Mimicking Finance[*]

Lauren H. Cohen Yiwen Lu Quoc H. Nguyen[†]

January 27, 2026

Abstract

We use frontier advancements in Artificial Intelligence and machine learning to extract and classify the part of key economic agents' behaviors that are predictable from past behaviors. Even the agents themselves might view these as novel (innovative) decisions; however, we show in strong contrast that a large percentage of these actions and behaviors can be predicted – and thus mimicked – in the absence of these individuals. In particular, we show that 71% of mutual fund managers' trade directions can be predicted in the absence of the agent making a single trade. For some managers, this increases to nearly all of their trades in a given quarter. Further, we find that manager behavior is more predictable and replicable for managers who have a longer history of trading and are in less competitive categories. The larger the ownership stake of the manager in the fund, the less predictable their behavior. Lastly, we show strong performance implications: less predictable managers strongly outperform their peers, while the most predictable managers significantly underperform. Even within each manager's portfolio, those stock positions that are more difficult to predict strongly outperform those that are easier to predict. Aggregating across the universe of fund managers each quarter, stocks whose position changes are least predictable additionally significantly outperform stocks whose position changes are most predictable across the universe. Our framework allows researchers to delineate and classify the portion of financial agents' action sets which are predictable from those which are novel responses to stimuli – open to being evaluated for value creation or destruction.

Keywords: Predictable portfolio choice, financial mimicking behavior, AI, machine learning.

JEL Classifications: C45, C53, C55, C82, G11, G23

[*]We would like to thank Suman Banerjee, Stefano Bonini, Brigham Brau, William Cong, Patricia Dechow, Winston Wei Dou, Cameron Ellis, Paul Gao, Jon Garfinkel, Itay Goldstein, David Hirshleifer, Gerard Hoberg, John Hund, Byoung-Hyoun Hwang, Zoran Ivkovich, Ron Kaniel, Scott Kominers, Seokwoo J. Lee, Bradford Levy, Erik Lie, Claudia Moise, Sean Myers, Anna Pavlova, Carmen Payne-Mann, Lin Peng, Tarun Ramadorai, Jonathan Reuter, Alberto Rossi, Denis Schweizer, Devin Shanthikumar, Richard Sloan, Jonathan Sokobin, Chester S. Spatt, Jess Stauth, Lorien Stice-Lawrence, Siew Hong Teoh, Jose Tribo, Wei Xiong, Jiajie Xu, Mao Ye, Yao Zeng, and Miao Ben Zhang for their valuable comments and suggestions. We also thank seminar participants at the University of Iowa, Harvard University, NBS at NTU, the UGA Fall Finance Conference, the Conference on Emerging Technologies in Accounting and Financial Economics (CETAFE) Conference, the 35th Conference on Financial Economics and Accounting (CFEA), the 2025 Society of Quantitative Analysts Boston Conference, the Behavioral Finance Working Group Conference, Behavioural Finance Working Group (BFWG), NBER Big Data, Artificial Intelligence, and Financial Economics conference, 6th RCF-ECGI Corporate Finance and Governance Conference, CUHK-RAPS-RCFS Conference on Asset Pricing and Corporate Finance, SQA Conference, and the AI in Finance Conference. All remaining errors are our own.

[†]Lauren H. Cohen, Correspondence Author, Harvard Business School and NBER, `lcohen@hbs.edu`; Yiwen Lu, The Wharton School, University of Pennsylvania, `yiwelu@wharton.upenn.edu`; Quoc H. Nguyen, DePaul University, `q.nguyen@depaul.edu`.

1 Introduction

Agents should not be compensated for aspects of their behavior and human capital that can be replicated in a low-cost way. Thus, equilibrium wages, hiring, and job tenure should be substantively affected by advances in technology and replicable skill acquisition. We explore this in the 54 trillion-dollar market of U.S. asset management.[1] In particular, we focus on the actions and behaviors of asset managers that are predictable given their past behaviors and observable fund, firm, and market conditions. Although much of the previous literature has focused on returns produced by asset managers within this space, we explore this first principle of trading behavior and relate it to realizations (including performance).

