS - X'\beta)^2 \right] + \alpha_1 \|\beta\|_1 + \alpha_2 \|\beta\|_2^2,$$

where the terms have the same meaning as in Lasso.

A.1.3. Elastic Net

Elastic Net combines the L1 and L2 penalties of Lasso and Ridge, respectively, in order to leverage the benefits of both regularization techniques. The objective function for Elastic Net is defined as:

$$L(\beta, \alpha_1, \alpha_2) = \sum \left[(EPS - X'\beta)^2 \right] + \alpha_1 \|\beta\|_1 + \alpha_2 \|\beta\|_2^2,$$

where α_1 and α_2 control the amount of regularization.
The hyperparameters α_1 and α_2 are chosen using cross-validation on the training set to avoid introducing any look-ahead bias.

A.2. Non-Linear Models

Non-linear models offer greater flexibility in capturing complex relationships between predictor variables and firm earnings. I consider two popular non-linear models:

A.2.1. Random Forest

Random Forest is an ensemble learning method that combines multiple regression trees. Each tree is built using a random subset of predictor variables and a random subset of observations. The final prediction is obtained by averaging the predictions of all the trees. Random Forest is regularized through the averaging of trees with different structures, reducing prediction variance and limiting overfitting. The hyperparameters, such as the number of trees and the maximum number of splits, can be chosen using cross-validation on the training set.

A.2.2. Gradient-Boosted Trees

Gradient-Boosted Trees (GBT) is another tree-based method that builds an ensemble of regression trees in a sequential manner. GBT starts by fitting a shallow tree of depth d to the data and calculates the residuals from this tree. Then, another shallow tree of depth d is fitted to the residuals, and this process is repeated for a specified number of iterations. The predicted values are formed by adding the predicted values from each tree, with a regularization factor λ applied to shrink the predicted values from subsequent trees. By growing trees sequentially on the residuals from previous trees, GBT reduces correlation among the trees and limits overfitting.

GBT has three hyperparameters: the number of iterations B, the depth of each tree d, and the regularization factor λ. These hyperparameters can be chosen using cross-validation on the training set.

Appendix B. Forecasts formation

B.1. Analyst Forecast Processing

We implement several procedures to ensure the robustness of analyst forecast data:

- **Outlier Control**: Earnings-per-share (EPS) forecasts are winsorized at the 5% level within each country on an annual basis to remove extreme values that may distort analysis.
- **Date Alignment**: Forecast dates and earnings release dates are standardized using UTC timestamps to ensure temporal consistency across countries and time zones.
- **Forecast Horizon Classification**: Forecasts are grouped into short-, mid-, and long-term horizons depending on the number of calendar days between the forecast date and the actual earnings announcement. The classification procedure follows Dessaint, Foucault, and Frésard (2021), and details are provided in Section 2.2.2.

B.2. Machine Forecast Formation

Following the common framework of current literature Van Binsbergen, Han, and Lopez-Lira (2020), De Silva and Thesmar (2021), Cao et al. (2021), I build the machine analyst using the following framework:

- **Build a Machine Analyst using algorithms**: I construct the machine analyst by leveraging supervised learning algorithms. Specifically, I utilize quasi-linear models such as Lasso, Ridge, Elastic Net, as well as non-linear models such as Random Forest and Gradient-Boosted Trees. These algorithms are chosen for their ability to capture complex patterns and relationships in the data.
- **Feed public information to the machine and generate forecasts**: The machine analyst is trained using a combination of publicly available information. This includes

financial statements, market data, macro information and other relevant variables. The trained machine analyst then generates forecasts based on this input.

- **Compare Man vs. Machine in forecasting accuracy**: I assess the forecasting accuracy of the machine analyst by comparing its predictions against the historical forecasts made by human analysts. This comparison allows us to evaluate the performance of the machine analyst in terms of accuracy and reliability.
- **Feed (analyst forecasts + public information) to the machine to build a (Man+Machine) analyst**: To further enhance the forecasting process, I combine the forecasts made by the human analysts with the information provided to the machine analyst. This fusion of inputs creates a hybrid forecasting approach, referred to as the (Man+Machine) analyst.
- **Compare Man vs. Machine vs (Man+Machine)**: Finally, I compare the forecasting performance of the human analyst, machine analyst, and the hybrid (Man+Machine) analyst. This comparison enables us to evaluate the relative strengths and weaknesses of each approach and identify the most effective forecasting strategy.

Appendix C. GMM Estimation Details

C.1. Moment Condition Derivation

The model structure is:

$$F_{ij} = x_i + z_i + b_{ij} + \eta_{ij}$$

Define residualized forecasts and outcomes:

$$F_{ij}^* = F_{ij} - \hat{E}[F_{ij} \mid X_i], \quad \pi_i^* = \pi_i - \hat{E}[\pi_i \mid X_i]$$

Based on the orthogonality of noise and residual soft information, the following moment conditions hold:

(A1) $$\mathbb{E}[\pi_i^* F_{ij}^*] = \alpha\Theta$$

(A2) $$\mathbb{E}[(F_{ij}^*)^2] = \alpha^2\Theta + \Sigma$$

(A3) $$\mathbb{E}[F_{ij}^* F_{ik}^*] = \alpha^2\Theta \quad (j \neq k)$$

C.2. Instrument Construction

Instruments Z_{it} are constructed from firm- or macro-level characteristics, including size bin dummies, lagged volatility measures, and country-level disclosure regimes. Moment conditions are multiplied by Z_{it} to yield over-identified GMM equations.

C.3. Parameter Interpretation

- Θ: Variance of soft information (incremental, unobserved signal used by analysts)
- α: Intensity of soft information usage
- Σ: Variance of idiosyncratic noise
- Δ: Bias from public model mis-specification

- $\Delta_p = (1-\alpha)^2\Theta$: Misuse of soft information (soft bias)

C.4. Estimation Procedure

Estimation is implemented via two-step GMM with robust standard errors. Each horizon is estimated separately, and moment weighting is optimized using the inverse of the moment covariance matrix. Identification is confirmed via Hansen's J-statistic.

C.5. Validation and Robustness

I validate the estimation framework through Monte Carlo simulations (not shown) and compare parameter estimates across alternative instrument sets. Results are robust to trimming of forecast tails and subsample exclusions.

Appendix D. International Datasets

D.1. List of Country/Region

The study considers the countries included in the MSCI ACWI Index, which represents stocks from 23 Developed Markets and 24 Emerging Markets, covering about 85% of the global investable equity market. See Table A1 for the full list of countries and regions.

TABLE A1. List of Countries and Regions in the MSCI ACWI Index

MSCI World Index (Developed Markets)	
Americas	Canada, United States
Europe & Middle East	Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom
Pacific	Australia, Hong Kong, Japan, New Zealand, Singapore
MSCI Emerging Markets Index	
Americas	Brazil, Chile, Colombia, Mexico, Peru
Europe, Middle East & Africa	Czech Republic, Egypt, Greece, Hungary, Kuwait, Poland, Qatar, Saudi Arabia, South Africa, Turkey, United Arab Emirates
Asia	Mainland China, India, Indonesia, Korea, Malaysia, Philippines, Taiwan, Thailand

D.2. List of Macro Shocks by Country/Region

The study examines 47 countries and regions and incorporates more than 200 macroeconomic shocks observed since the widespread adoption of machine learning technologies in the 1990s.

TABLE A2. List of Macro Shocks by Country/Region

Country/ Region	**Early 1990 Recession**	**2000 Dot-com Bubble**	**2008 Global Financial Crisis**	**2020 Covid Pandemic**	**Other Macro Shock(s)**
Australia	✓	✓	✓	✓	
Austria	✓	✓	✓	✓	
Belgium	✓	✓	✓	✓	
Brazil	✓	✓	✓	✓	1994 Mexican Peso Crisis
Canada	✓	✓	✓	✓	
Chile	✓	✓	✓	✓	1994 Mexican Peso Crisis
China	✓	✓	✓	✓	1997 Asian Financial Crisis 2015 Chinese Stock Market Crash
Colombia	✓	✓	✓	✓	
Czech Republic	✓	✓	✓	✓	1998 Russian Financial Crisis
Denmark	✓	✓	✓	✓	
Egypt	✓	✓	✓	✓	
Finland	✓	✓	✓	✓	
France	✓	✓	✓	✓	
Germany	✓	✓	✓	✓	
Greece	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Japan	✓	✓	✓	✓	1991 Japanese Asset Bubble Burst; 1997 Asian Financial Crisis
Korea	✓	✓	✓	✓	1997 Asian Financial Crisis

Continued on next page

Country/ Region	Early 1990 Recession	2000 Dot-com Bubble	2008 Global Financial Crisis	2020 Covid Pandemic	Other Macro Shock(s)
Kuwait	✓	✓	✓	✓	
Hong Kong	✓	✓	✓	✓	1997 Asian Financial Crisis
Hungary	✓	✓	✓	✓	1998 Russian Financial Crisis
India	✓	✓	✓	✓	
Indonesia	✓	✓	✓	✓	1997 Asian Financial Crisis
Ireland	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Israel	✓	✓	✓	✓	
Italy	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Malaysia	✓	✓	✓	✓	1997 Asian Financial Crisis
Mexico	✓	✓	✓	✓	1994 Mexican Peso Crisis
Netherlands	✓	✓	✓	✓	
Norway	✓	✓	✓	✓	
Philippines	✓	✓	✓	✓	1997 Asian Financial Crisis
Poland	✓	✓	✓	✓	1998 Russian Financial Crisis
Portugal	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Peru	✓	✓	✓	✓	
Qatar	✓	✓	✓	✓	
Saudi Arabia	✓	✓	✓	✓	
Singapore	✓	✓	✓	✓	
South Africa	✓	✓	✓	✓	
Spain	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Sweden	✓	✓	✓	✓	
Switzerland	✓	✓	✓	✓	

Continued on next page

Country/ Region	Early 1990 Recession	2000 Dot-com Bubble	2008 Global Financial Crisis	2020 Covid Pandemic	Other Macro Shock(s)
United Kingdom	✓	✓	✓	✓	2010 European Sovereign Debt Crisis
Taiwan	✓	✓	✓	✓	
Thailand	✓	✓	✓	✓	1997 Asian Financial Crisis
Turkey	✓	✓	✓	✓	
United Arab Emirates	✓	✓	✓	✓	
United States	✓	✓	✓	✓	
New Zealand	✓	✓	✓	✓	

Explanations

- **1991 Japanese Asset Bubble Burst**: Japan experienced a significant economic bubble in the late 1980s, primarily driven by rapid increases in real estate and stock market prices. The bubble burst in 1991, leading to a prolonged period of economic stagnation known as the ”Lost Decade.”
    - Affected Country: Japan
- **1994 Mexican Peso Crisis (Tequila Crisis)**: A sudden devaluation of the Mexican peso in December 1994 triggered a financial Crisis. This led to severe economic and social disruptions in Mexico and impacted other emerging markets.
    - Affected Country: Mexico
    - Countries affected by ripple effects: Argentina and other Latin American countries
- **1997 Asian Financial Crisis**: Starting in Thailand with the collapse of the Thai baht, this Crisis spread to several Asian countries including Indonesia, South Korea, and

Malaysia. It resulted in severe economic downturns and required international financial intervention.

  - Affected Countries: Thailand, Indonesia, South Korea, Malaysia, Philippines, Hong Kong, Laos
  - Countries affected by ripple effects: Japan, China, and other emerging markets

- **1998 Russian Financial Crisis**: Russia devalued the ruble and defaulted on its debt in August 1998 due to a collapse in commodity prices and political instability. This Crisis had ripple effects on global financial markets.

  - Affected Country: Russia
  - Countries affected by ripple effects: Emerging markets in Eastern Europe and global financial markets

- **2000 Dot-com Bubble**: The rapid rise and subsequent collapse of internet-based companies' stock prices around the turn of the millennium. The NASDAQ Composite index, which includes many tech stocks, lost nearly 78

  - Affected Countries: Primarily the United States, but also global markets with tech stocks

- **2007-2008 Global Financial Crisis**: Triggered by the collapse of the subprime mortgage market in the United States, this macro shock led to the failure of major financial institutions, bailouts of banks by national governments, and significant downturns in stock markets worldwide.

  - Affected Countries: United States, United Kingdom, Iceland, Ireland, Spain, Greece, Portugal, Italy
  - Countries affected by ripple effects: Global impact, affecting most economies worldwide

- **2010 European Sovereign Debt Crisis**: Several Eurozone countries, including Greece, Ireland, Portugal, and Spain, faced high government debt levels and rising borrowing

costs. This led to austerity measures and financial assistance from the European Union and International Monetary Fund.

  - Affected Countries: Greece, Ireland, Portugal, Spain, Italy, Cyprus
  - Countries affected by ripple effects: Entire Eurozone and global markets

- **2015 Chinese Stock Market Crash**: China's stock markets saw dramatic rises and falls in 2015, leading to a global sell-off. The Shanghai Stock Exchange fell by 32

  - Affected Country: China
  - Countries affected by ripple effects: Global markets, especially emerging markets closely tied to China's economy

- **2020 COVID-19 Pandemic**: The global outbreak of COVID-19 led to unprecedented economic shutdowns and contractions. Stock markets plummeted in March 2020, and many countries implemented stimulus measures to support their economies.

  - Affected Countries: Global impact, with virtually all countries affected to varying degrees. The major economies hit include the United States, China, member states of the European Union, India, Brazil, and many others.

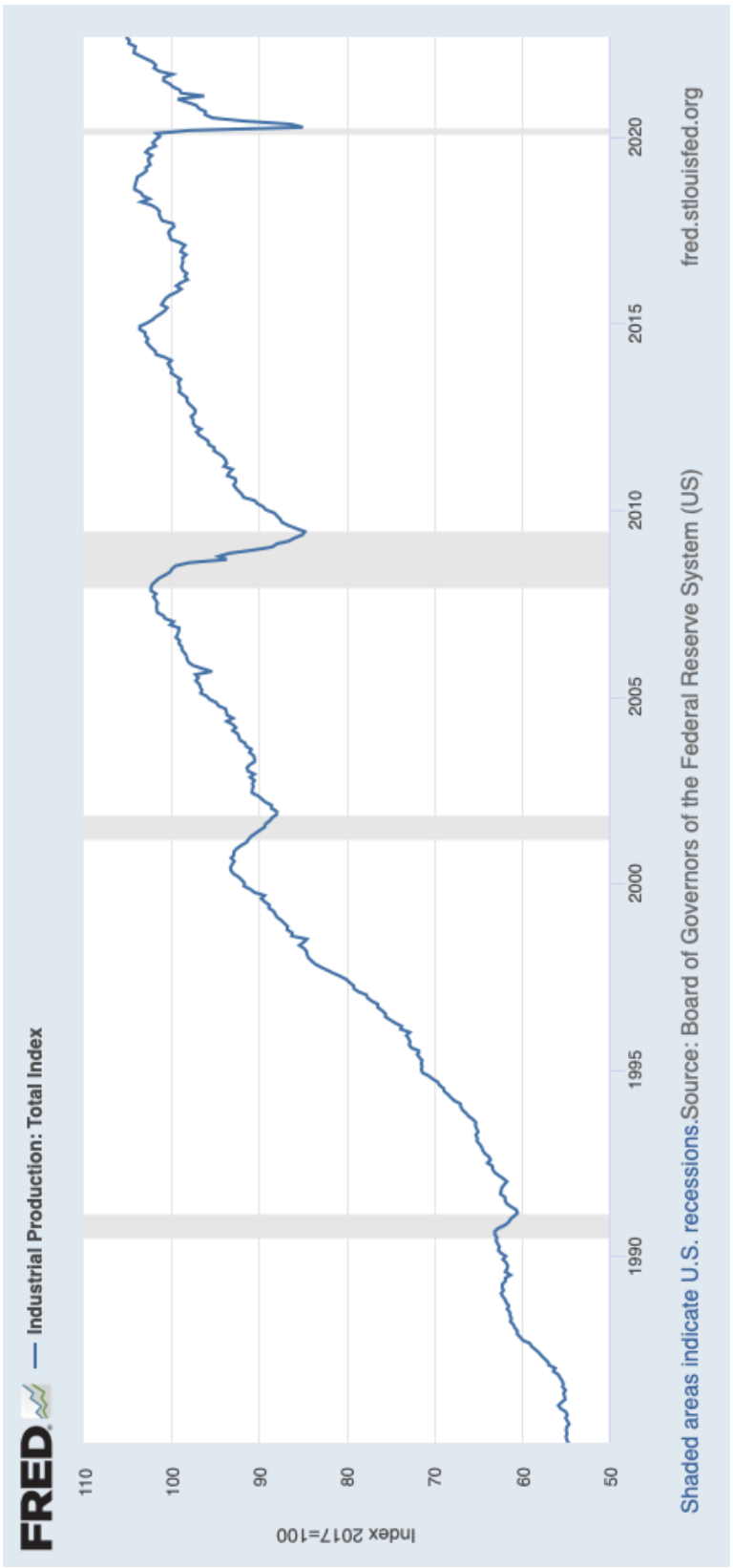


FIGURE A1. Macro shocks considered in United States

Appendix E. Data Inputs

E.1. Fundamental Variables

E.1.1. Compustat

Owing to the variations in the available variables lists between Compustat North America (covering the United States and Canada) and Compustat Global (covering all other countries/regions), the fundamental variables gathered for the United States and Canada differ slightly from those in other countries within the international datasets. Efforts were made to collect comparable variables to minimize the differences in fundamental variables as much as feasible.

It's important to mention that for some variables, the naming conventions in the Compustat Global dataset differ from those in the Compustat North America dataset.

The comprehensive lists of Computstat variables for North America and other countries/regions in the international datasets are detailed in Table A3 and Table A4, respectively.

TABLE A3. Fundamental Variables from Compustat North America

Variable	**Required non-missing?**
Total assets	✓
Total liabilities	✓
Revenue	✓
SG&A expense	
R&D expense	
Cost of goods sold	✓
Current assets	
Current liabilities	
Cash	
Cash and short-term investments	
Income tax expense	
Total long-term debt	
Total long-term debt due within one-year	
Debt in current liabilities	
Depreciation expense	
EBIT	
EBITDA	✓
Interest expense	
Interest paid	
Capital expenditures	
Goodwill	
Income tax payable	
Income tax expense	
Total income tax	
Net income	
Common dividends	

Continued on next page

Variable	Required non-missing?
Purchase of common and preferred stock	
Sale of common and preferred stock	
Subordinated debt	
Gross profit	✓
Operating cash flow	✓
Common shares outstanding	
Stock price at fiscal year end	✓
Extraordinary items	
Common ESOP obligation	
Special items	
Acquisitions	
Capitalized leases (due within two-years)	
Capitalized leases (due within three-years)	
Capitalized leases (due within four years)	
Capitalized leases (due within five years)	
Interest and related income (total)	
Total intangible assets	
Marketable securities adjustment	
Net PPE	✓
Nonoperating income	
Tax loss carryforward	
Pension and retirement expense	
Preferred stock value	

TABLE A4. Fundamental Variables from Compustat Global

Compustat Name	**Variable**	**Required non-missing?**
act	Total Current Assets	
aqc	Acquisitions	
at	Total assets	✓
capfl	Capital Element of Finance Lease Rental Payments	
capx	Capital Expenditures	
ceq	Total Common/Ordinary Equity	
ch	Cash	
che	Cash and Short-Term Investments	
cogs	Cost of Goods Sold	✓
dd1	ong-Term Debt Due in One Year	
dlc	Total Debt in Current Liabilities	
dltt	Total Long-Term Debt	
dp	Depreciation and Amortization	✓
dvc	Dividends Common/Ordinary	
ebit	Earnings Before Interest and Taxes	✓
gdwl	Goodwill	
ib	Income Before Extraordinary Items	
idit	Total Interest and Related Income	
intan	Total Intangible Assets	
intpn	Net Interest Paid	
lct	Total Current Liabilities	
lt	Total Liabilities	✓
nopi	Nonoperating Income (Expense)	
oancf	Operating Activities - Net Cash Flow	✓
opprft	Operating Profit	

Continued on next page

Compustat Name	Variable	Required non-missing?
pi	Pretax Income	
ppent	Property, Plant and Equipment - Total (Net)	✓
prstkc	Purchase of Common and Preferred Stock	
pstk	Preferred/Preference Stock (Capital) - Total	
pstkn	Preferred/Preference Stock - Nonredeemable	
pstkr	Preferred/Preference Stock - Redeemable	
revt	Total Revenue	✓
sale	Sales/Turnover (Net)	
spi	Special Items	
sstk	Sale of Common and Preferred Stock	
tx	Taxation	
txdb	Deferred Taxes (Balance Sheet)	
txp	Income Taxes Payable	
txpd	Income Taxes Paid	
txt	Income Taxes - Total	
unnp	Unappropriated Net Profit (Stockholders' Equity)	
xi	Extraordinary Items	
xint	Interest and Related Expense - Total	
xpr	Pension and Retirement Expense	
xrd	Research and Development Expense	
xsga	Selling, General and Administrative Expense	
ajexi	Adjustment Factor (International Issue)-Cumulative by Ex-Date	
cshoi	Com Shares Outstanding - Issue	
cshpria	Common Shares Used to Calculate Earnings Per Share (Basic) - As Reported	

Continued on next page

Compustat Name	Variable	Required non-missing?
epsexcon	Earnings Per Share (Basic) - Excluding Extraordinary Items - Consolidated	
epsexnc	Earnings Per Share (Basic) - Excluding Extraordinary Items - Nonconsolidated	
epsincon	Earnings Per Share (Basic) - Including Extraordinary Items - Consolidated	
epsinnc	Earnings Per Share (Basic) - Including Extraordinary Items - Nonconsolidated	

E.1.2. CRSP and Datastream

For gathering firm-specific data on stock prices and market capitalization, I utilized resources from CRSP and Datastream. Following the literature, CRSP is commonly employed for accessing stock price data for firms within North America, specifically the United States and Canada. However, given CRSP's inaccessibility outside of North America, this study relies on Datastream to obtain information on stock prices and market valuations, which are crucial inputs for the econometric forecasting algorithms.

The comprehensive lists of price-related variables used for constructing market valuation inputs are detailed in Table A5 for CRSP (North America) and Table A6 for Datastream (international markets).

TABLE A5. Fundamental Variables from CRSP

Variable	**Required non-missing?**
SIC 2-digit industry code dummie	✓
Return over prior month to fiscal year end t	✓
Return over year prior to fiscal year end t, excluding last month	✓
Market capitalization at the end of year t	✓

TABLE A6. Fundamental Variables from Datastream

Variable	**Required non-missing?**
Adjusted close price on market date t	✓
Return over the day prior to market date t	✓
Market capitalization on market date t	✓

E.2. Macroeconomic Variables

To gather information on the macroeconomic conditions at the country level, I sourced data from Trading Economics and FactSet, serving as crucial inputs for the econometric forecasting algorithms.

TABLE A7. Macro Variables

Variable	**Required non-missing?**
balance_of_trade	✓
consumer_confidence	✓
consumer_price_index_cpi	✓
crude_oil_rigs	✓
currency	✓
exports	✓
gdp_growth_rate	✓
government_bond_10y	✓
government_debt_to_gdp	✓
gross_fixed_capital_formation	✓
gross_national_product	✓
housing_index	✓
imports	✓
inflation_rate	✓
interbank_rate	✓
interest_rate	✓
labor_force_participation_rate	✓
labour_costs	✓
manufacturing_pmi	✓
money_supply_m0	✓

Continued on next page

Variable	**Required non-missing?**
money_supply_m1	✓
money_supply_m2	✓
money_supply_m3	✓
productivity	✓
retail_sales_mom	✓
services_pmi	✓
stock_market	✓
unemployment_rate	✓

Note: This is the standard list of macroeconomic variables used in our analysis. For most countries in my sample, these variables are fully available. However, due to limitations in public data disclosure, a few countries may have incomplete macroeconomic coverage.

Appendix F. Descriptive Statistics

Table A8 and Table A9 provide summary statistics of the key variables used in the empirical analysis, separately at the firm-forecast horizon level and the analyst-forecast level. The statistics are computed across all available forecast horizons (1 to 4 quarters and 1 to 4 years) and both quarterly and annual frequencies.

Table A8 summarizes firm-level forecast information, including the consensus forecast F_{it}^h, realized earnings π_{it+h}, and the normalized forecast error $(F_{it}^h - \pi_{it+h})/P_{it}$. It also includes the number of analysts N_{it} and firm characteristics such as total assets. These statistics help characterize the typical magnitude and dispersion of analyst forecasts across firms and time.

The following two tables are both calculated using the data in United Kingdom.

TABLE A8. Firm-level Summary Statistics Across Forecast Horizons

	Count	Mean	SD	10%	25%	50%	75%	90%
$F_{it}^{h=0.25}$	4,197	20.068	30.228	0.100	0.579	9.023	24.542	53.972
$N_{it}^{h=0.25}$	4,197	3.945	4.125	1	1	2	5	9
Total Assets$_{it}^{h=0.25}$	4,197	4,302.910	21,511.639	23.218	74.407	328.477	1,419.900	5,346.100
$F_{it}^{h=0.25} - \pi_{it+h}^{h=0.25}$_q1	4,197	0.068	1.317	-0.933	-0.279	-0.001	0.285	1.135
$F_{it}^{h=0.5}$	4,386	19.610	29.504	0.100	0.650	8.661	24.327	51.303
$N_{it}^{h=0.5}$	4,386	4.056	4.028	1	1	3	5	10
Total Assets$_{it}^{h=0.5}$	4,386	4,095.843	21,054.611	21.644	69.484	289.050	1,319.720	4,826.575
$F_{it}^{h=0.5} - \pi_{it+h}^{h=0.5}$_q2	4,386	0.117	1.497	-1.122	-0.345	0.004	0.416	1.535
$F_{it}^{h=0.75}$	4,194	20.388	30.150	0.100	0.656	9.500	25.032	54.347
$N_{it}^{h=0.75}$	4,194	4.111	4.014	1	1	3	6	10
Total Assets$_{it}^{h=0.75}$	4,194	4,215.651	21,423.881	23.470	73.118	314.984	1,395.782	5,091.200
$F_{it}^{h=0.75} - \pi_{it+h}^{h=0.75}$_q3	4,194	0.220	1.733	-1.286	-0.355	0.028	0.613	2.027
$F_{it}^{h=1^*}$	4,509	20.300	29.755	0.100	1.038	9.690	24.708	53.296
$N_{it}^{h=1^*}$	4,509	4.627	4.521	1	1	3	7	11
Total Assets$_{it}^{h=1^*}$	4,509	3,702.871	19,948.612	16.899	55.112	242.197	1,119	4,382.260
$F_{it}^{h=1^*} - \pi_{it+h}^{h=1^*}$_q4	4,509	0.309	2.030	-1.675	-0.430	0.058	0.862	2.735
$F_{it}^{h=1}$	3,432	23.132	31.921	0.100	1.280	11.970	28.870	61.833
$N_{it}^{h=1}$	3,432	3.561	3.336	1	1	2	5	9
Total Assets$_{it}^{h=1}$	3,432	4,974.461	22,849.198	41.325	121.149	474.606	1,824.050	6,679.060
$F_{it}^{h=1} - \pi_{it+h}^{h=1}$_a1	3,432	0.098	1.433	-1.175	-0.339	0.007	0.413	1.392
$F_{it}^{h=2}$	3,035	25.903	33.171	0.192	2.980	14.667	33.491	69.456
$N_{it}^{h=2}$	3,035	3.618	3.413	1	1	2	5	9
Total Assets$_{it}^{h=2}$	3,035	5,011.709	22,420.154	42.481	128.426	485.300	1,864.655	7,050.120
$F_{it}^{h=2} - \pi_{it+h}^{h=2}$_a2	3,035	0.409	2.314	-1.969	-0.501	0.105	1.231	3.417
$F_{it}^{h=3}$	2,092	30.554	37.030	0.306	3.872	17.835	40.677	84.878
$N_{it}^{h=3}$	2,092	3.352	3.046	1	1	2	5	8
Total Assets$_{it}^{h=3}$	2,092	6,484.571	25,784.675	58.069	198.136	716.563	2,576.054	10,608.690
$F_{it}^{h=3} - \pi_{it+h}^{h=3}$_a3	2,092	0.650	2.689	-2.387	-0.425	0.196	2.032	4.399
$F_{it}^{h=4}$	612	36.545	45.005	0.476	1.697	17.336	54.334	118.851
$N_{it}^{h=4}$	612	1.482	0.965	1	1	1	2	3
Total Assets$_{it}^{h=4}$	612	15,110.166	42,826.644	115.450	557.601	2,091.202	8,532	31,356.700
$F_{it}^{h=4} - \pi_{it+h}^{h=4}$_a4	612	1.052	2.780	-1.772	-0.042	0.243	2.840	5.130

Table A9 presents the distribution of individual analyst forecast errors, defined as $(\pi_{it+h} - F_t^j \pi_{it+h})/P_{it}$. These statistics reveal the dispersion of analyst-level expectations around both the realized outcomes and the consensus, shedding light on the extent of heterogeneity in analyst beliefs and potential informational frictions.

Overall, the descriptive statistics suggest substantial variation in forecast errors across horizons and analysts, motivating the need for a decomposition of forecast inaccuracy into its structural components.

TABLE A9. Analyst-level Forecast Error Summary Statistics

	Count	Mean	SD	10%	25%	50%	75%	90%
$\pi_{it+h}^{h=0.25} - F_t^j \pi_{it+h}^{h=0.25}$	51,402,370	-0.105	0.176	-0.369	-0.241	-0.001	0.005	0.009
$\pi_{it+h}^{h=0.5} - F_t^j \pi_{it+h}^{h=0.5}$	51,383,506	-0.051	0.133	-0.198	-0.140	-0	0.011	0.025
$\pi_{it+h}^{h=0.75} - F_t^j \pi_{it+h}^{h=0.75}$	51,372,185	0.697	1.624	-0.100	-0.005	0.012	0.117	4.464
$\pi_{it+h}^{h=1^*} - F_t^j \pi_{it+h}^{h=1^*}$	51,411,992	1.073	1.908	-0.012	-0.009	0.016	0.399	4.653
$\pi_{it+h}^{h=1} - F_t^j \pi_{it+h}^{h=1}$	19,312,551	0.023	0.311	-0.003	0.004	0.009	0.012	0.014
$\pi_{it+h}^{h=2} - F_t^j \pi_{it+h}^{h=2}$	19,318,903	0.086	0.393	0.019	0.038	0.062	0.079	0.105
$\pi_{it+h}^{h=3} - F_t^j \pi_{it+h}^{h=3}$	974,232	0.597	1.727	-0.201	-0.036	0.056	0.279	3.562
$\pi_{it+h}^{h=4} - F_t^j \pi_{it+h}^{h=4}$	18,457,347	0.178	0.313	0.086	0.147	0.148	0.167	0.277

Appendix G. Key Elements of the Structural Estimation

Table A10 summarizes the key structural parameters in our estimation framework, along with their economic interpretation and sources of identification.

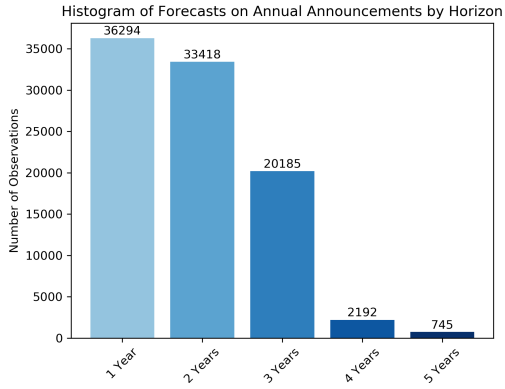
TABLE A10. Structural Parameters and Their Identification

Parameter	**Economic Meaning**	**Identification Source**
α	Analyst responsiveness to soft information Z_{it}	Covariance structure of F^*_{ijt} and E^*_{it}
$\theta = \text{Var}(Z)$	Importance of soft information	Cross-sectional variance of residual forecasts
$\Sigma = \text{Var}(\eta)$	Analyst-specific forecast noise	Variation in F^*_{ijt} across analysts
Δ	Magnitude of model disagreement (common bias)	Mean of $\delta_{it} = (F^f_{it} - F^e_{it})^2$
$\Delta_p = (1-\alpha)^2 \cdot \theta$	Bias induced by soft information	Constructed from α and θ

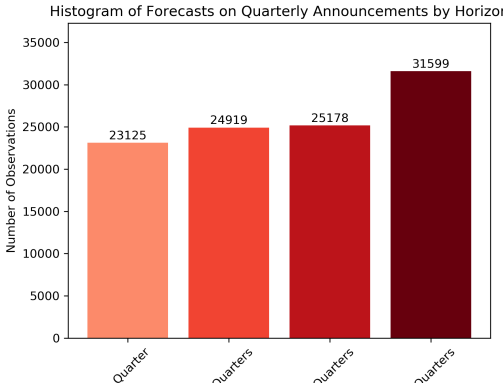
Appendix H. Detailed Empirical Results

H.1. Overview

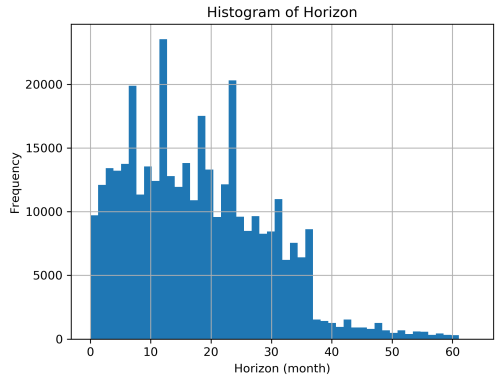
This section presents the distribution of expectation horizons in the United Kingdom, using both annual and quarterly forecast frequencies. Figure A2 provides four complementary views of this distribution, including raw counts and percentage-based representations. The UK is used as a representative example, as the distribution of forecast horizons in this country is broadly similar to those observed in other countries in the sample.



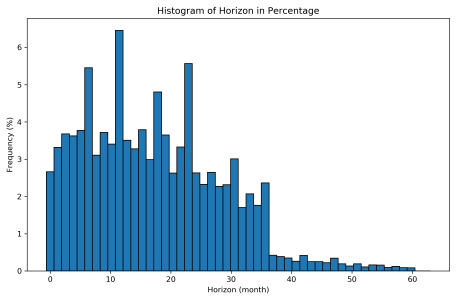
(a) Annual Horizon Counts



(b) Quarterly Horizon Counts



(c) Raw Forecast Horizon Distribution



(d) Forecast Horizons (Percentage)

FIGURE A2. Different views of expectation horizon distributions in the United Kingdom. *Note:* Due to space constraints, only results for the United Kingdom are presented as a representative example. The patterns observed here are broadly consistent with those found across the other 46 countries in the dataset.

A comparison of forecast accuracy between human analysts and machine learning models over the past decade in the United States is presented.[2]

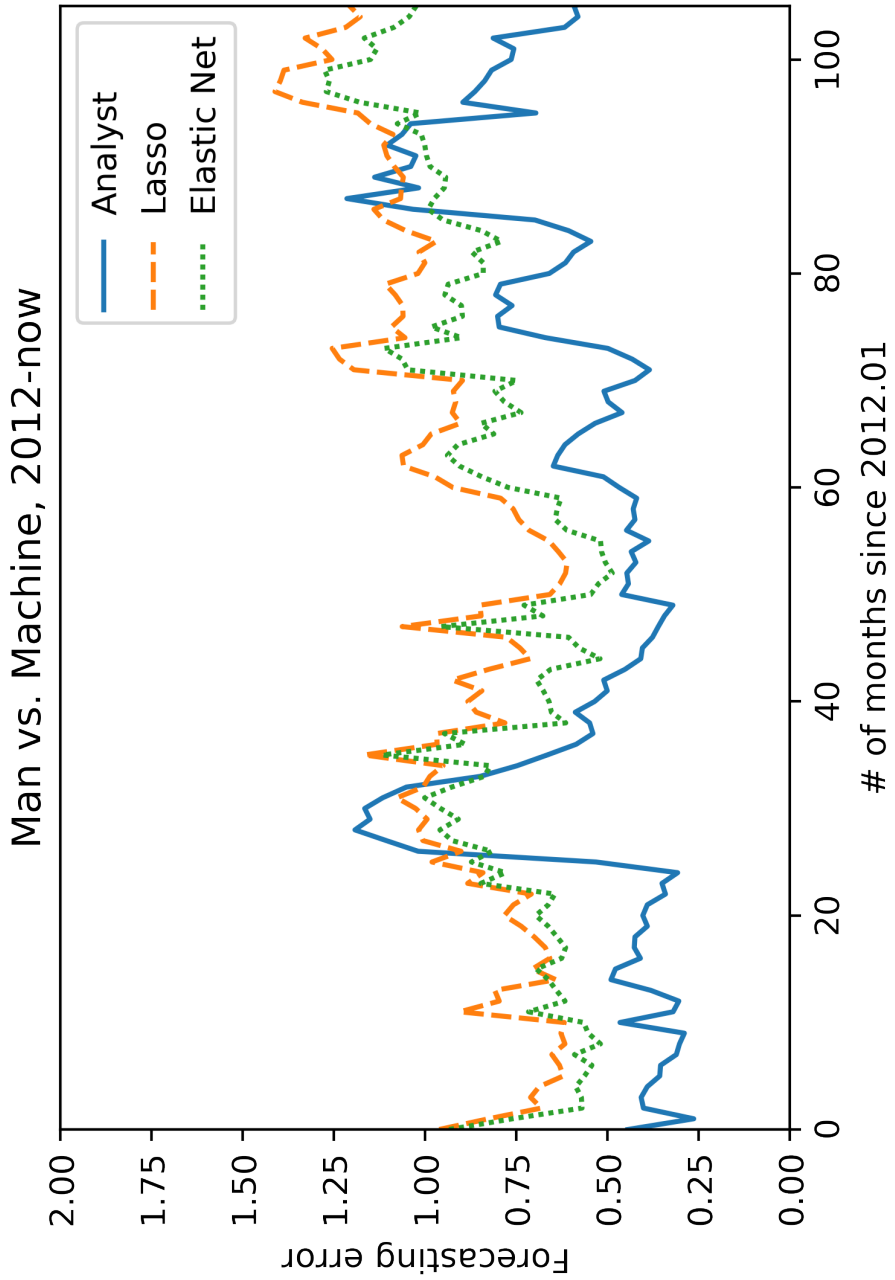


FIGURE A3. Recent 10 years history in United States

[2]All machine learning algorithms were trained using the same input data. For clarity and visual simplicity, the figure below reports results from two representative models, Lasso and Elastic Net. Other algorithms, including Ridge regression, Random Forest, and Gradient-Boosted Trees, yielded broadly similar performance.

Appendix I. Forecasting Accuracy Shock-Time Dynamics

Figure A4 illustrates the evolution of forecasting accuracy over the course of financial crises. Two core dynamics are highlighted:

- **Early phase of shock:** Human forecasters exhibit a significant advantage, driven by their ability to rapidly process soft information such as policy announcements, sentiment shifts, and idiosyncratic disruptions.
- **Later stage of crisis:** As uncertainty persists, forecast bias and noise in human predictions increase. Meanwhile, machine learning models gradually adapt to the evolving data environment, narrowing or even reversing the initial performance gap.

These dynamics reinforce the structural interpretation of forecast errors discussed in the main text and highlight the time-varying nature of informational advantages between human and machine forecasters during systemic disruptions.

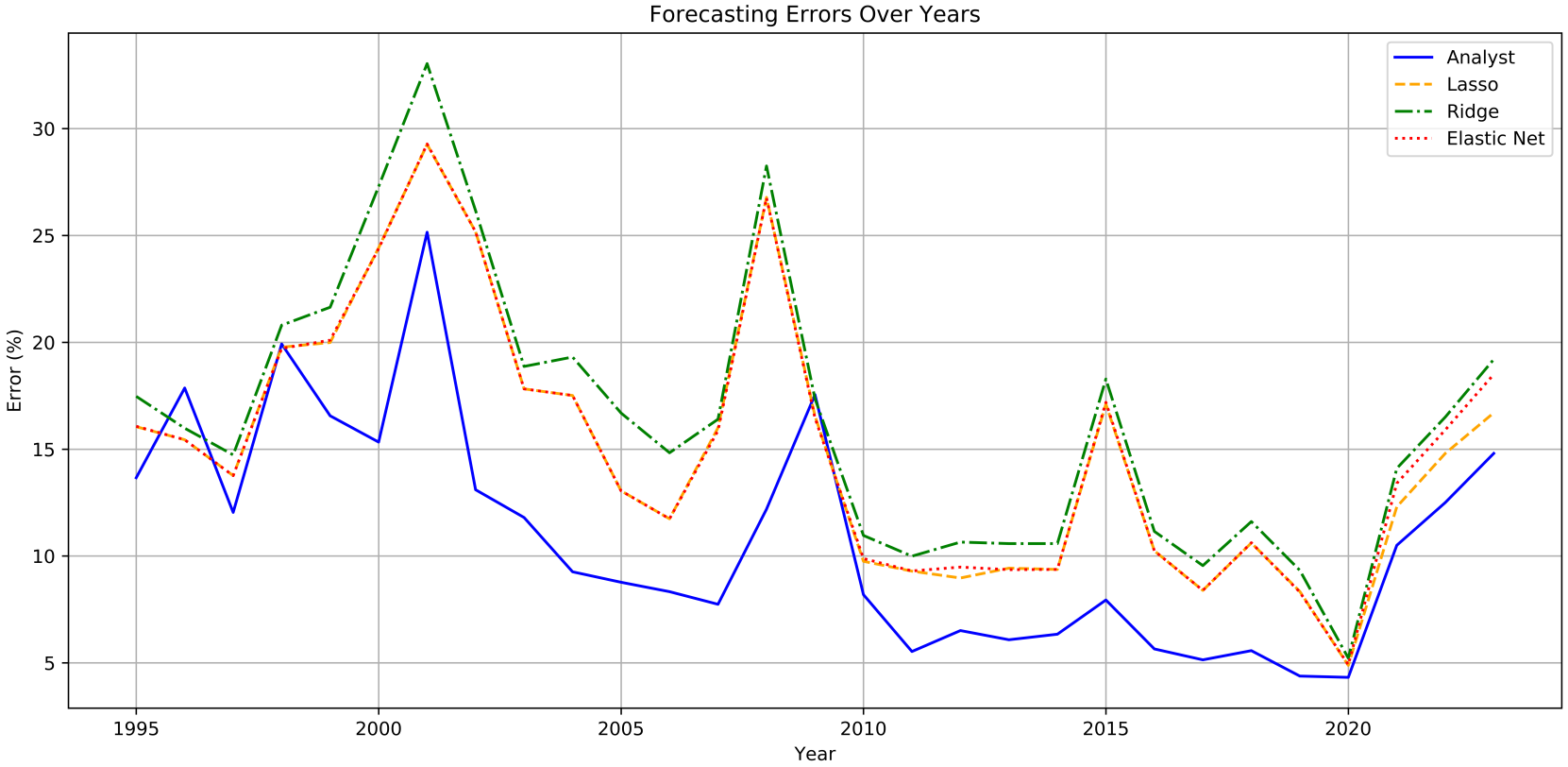


FIGURE A4. Human Forecasting Performance over Crisis Phases in the United States

Appendix J. Term Structure

J.1. General Trends in the Term Structure: Man vs Machine

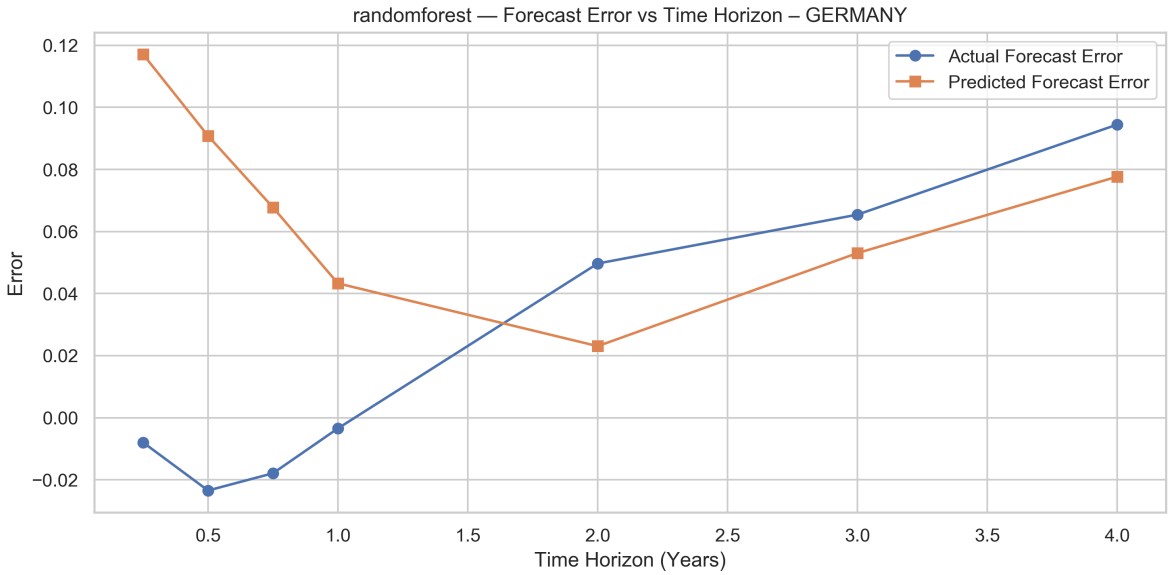


FIGURE A5. Term Structure of Forecast Errors over Time.

Figure A5 illustrates the distinct term structures of forecast errors for human analysts and machine learning models, as discussed in Section **??**. This figure highlights that human forecast errors generally increase with the forecast horizon, suggesting a decline in accuracy over longer periods. Conversely, the machine learning model's forecast error remains relatively stable across horizons, indicating its insensitivity to the forecast horizon. This divergence underscores the differing strengths of human forecasters, particularly in short-term interpretation, versus the consistent pattern-recognition capabilities of machine learning models across various horizons.

J.2. Decomposed Components

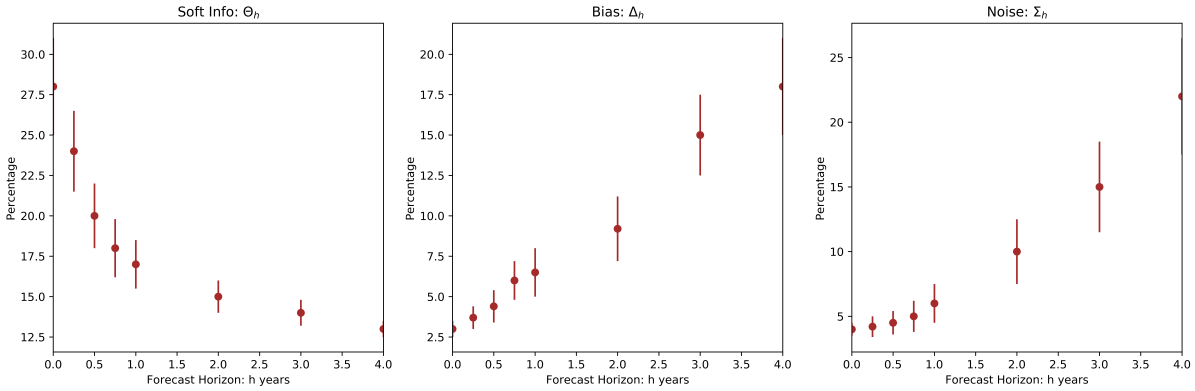
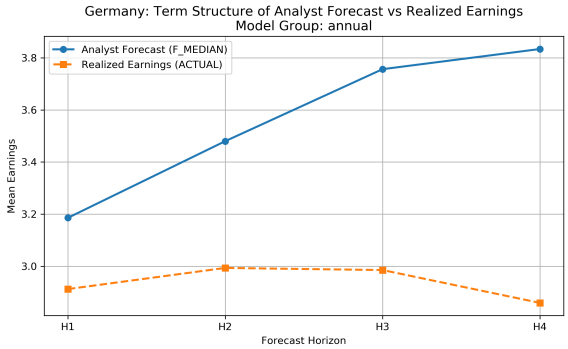


FIGURE A6. Term Structures of Soft Information, Bias, and Noise in Human Forecasts

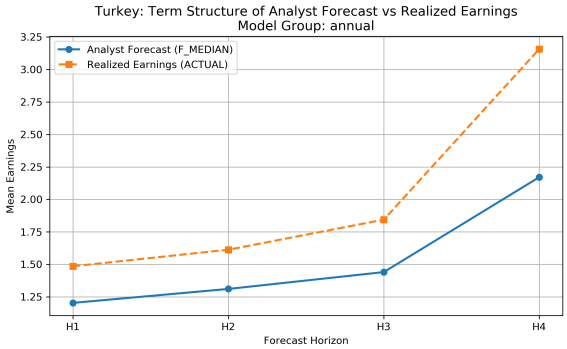
Figure A6 visually presents the decomposition of human forecast adjustments into soft information, bias, and noise across varying forecast horizons. As discussed in Section 3.5.2, this figure illustrates the decreasing contribution of soft information and the increasing impact of behavioral bias and noise on human forecasts as the horizon lengthens.

J.3. Analyst Optimism vs Pessimism

3A. Annual Forecast Errors
3B. Quarterly Forecast Errors

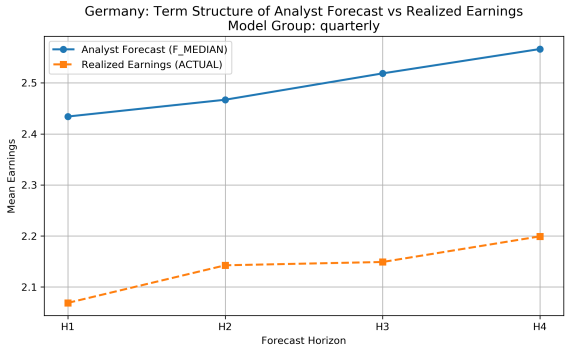


Germany

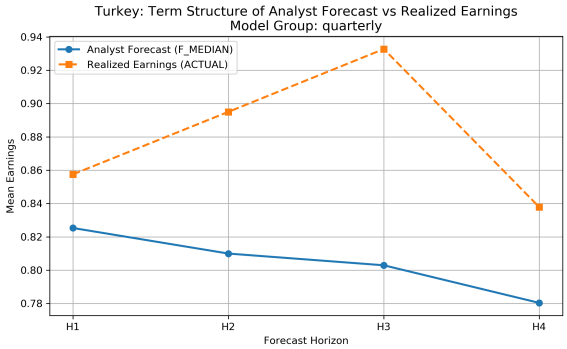


Turkey

FIGURE A7. Term structure of analyst forecast errors based on **annual** earnings forecasts. Forecasts in Germany tend to exhibit a downward bias, whereas forecasts in Turkey display an upward bias. These patterns reflect differences between structured and noisy forecasting environments across countries.



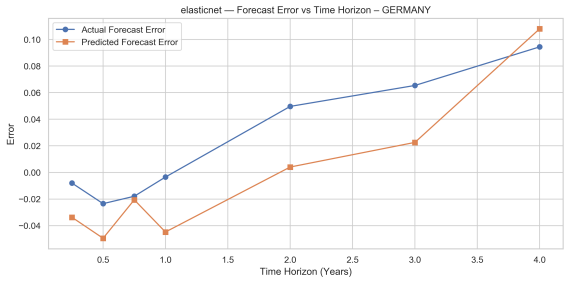
Germany



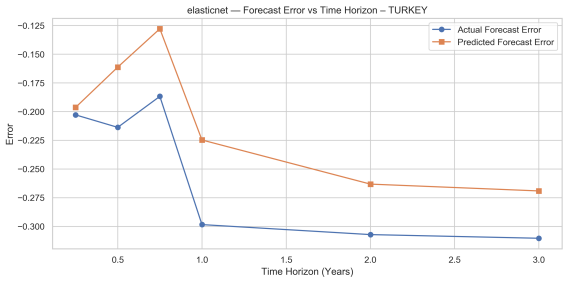
Turkey

FIGURE A8. Term structure of analyst forecast errors based on **quarterly** earnings forecasts. The results echo the annual pattern: German forecasts show consistent pessimism, while Turkish forecasts show optimism. This suggests the bias is persistent across reporting frequencies.

J.4. Machines Learn Human Forecast Bias



Structured Environment (e.g., Germany)



Noisy Forecast Environment (e.g., Turkey)

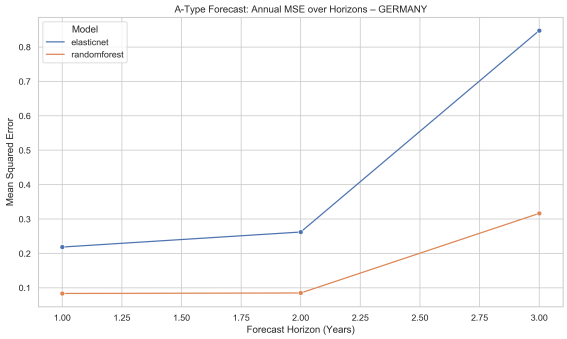
FIGURE A9. Forecast error term structures across horizons for representative structured and noisy forecast environments. Blue lines denote actual analyst forecast errors; orange lines indicate machine-predicted errors using elastic net. Machines consistently capture the direction of analyst bias but forecast more conservatively.

J.5. Cross-Country MSE Patterns in E-type and AE-type Forecasts

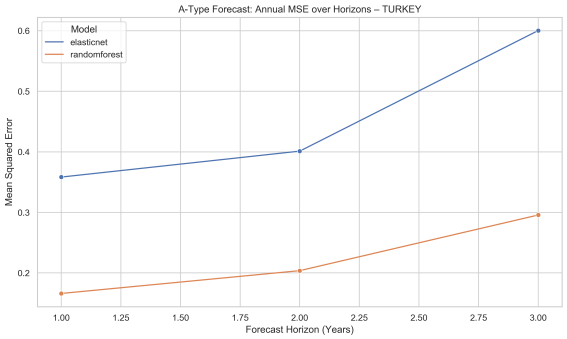
Appendix K. Forecast Accuracy Figures for the United Kingdom

This appendix presents a detailed set of forecast performance figures that aim to illustrate cross-model and cross-setting variations in forecasting accuracy. The visualizations compare results across different machine learning algorithms, forecast horizons, and data frequencies (annual vs. quarterly), and also include comparisons between human analysts and machine-generated forecasts. These figures highlight how forecasting performance evolves across these multiple dimensions.

Due to space constraints, results are presented only for a single representative country—the United Kingdom. Similarly, while all machine learning models were trained on the same input data, only two representative algorithms (Lasso and Elastic Net) are shown here. Other models, including Ridge regression, Random Forest, and Gradient-Boosted Trees, were also tested and yielded broadly similar patterns.

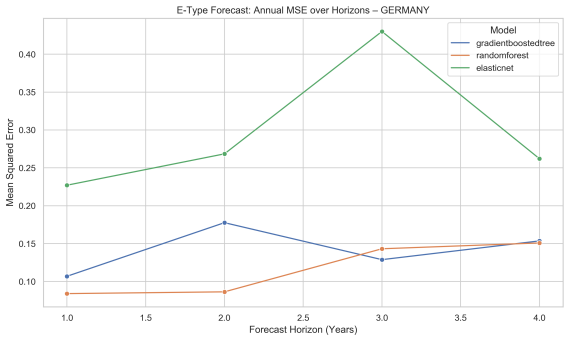


Structured Environment (e.g., Germany)

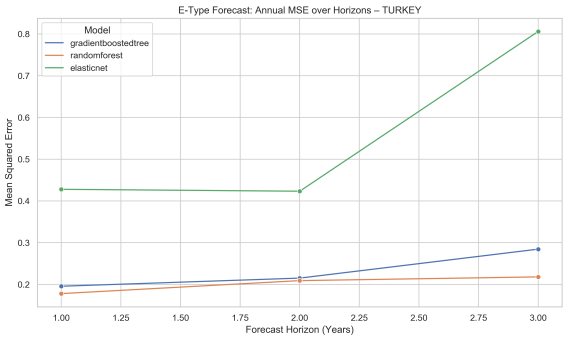


Noisy Forecast Environment (e.g., Turkey)

FIGURE A10. Mean squared error (MSE) of *A-type forecasts*—machine predictions of analyst forecasts—across forecast horizons. Both elastic net and random forest models exhibit rising MSE with horizon, implying shared recognition of increasing human uncertainty across models and countries.

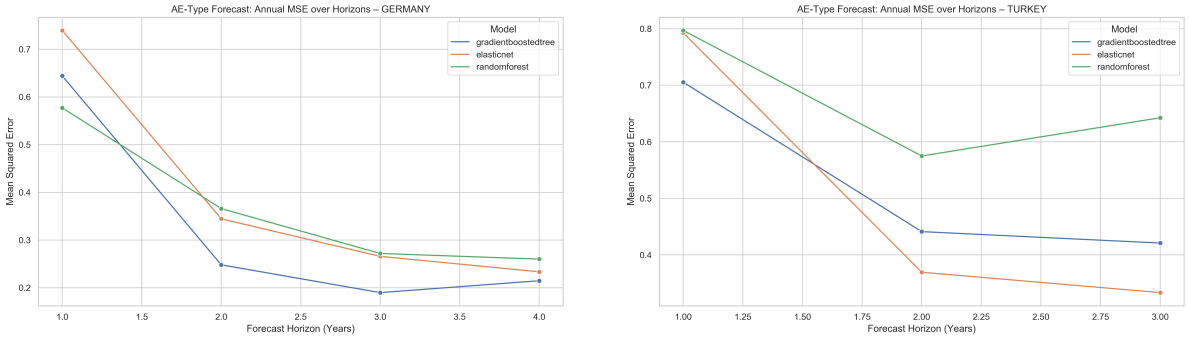


Structured Environment (e.g., Germany)



Noisy Forecast Environment (e.g., Turkey)

FIGURE A11. MSE of *E-type forecasts*—machine predictions of actual earnings—across forecast horizons. As expected, MSE increases with horizon. The increase is more structured in Germany and more erratic in Turkey, consistent with cross-country differences in data quality and firm behavior.



Structured Environment (e.g., Germany) Noisy Forecast Environment (e.g., Turkey)

FIGURE A12. MSE of *AE-type forecasts*—machine predictions of analyst forecast errors—across forecast horizons. MSE decreases with horizon, suggesting that long-term analyst conservatism is more predictable. The pattern is clearer and more stable in Germany.

K.1. Aggregate MSE Comparison (Bar Charts)

These plots summarize average forecast errors across models and horizons, separated by data frequency (annual or quarterly).

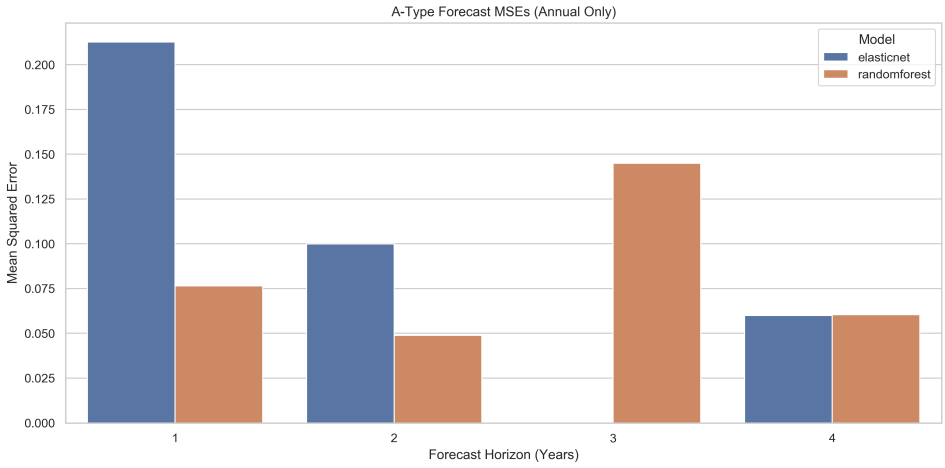


FIGURE A13. A Type Forecast Mse Annual

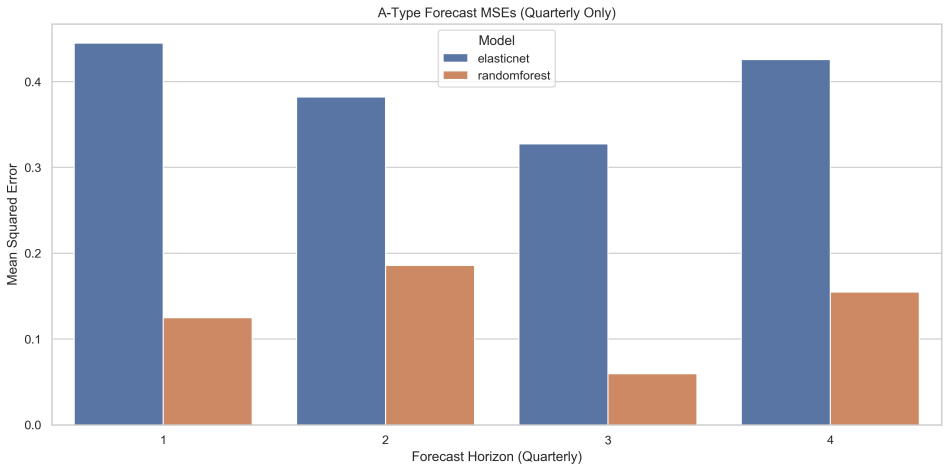


FIGURE A14. A Type Forecast Mse Quarterly

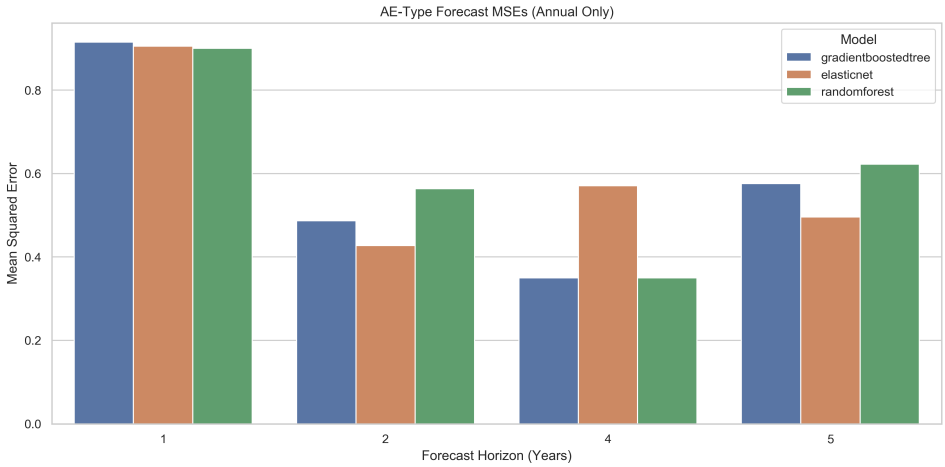


FIGURE A15. Ae Type Forecast Mse Annual

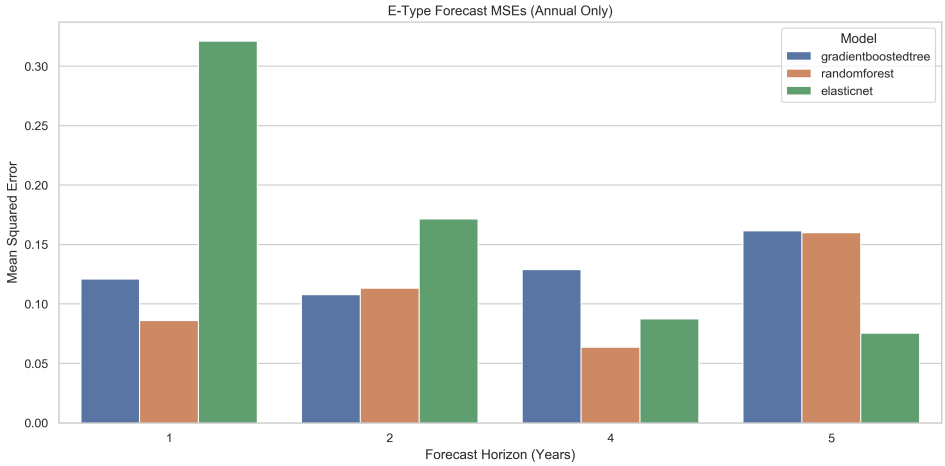


FIGURE A16. E Type Forecast Mse Annual

K.2. Model Performance Overview

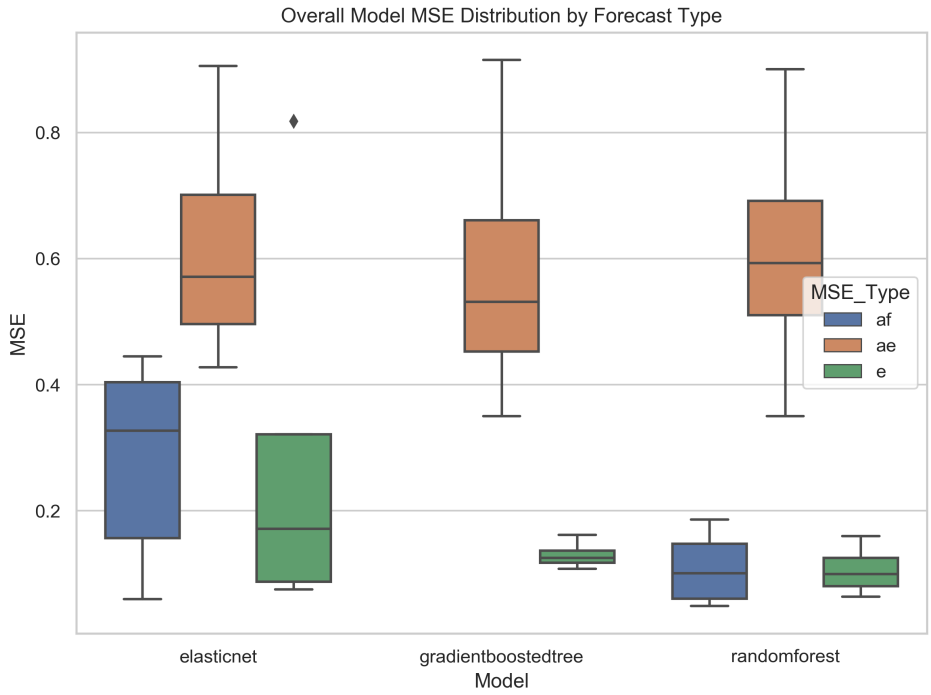


FIGURE A17. Model Mse Boxplot Overall

K.3. Cross Variation by Forecast Type, Algorithm, Frequency & Horizon

K.3.1. Actual vs Predicted

This subsection shows actual vs. predicted values for the Elastic Net model[3] using quarterly data[4]. The figures highlight how prediction accuracy evolves as the forecast horizon increases.

[3] Elastic Net is a regularized linear regression method that combines Lasso and Ridge penalties.

[4] Quarterly frequency corresponds to Horizon = 1 to 3, approximately covering forecast periods of 1 to 3 years.

Elastic Net — Quarterly.

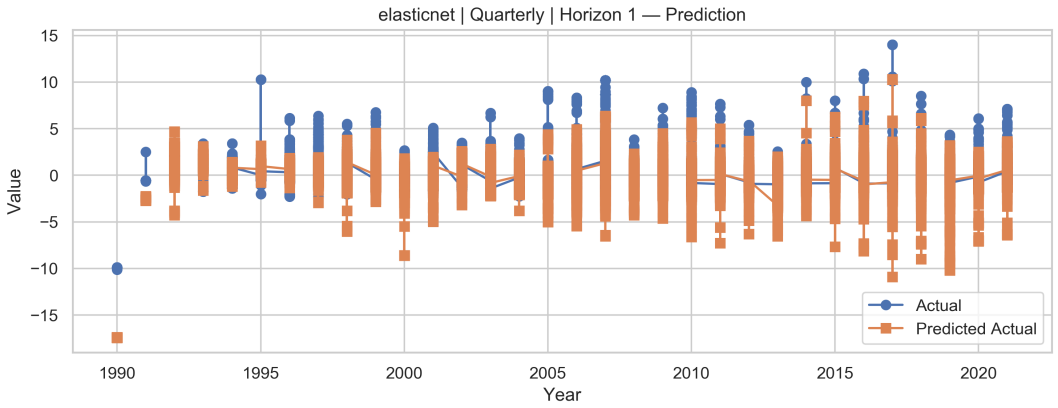


FIGURE A18. Elasticnet Actual Vs Predicted Quarterly H1

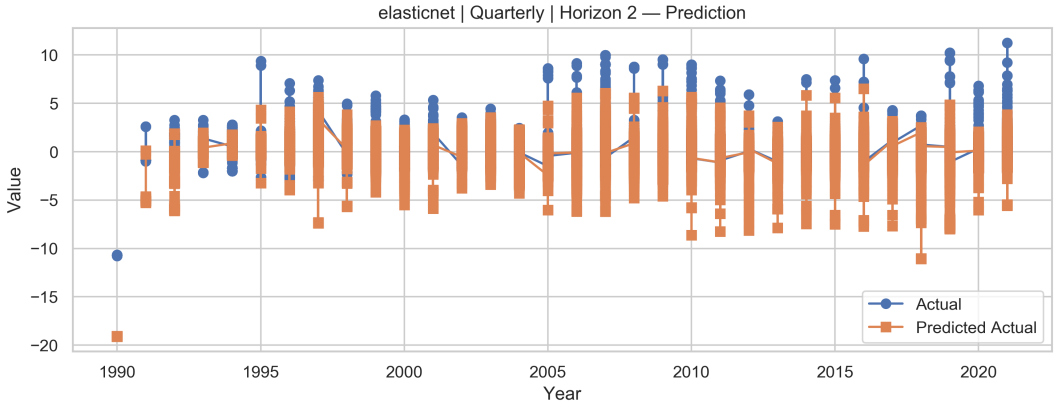


FIGURE A19. Elasticnet Actual Vs Predicted Quarterly H2

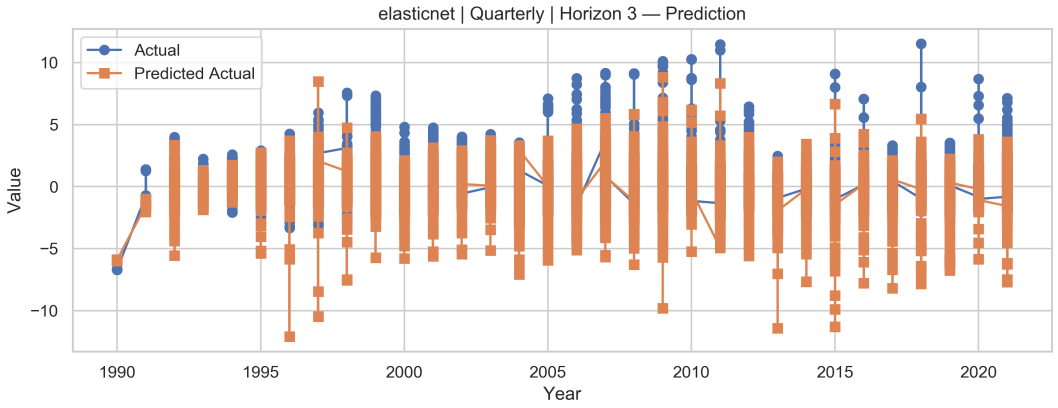


FIGURE A20. Elasticnet Actual Vs Predicted Quarterly H3

Elastic Net — Annual.

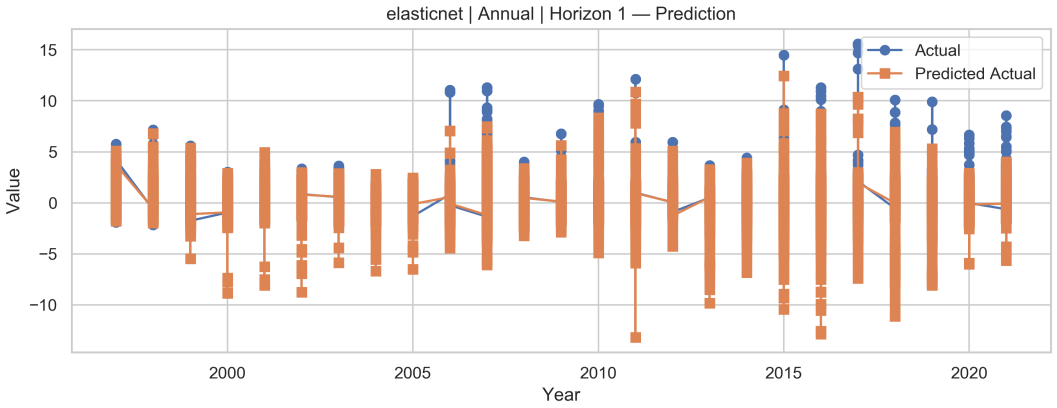


FIGURE A21. Elasticnet Actual Vs Predicted Annual H1

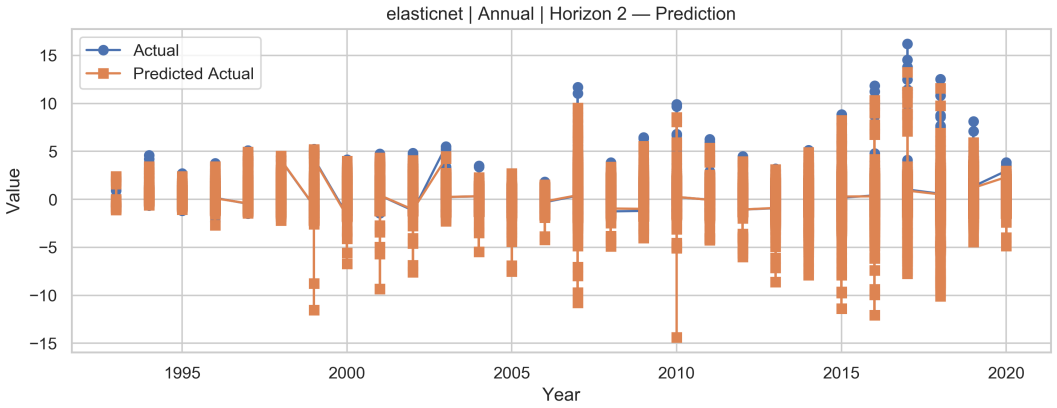


FIGURE A22. Elasticnet Actual Vs Predicted Annual H2

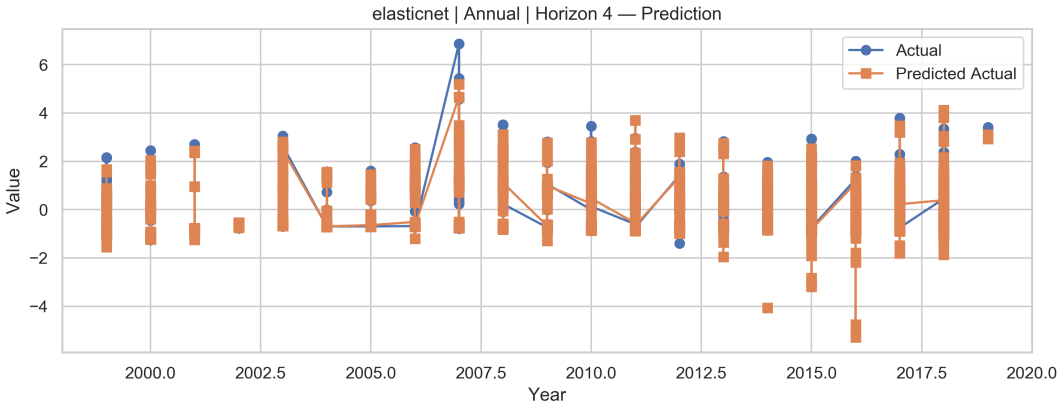


FIGURE A23. Elasticnet Actual Vs Predicted Annual H4

Gradient Boosted Tree — Quarterly.

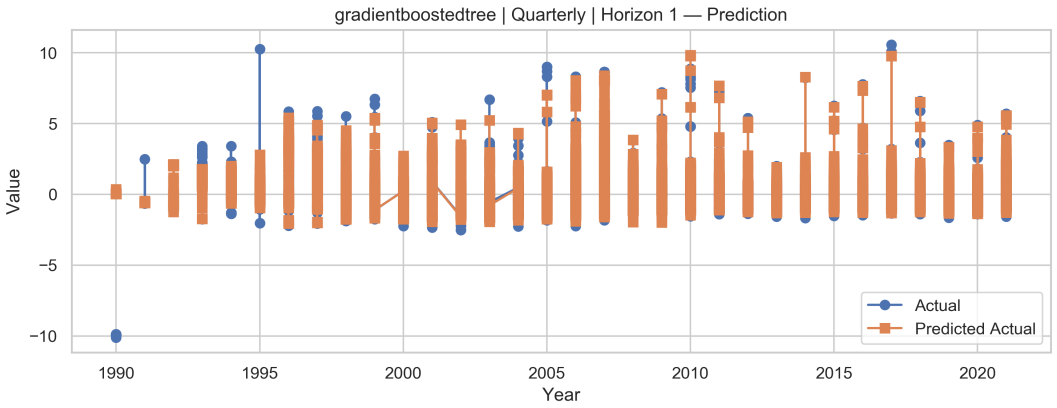


FIGURE A24. Gradient Boosted Tree Actual Vs Predicted Quarterly H1

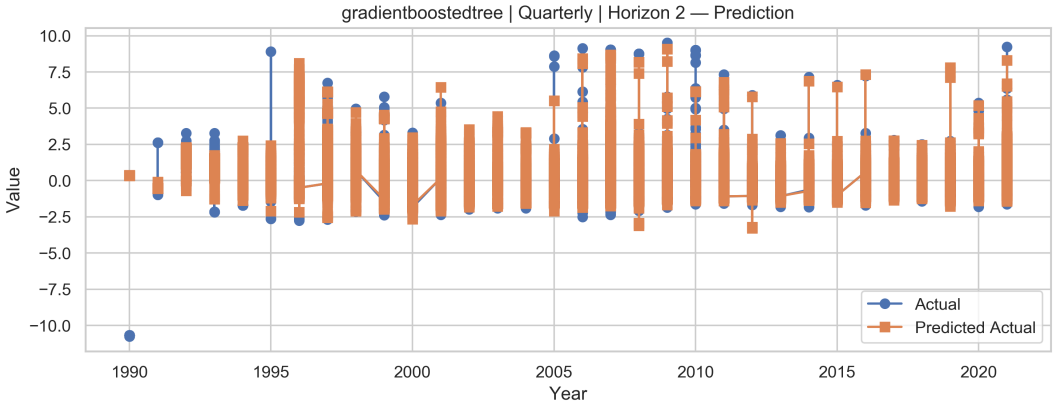


FIGURE A25. Gradient Boosted Tree Actual Vs Predicted Quarterly H2

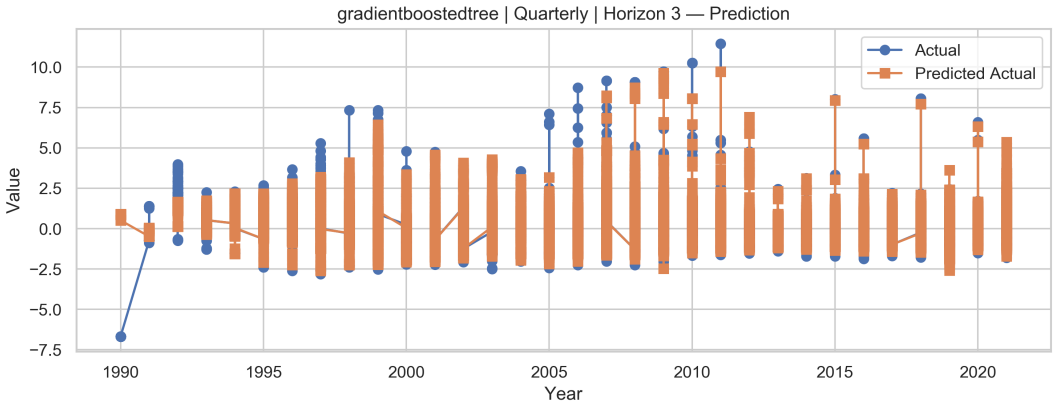


FIGURE A26. Gradient Boosted Tree Actual Vs Predicted Quarterly H3

Gradient Boosted Tree — Annual.

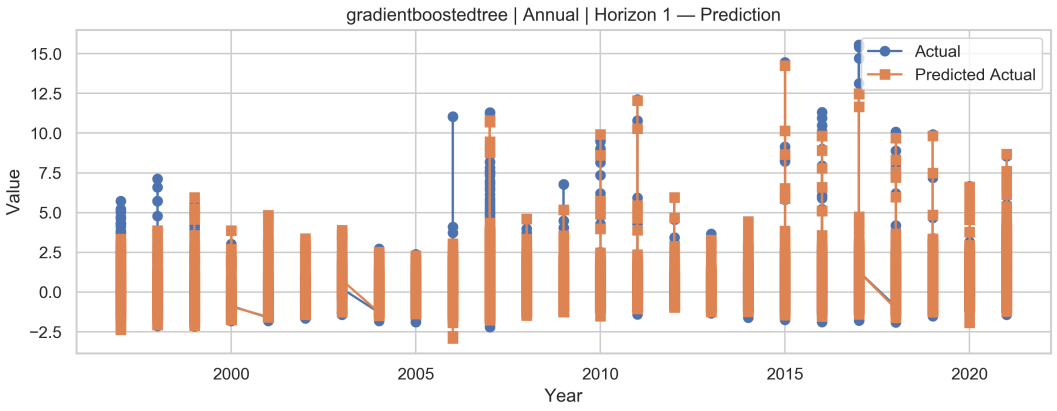


FIGURE A27. Gradient Boosted Tree Actual Vs Predicted Annual H1

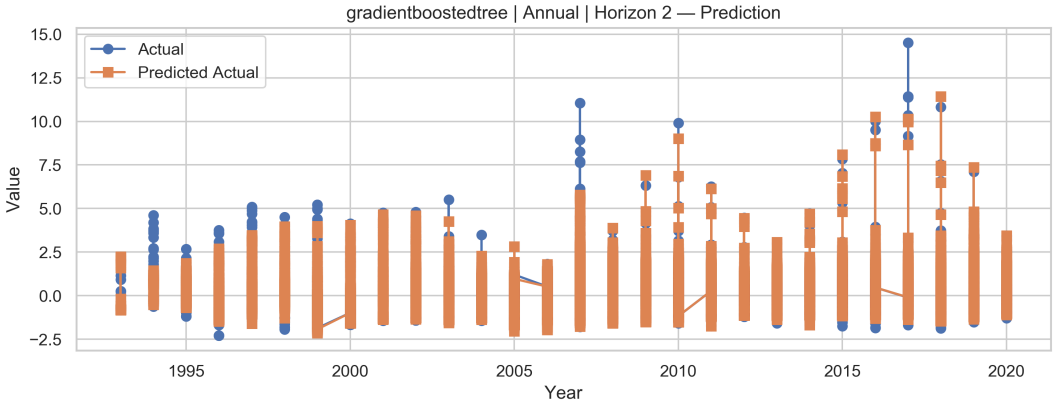


FIGURE A28. Gradient Boosted Tree Actual Vs Predicted Annual H2

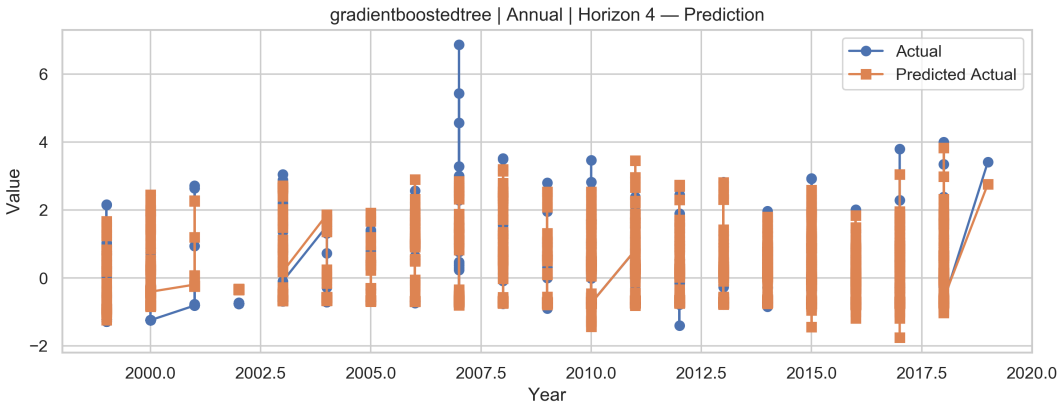


FIGURE A29. Gradient Boosted Tree Actual Vs Predicted Annual H4

Random Forest — Quarterly.

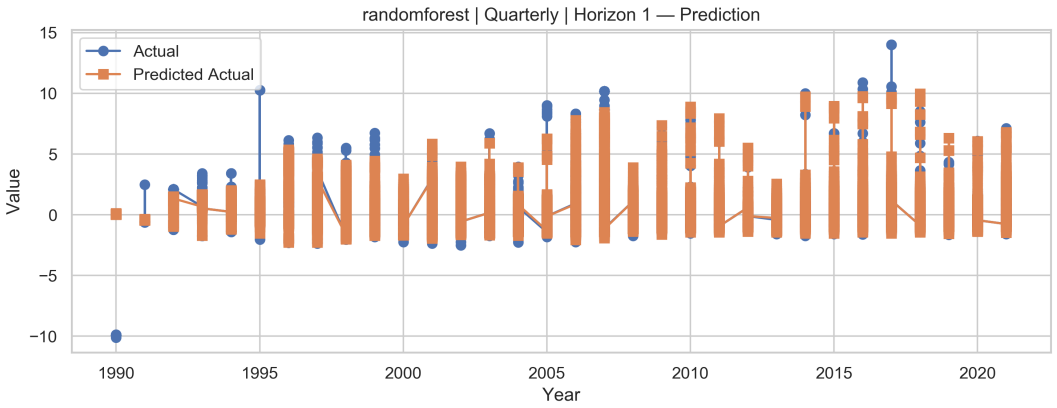


FIGURE A30. Random Forest Actual Vs Predicted Quarterly H1

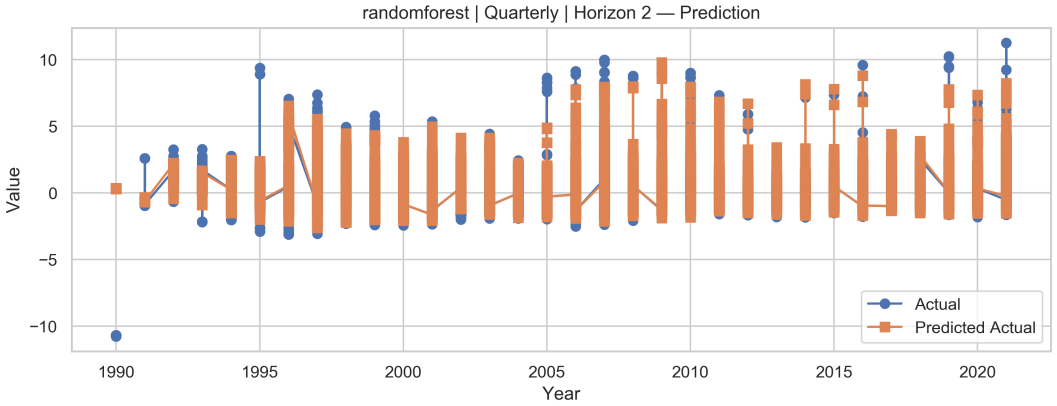


FIGURE A31. Random Forest Actual Vs Predicted Quarterly H2

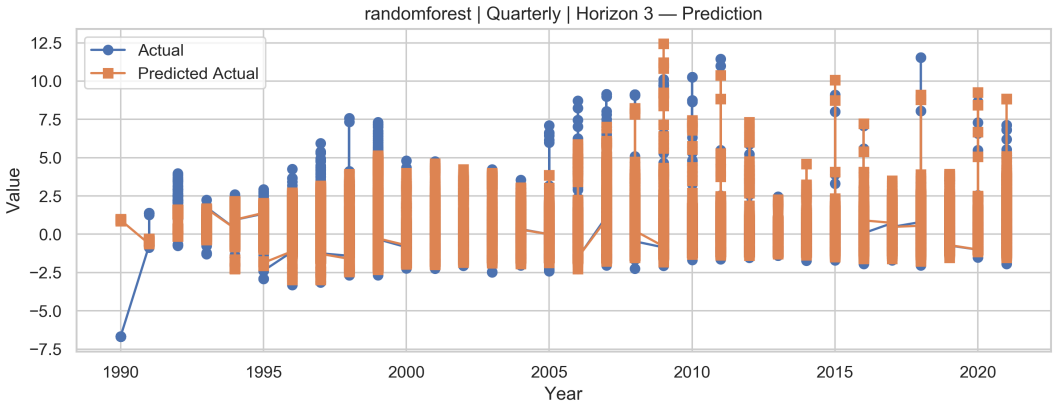


FIGURE A32. Random Forest Actual Vs Predicted Quarterly H3

Random Forest — Annual.

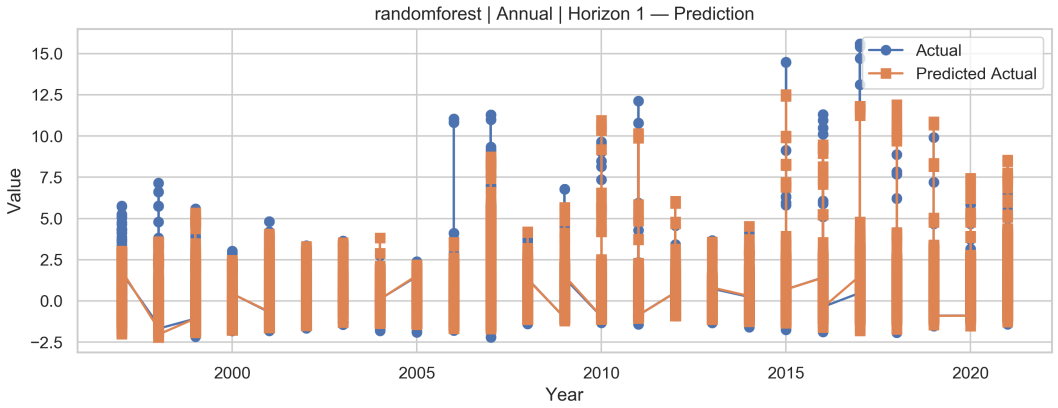


FIGURE A33. Random Forest Actual Vs Predicted Annual H1

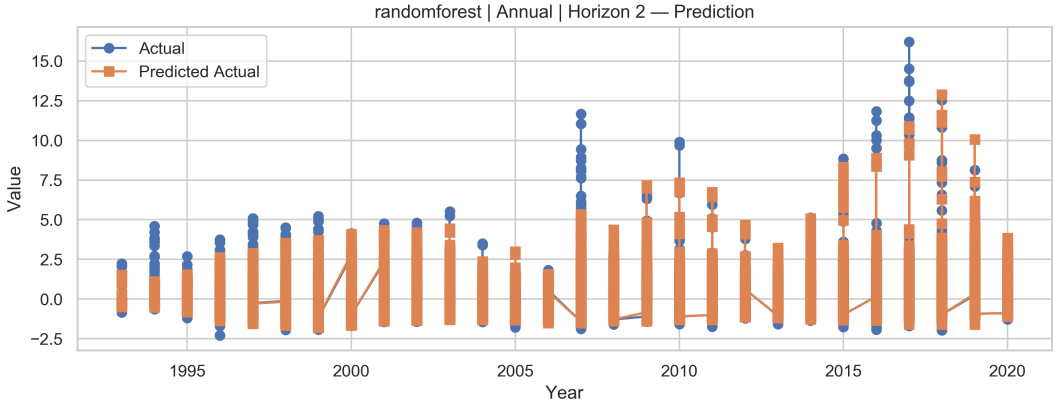


FIGURE A34. Random Forest Actual Vs Predicted Annual H2

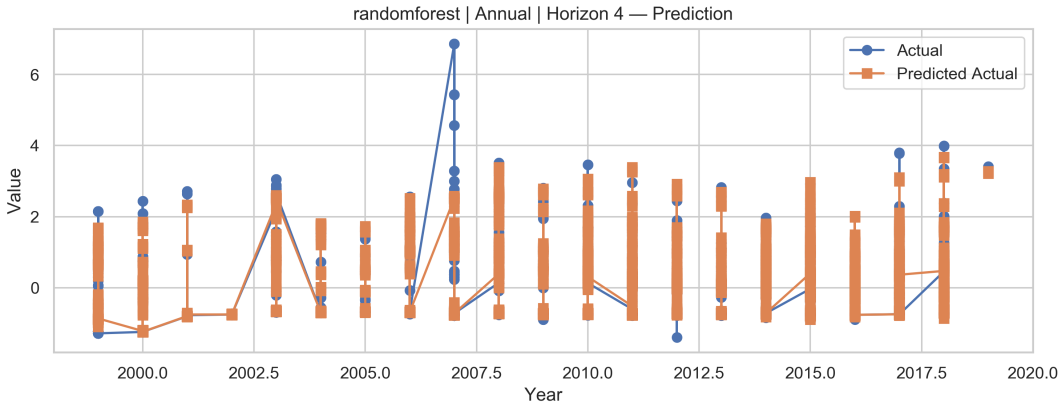


FIGURE A35. Random Forest Actual Vs Predicted Annual H4

K.3.2. Forecast Error

This subsection presents the forecast errors and their predicted counterparts for each model across multiple forecast horizons. The figures allow visual comparison of actual forecast error magnitudes and how well each algorithm captures them. Results are organized by frequency (annual and quarterly) and forecast horizon.[5]

[5]Forecast error is defined as the difference between the forecast and the realized outcome. All models were trained using the same input features.

Elastic Net — Quarterly.

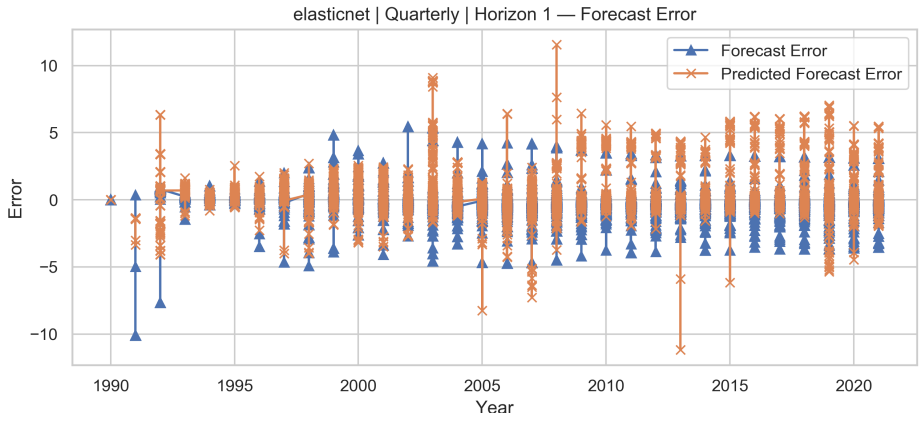


FIGURE A36. Elasticnet Forecast Error Quarterly H1

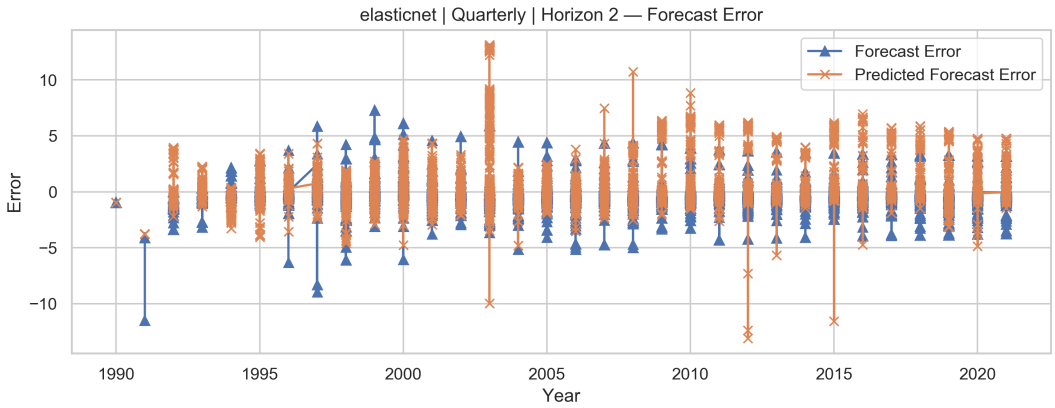


FIGURE A37. Elasticnet Forecast Error Quarterly H2

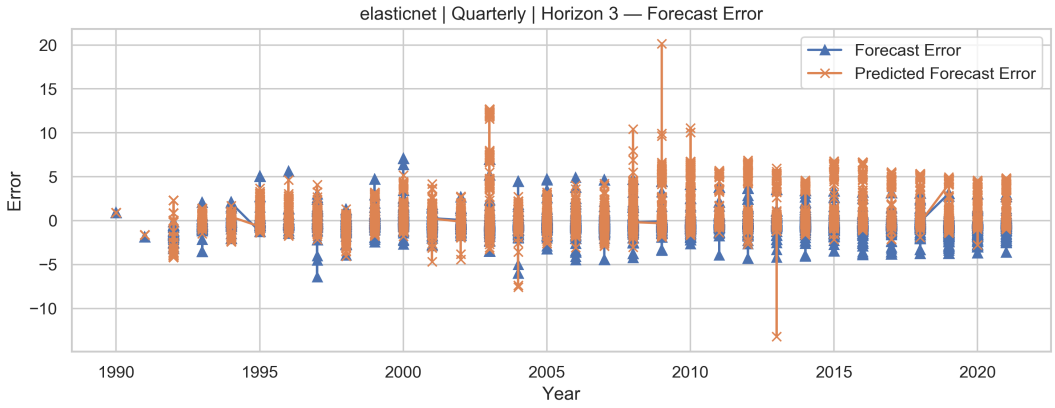


FIGURE A38. Elasticnet Forecast Error Quarterly H3

Elastic Net — Annual.

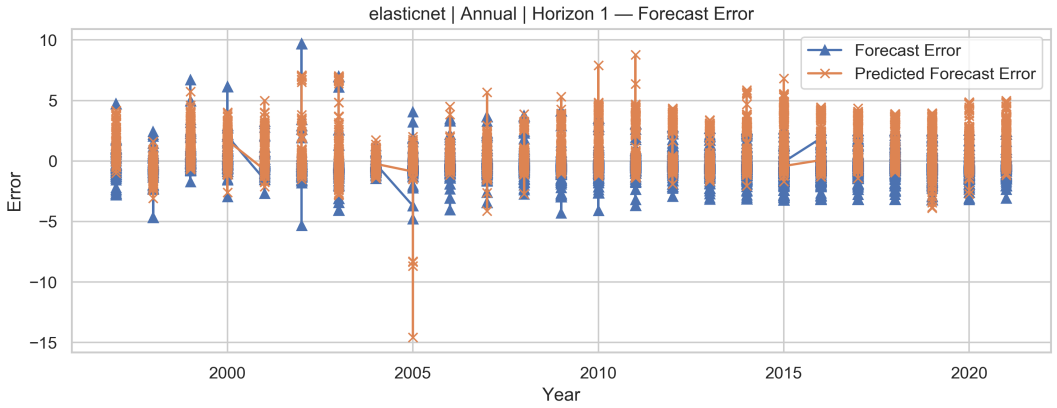


FIGURE A39. Elasticnet Forecast Error Annual H1

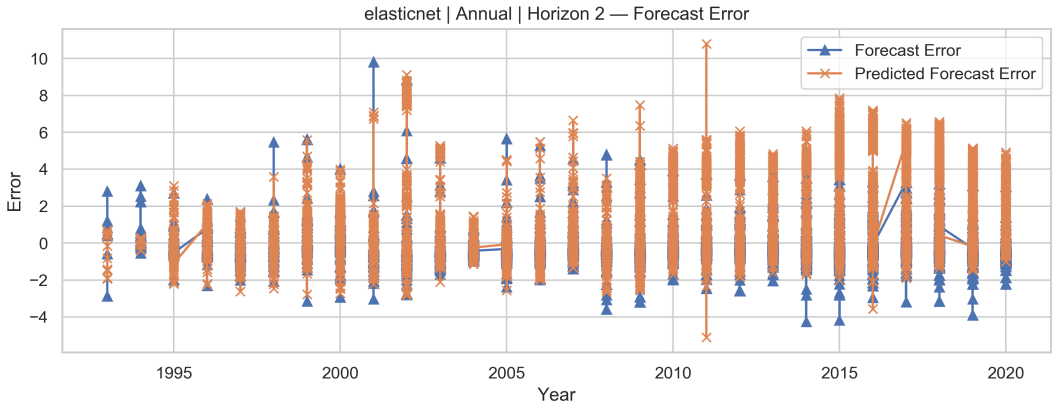


FIGURE A40. Elasticnet Forecast Error Annual H2

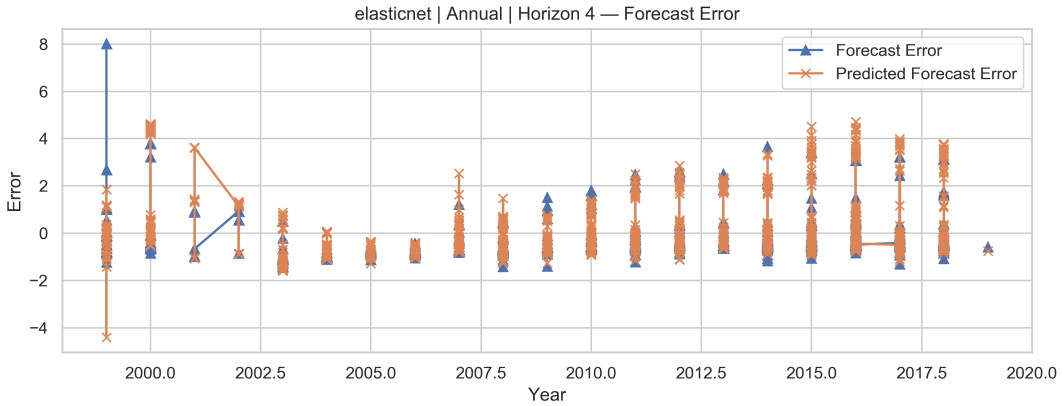


FIGURE A41. Elasticnet Forecast Error Annual H4

Gradient Boosted Tree — Quarterly.

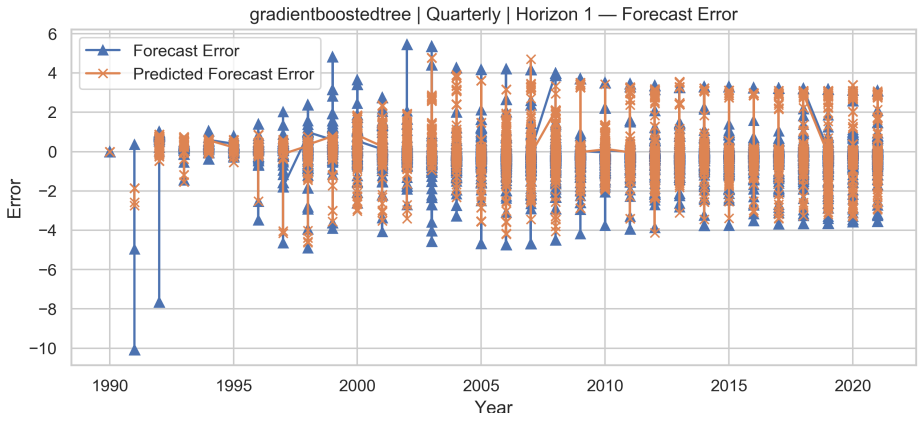


FIGURE A42. Gradient Boosted Tree Forecast Error Quarterly H1

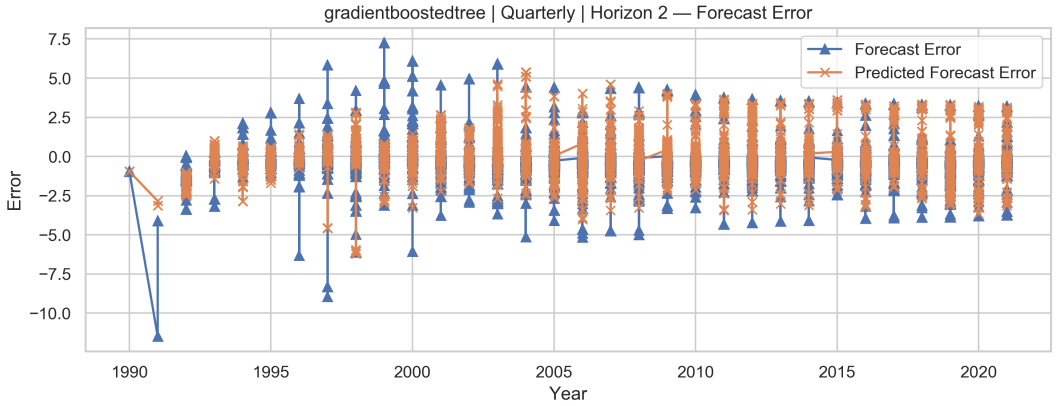


FIGURE A43. Gradient Boosted Tree Forecast Error Quarterly H2

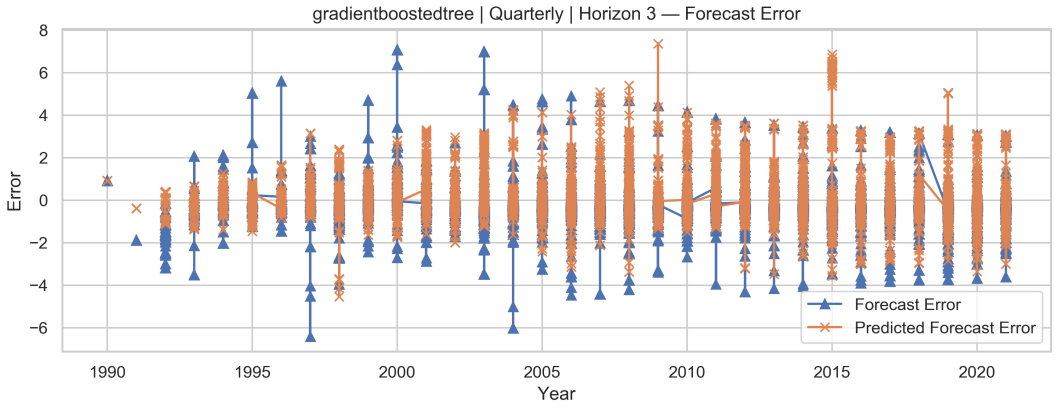


FIGURE A44. Gradient Boosted Tree Forecast Error Quarterly H3

Gradient Boosted Tree — Annual.

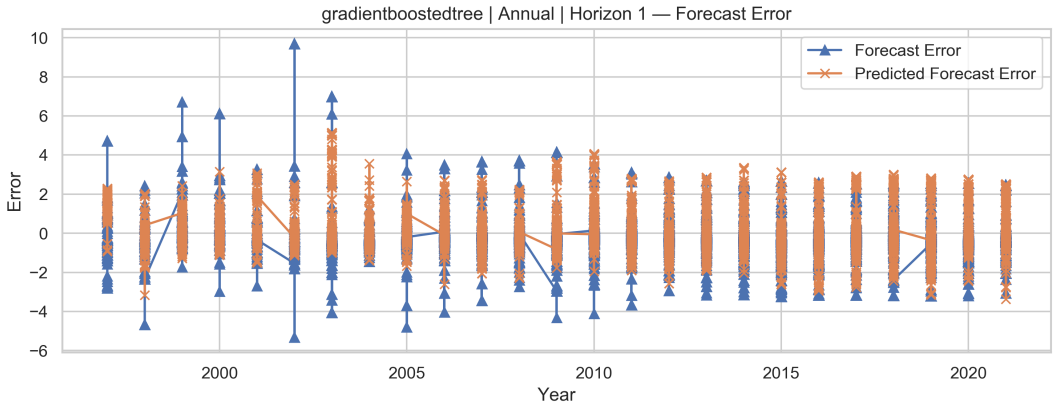


FIGURE A45. Gradient Boosted Tree Forecast Error Annual H1

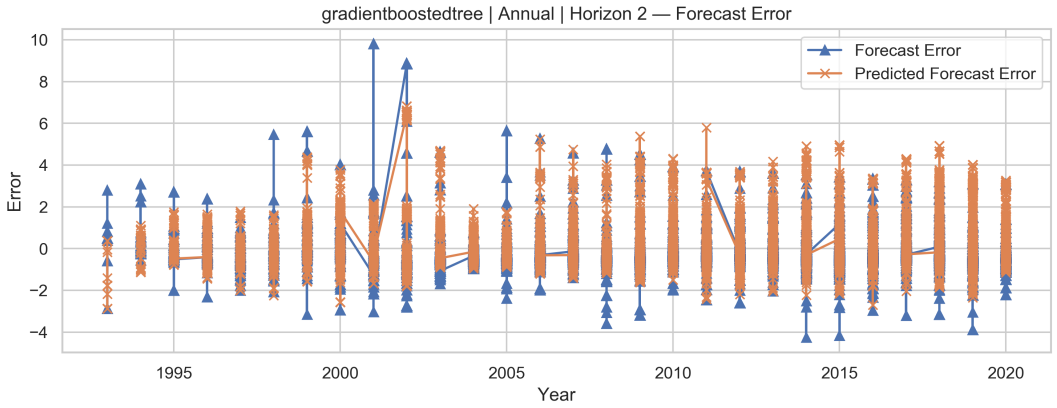


FIGURE A46. Gradient Boosted Tree Forecast Error Annual H2

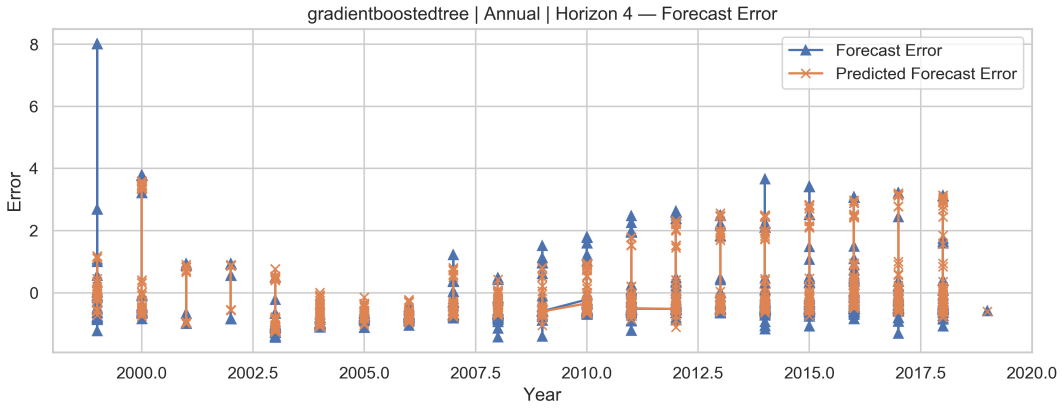


FIGURE A47. Gradient Boosted Tree Forecast Error Annual H4

Random Forest — Quarterly.

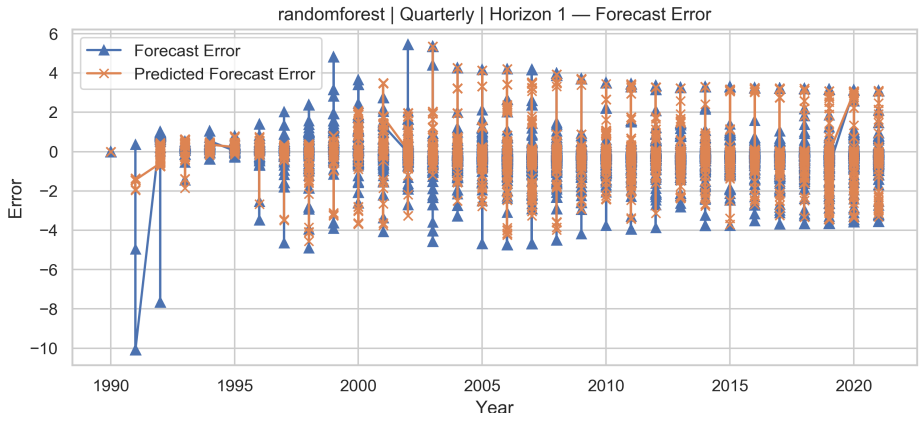


FIGURE A48. Random Forest Forecast Error Quarterly H1

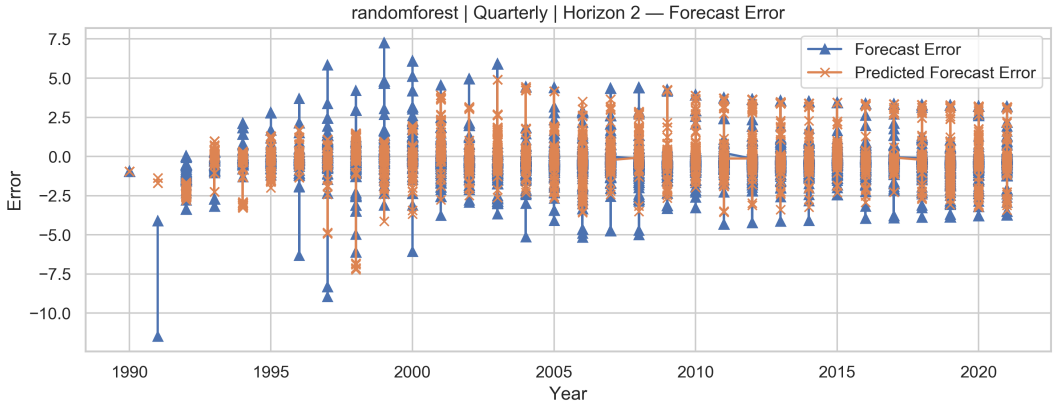


FIGURE A49. Random Forest Forecast Error Quarterly H2

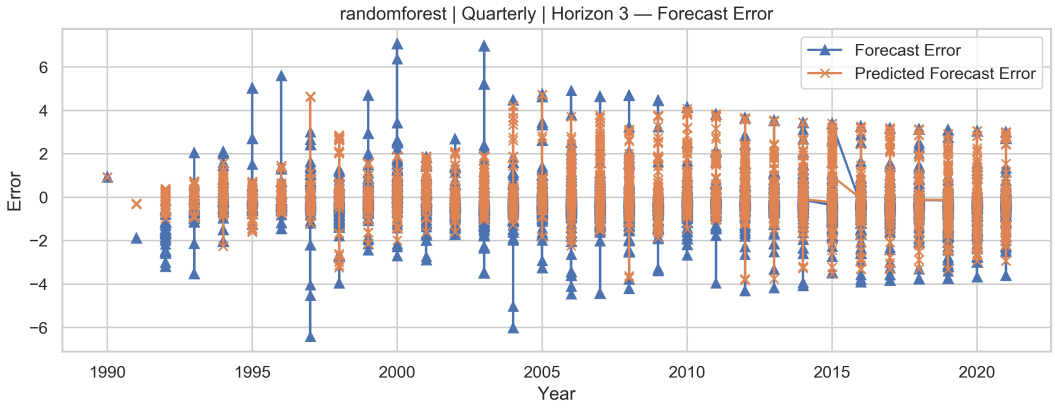


FIGURE A50. Random Forest Forecast Error Quarterly H3

Random Forest — Annual.

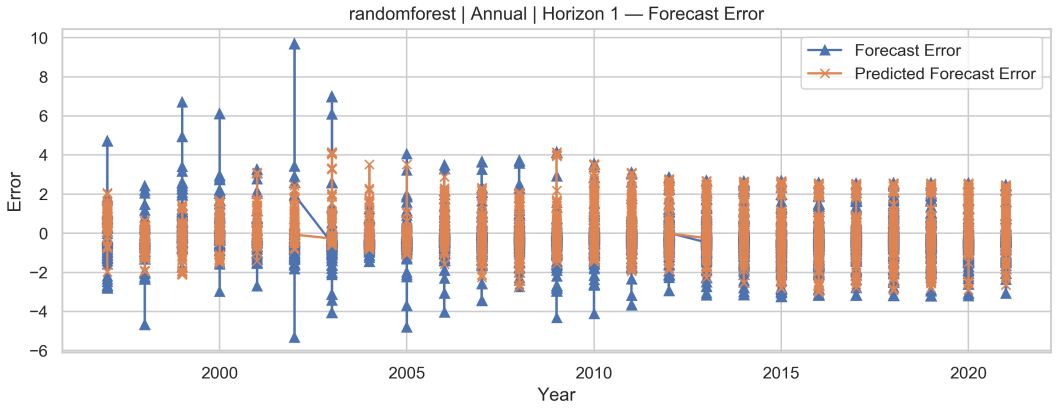


FIGURE A51. Random Forest Forecast Error Annual H1

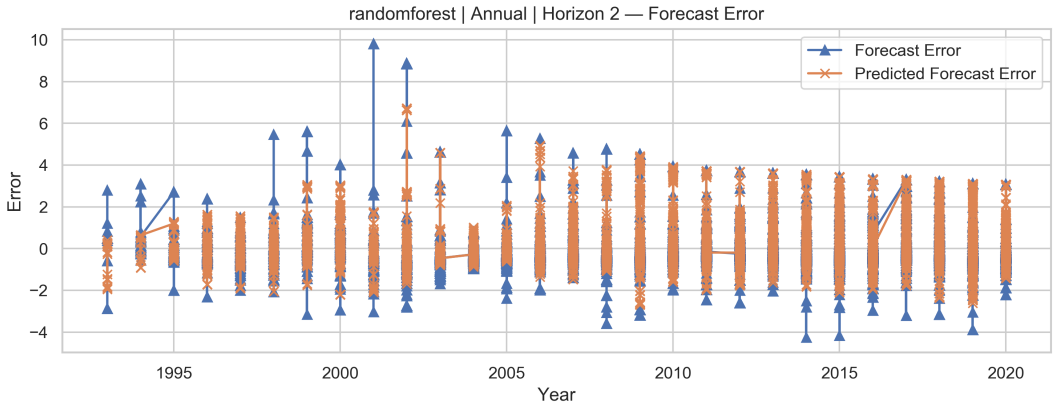


FIGURE A52. Random Forest Forecast Error Annual H2

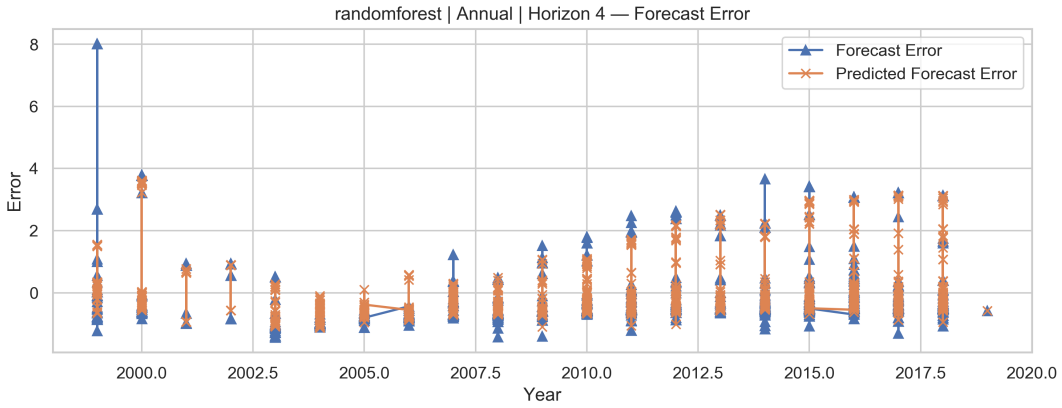


FIGURE A53. Random Forest Forecast Error Annual H4

K.4. Term Structure of Forecast Accuracy

This subsection presents the term structure of forecast accuracy under different machine learning specifications and data frequencies. Each plot shows how the mean squared error (MSE) changes with forecast horizon for multiple forecast types—namely A-type, AE-type, and E-type forecasts. Colored lines represent different forecast types, while each panel reflects either annual or quarterly frequency. These trends reveal how forecast difficulty and model performance vary across forecast horizon, frequency, and forecast component.

A-Type Forecast — Annual Frequency.

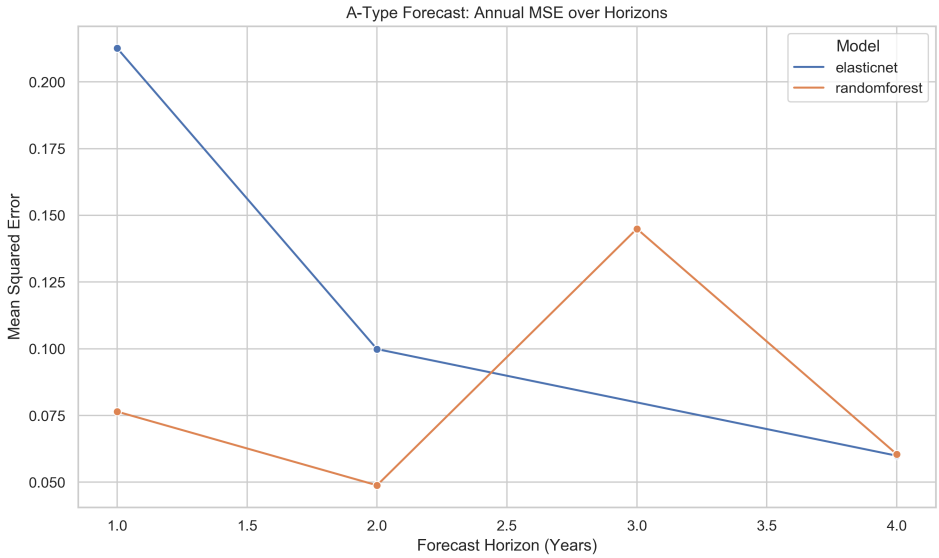


FIGURE A54. Term Structure of A-Type Forecast Accuracy (Annual)

A-Type Forecast — Quarterly Frequency.

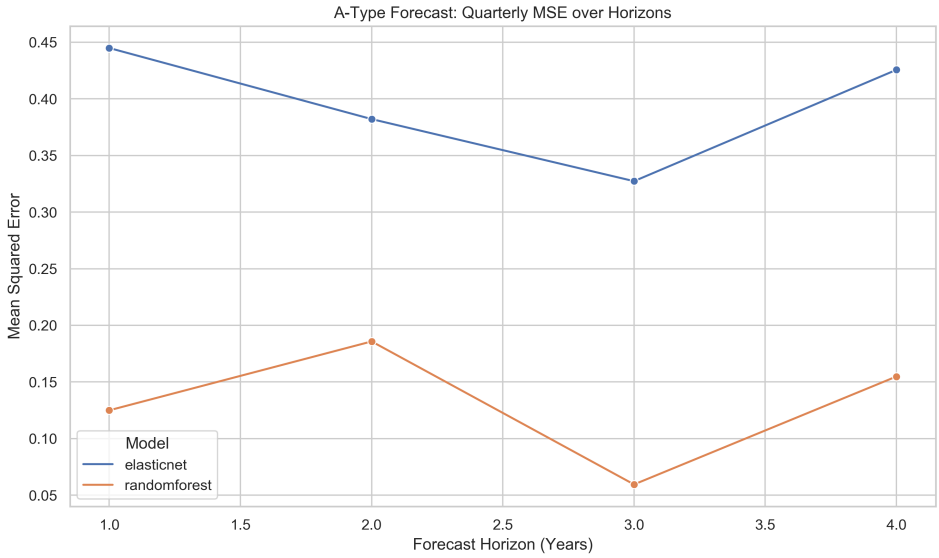


FIGURE A55. Term Structure of A-Type Forecast Accuracy (Quarterly)

AE-Type Forecast — Annual Frequency.

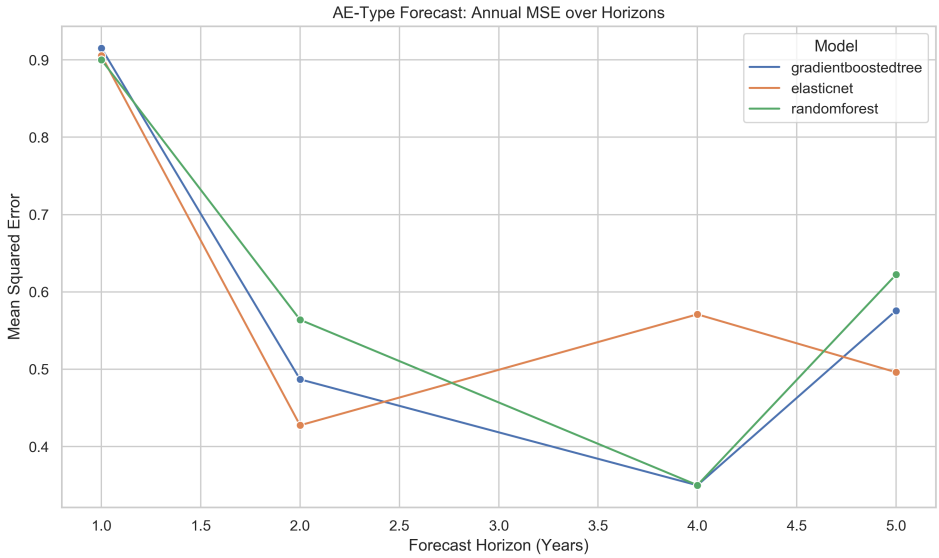


FIGURE A56. Term Structure of AE-Type Forecast Accuracy (Annual)

E-Type Forecast — Annual Frequency.

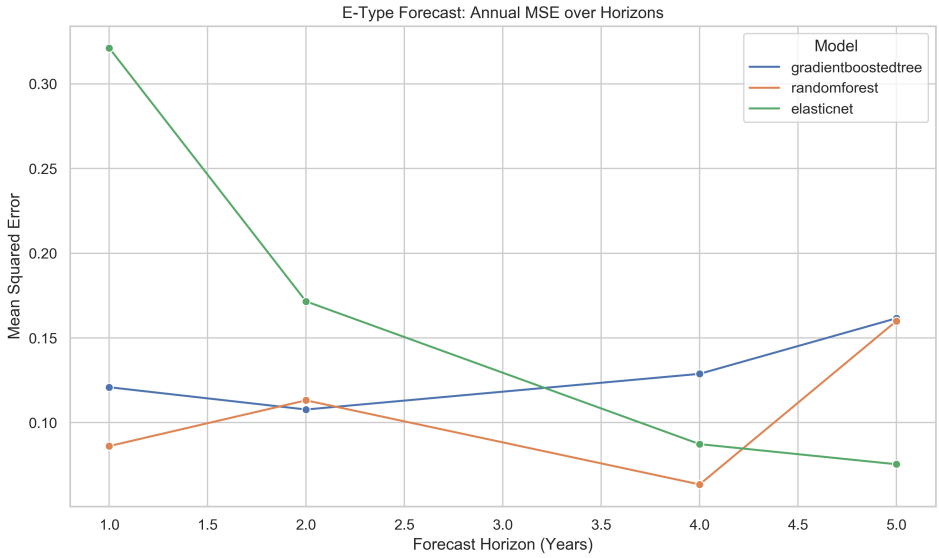


FIGURE A57. Term Structure of E-Type Forecast Accuracy (Annual)

Elastic Net — Forecast Error Term Structure.

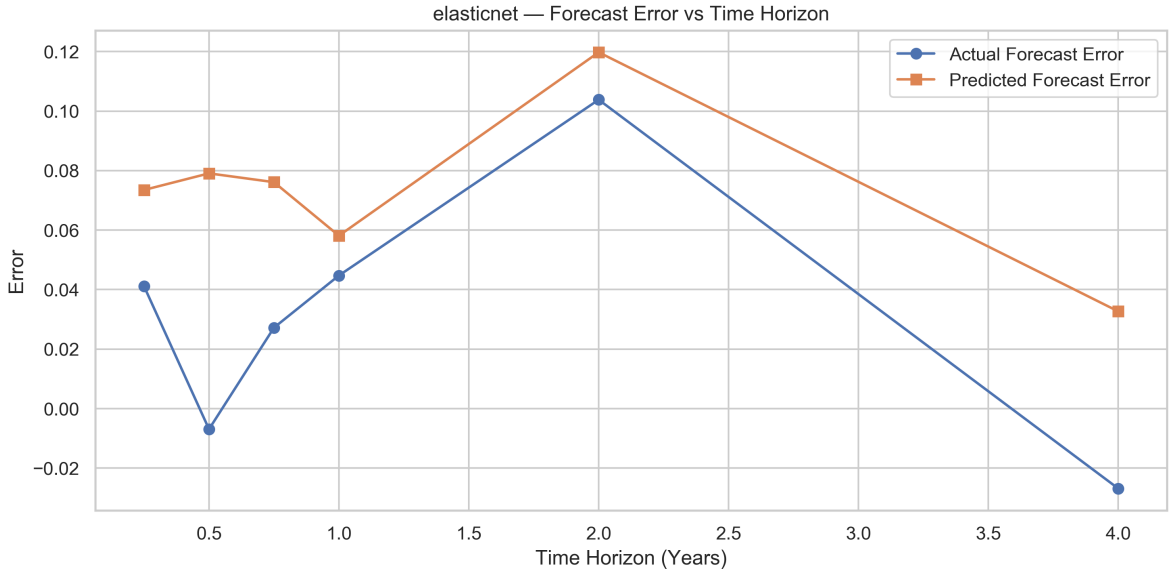


FIGURE A58. Forecast Error Term Structure — Elastic Net

Gradient Boosted Tree — Forecast Error Term Structure.

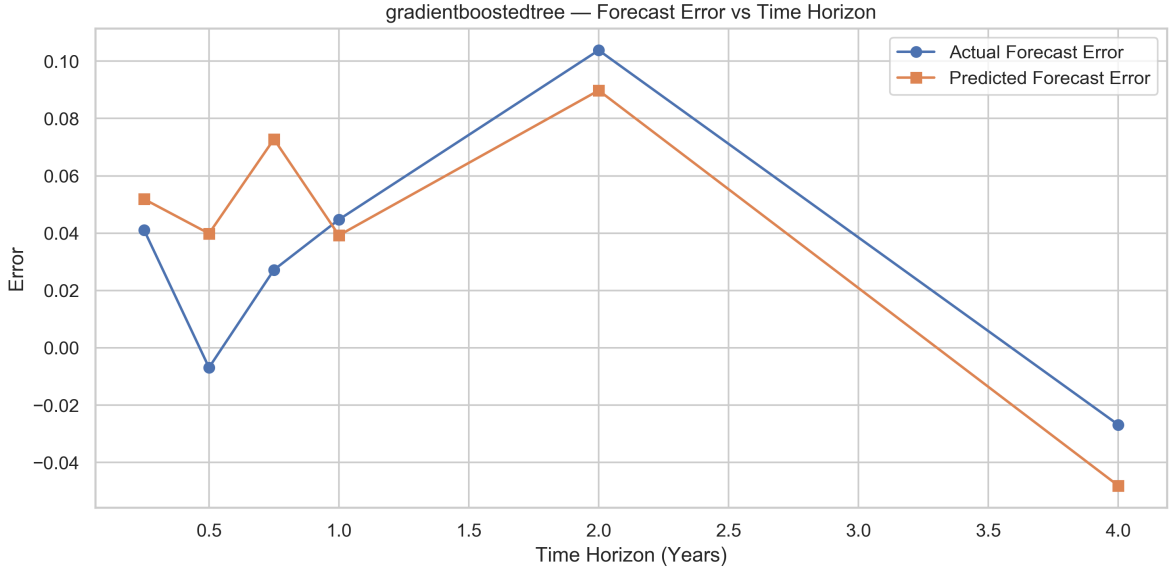


FIGURE A59. Forecast Error Term Structure — Gradient Boosted Tree

	Count	Mean	SD	10%	25%	50%	75%	
$F_{it}^{h=0.25}$	4,197	20.068	30.228	0.100	0.579	9.023	24.542	5
$N_{it}^{h=0.25}$	4,197	3.945	4.125	1	1	2	5	
Total Assets$_{it}^{h=0.25}$	4,197	4,302.910	21,511.639	23.218	74.407	328.477	1,419.900	5,3
$F_{it}^{h=0.25} - \pi_{it+h}^{h=0.25}$_q1	4,197	0.068	1.317	-0.933	-0.279	-0.001	0.285	1
$F_{it}^{h=0.5}$	4,386	19.610	29.504	0.100	0.650	8.661	24.327	5
$N_{it}^{h=0.5}$	4,386	4.056	4.028	1	1	3	5	
Total Assets$_{it}^{h=0.5}$	4,386	4,095.843	21,054.611	21.644	69.484	289.050	1,319.720	4,8
$F_{it}^{h=0.5} - \pi_{it+h}^{h=0.5}$_q2	4,386	0.117	1.497	-1.122	-0.345	0.004	0.416	1
$F_{it}^{h=0.75}$	4,194	20.388	30.150	0.100	0.656	9.500	25.032	5
$N_{it}^{h=0.75}$	4,194	4.111	4.014	1	1	3	6	
Total Assets$_{it}^{h=0.75}$	4,194	4,215.651	21,423.881	23.470	73.118	314.984	1,395.782	5,0
$F_{it}^{h=0.75} - \pi_{it+h}^{h=0.75}$_q3	4,194	0.220	1.733	-1.286	-0.355	0.028	0.613	2
$F_{it}^{h=1^*}$	4,509	20.300	29.755	0.100	1.038	9.690	24.708	5
$N_{it}^{h=1^*}$	4,509	4.627	4.521	1	1	3	7	
Total Assets$_{it}^{h=1^*}$	4,509	3,702.871	19,948.612	16.899	55.112	242.197	1,119	4,3
$F_{it}^{h=1^*} - \pi_{it+h}^{h=1^*}$_q4	4,509	0.309	2.030	-1.675	-0.430	0.058	0.862	2
$F_{it}^{h=1}$	3,432	23.132	31.921	0.100	1.280	11.970	28.870	6
$N_{it}^{h=1}$	3,432	3.561	3.336	1	1	2	5	
Total Assets$_{it}^{h=1}$	3,432	4,974.461	22,849.198	41.325	121.149	474.606	1,824.050	6,6
$F_{it}^{h=1} - \pi_{it+h}^{h=1}$_a1	3,432	0.098	1.433	-1.175	-0.339	0.007	0.413	1
$F_{it}^{h=2}$	3,035	25.903	33.171	0.192	2.980	14.667	33.491	6
$N_{it}^{h=2}$	3,035	3.618	3.413	1	1	2	5	
Total Assets$_{it}^{h=2}$	3,035	5,011.709	22,420.154	42.481	128.426	485.300	1,864.655	7,0
$F_{it}^{h=2} - \pi_{it+h}^{h=2}$_a2	3,035	0.409	2.314	-1.969	-0.501	0.105	1.231	3
$F_{it}^{h=3}$	2,092	30.554	37.030	0.306	3.872	17.835	40.677	8
$N_{it}^{h=3}$	2,092	3.352	3.046	1	1	2	5	
Total Assets$_{it}^{h=3}$	2,092	6,484.571	25,784.675	58.069	198.136	716.563	2,576.054	10,0
$F_{it}^{h=3} - \pi_{it+h}^{h=3}$_a3	2,092	0.650	2.689	-2.387	-0.425	0.196	2.032	4
$F_{it}^{h=4}$	612	36.545	45.005	0.476	1.697	17.336	54.334	11
$N_{it}^{h=4}$	612	1.482	0.965	1	1	1	2	
Total Assets$_{it}^{h=4}$	612	15,110.166	42,826.644	115.450	557.601	2,091.202	8,532	31,3
$F_{it}^{h=4} - \pi_{it+h}^{h=4}$_a4	612	1.052	2.780	-1.772	-0.042	0.243	2.840	5

	Count	Mean	SD	10%	25%	50%	75%	90%
$\pi_{it+h}^{h=0.25} - F_t^j \pi_{it+h}^{h=0.25}$	51,402,370	-0.105	0.176	-0.369	-0.241	-0.001	0.005	0.009
$\pi_{it+h}^{h=0.5} - F_t^j \pi_{it+h}^{h=0.5}$	51,383,506	-0.051	0.133	-0.198	-0.140	-0	0.011	0.025
$\pi_{it+h}^{h=0.75} - F_t^j \pi_{it+h}^{h=0.75}$	51,372,185	0.697	1.624	-0.100	-0.005	0.012	0.117	4.464
$\pi_{it+h}^{h=1^*} - F_t^j \pi_{it+h}^{h=1^*}$	51,411,992	1.073	1.908	-0.012	-0.009	0.016	0.399	4.653
$\pi_{it+h}^{h=1} - F_t^j \pi_{it+h}^{h=1}$	19,312,551	0.023	0.311	-0.003	0.004	0.009	0.012	0.014
$\pi_{it+h}^{h=2} - F_t^j \pi_{it+h}^{h=2}$	19,318,903	0.086	0.393	0.019	0.038	0.062	0.079	0.105
$\pi_{it+h}^{h=3} - F_t^j \pi_{it+h}^{h=3}$	974,232	0.597	1.727	-0.201	-0.036	0.056	0.279	3.562
$\pi_{it+h}^{h=4} - F_t^j \pi_{it+h}^{h=4}$	18,457,347	0.178	0.313	0.086	0.147	0.148	0.167	0.277