We are the first to explore the fundamental aspects of financial agent behavior that can be learned and repeated given advances in technology, and thus should now have a lower equilibrium price. In particular, we train simple, straightforward models to mimic asset manager behavior at the fund level. We find that a large percentage of the trading behavior of managers can be predicted and mimicked. In particular, 71% of portfolio managers' trades on average can be predicted in a given quarter given their past trading history. Given that this trading is a central component of the asset managers' profession, this represents an important impact of technological advancement that asset management labor supply and demand must confront.

We explore this in further detail to understand the cross-sectional and time-series of mimicking ability of the fund management universe. Namely, using data from 1990 to 2023, we find a number of empirical regularities regarding the precision with which we are able to predict asset manager trading behavior in the active equity fund manager universe. First, as mentioned above, using only basic and streamlined AI models, we are able to predict 71% of the (buy, sell, and hold) decisions of active equity fund managers each quarter. This predictability is relatively stable across the sample period, and significantly larger than that of a baseline max frequency

[1]Office of Financial Research, *Annual Report 2024* (Washington, DC: U.S. Department of the Treasury, 2024).

(also termed a Majority Class or Zero Rule Classifier) model ($p < 0.0001$). Further, when we explore this predictability of fund manager behavior across styles, we find that while large and stable across categories, it peaks for Mid-cap Blend funds at 75% predictability, on average.

Turning to fund-manager characteristics' relations to trading predictability, we find that a number of these characteristics is significantly related. In particular, we find that seasoned managers (longer tenures) are significantly easier to predict the trading of. Further, managers who manage many funds, along with those who manage across multiple strategies, have on average significantly higher trading predictability. This could be due, for instance, to constraints on time resulting in standardization or commonalities of strategies that are easier for the model to capture. In addition, managerial ownership in the fund was significantly related to predictability. Namely, the higher the ownership stake of managers in the fund itself, the less predictable their trading behavior.

We also explore a number of fund-specific characteristics. Similar to manager-level tenure, we find that older funds have a significantly higher predictability of trading. In contrast, competition, both at the category-level and within fund-family, appears to reduce trading behavior predictability. The more competition faced by managers, the more dynamic and unpredictable strategies they appear to implement. Further, larger fund management teams also exhibit trading behavior that is less predictable on average. With regard to fund size and trading activities themselves, we find that: larger funds, those with higher portfolio turnover, higher levels of investor inflows, and higher management fees, all exhibit trading behavior that is more variable and less predictable.

We then move on to explore the characteristics of the underlying securities themselves that might be related to the predictability of fund-manager trading behavior in them. First, we find that portfolio position rank of a security has a large and significant relation with predictability. Specifically, firms that are larger relative positions (e.g., the top 5 held positions in a managers' portfolio) exhibit more dynamic variation in trading, and are less predictable relative to lower ranked positions. Further, in terms of firm-level characteristics: value firms, firms with higher free cash flow, high payout ratios, and low asset turnover ratios all are easier to predict the trading of. This could be consistent with those firms that pursue more stable and easy-to-predict firm-level operating strategies. In contrast, growth, high R&D, and high earnings-surprise firms all have trading behavior surrounding them that is more difficult to predict.

Lastly, we examine the potential performance implications of a fund-manager having more (or less) predictable quarterly trading behavior. We find evidence of a strong and significant relationship. Namely, peer outperformance is nearly monotonic across trading predictability. Managers whose trading is the least predictable significantly outperform their peers, while those whose trading is most predictable significantly underperform their peers. This performance gap appears in the first quarter following trading, and continues to accrue for four quarters following trading, resulting in a significant cumulative performance gap of 0.79% ($t = 3.05$), unaffected by controls for known risk determinants.

This outperformance is concentrated in the securities that are least predictable, both within- and across-fund managers. When aggregating across managers, stocks in which the behavior of fund managers is least predictable strongly outperform stocks in which the behavior of fund managers are most predictable. The long-short (Q1–Q5) difference between the two generate an average annualized return of 424 basis points ($t = 5.74$) in the next quarter. Moreover, across all quintile portfolios, the returns are monotonically decreasing in the predictability of fund manager behavior in the security. Again, adjusting these returns for known risk-determinants has nearly no effect on these spreads, resulting in alphas of essentially the same magnitude and significance. Moreover, these returns are large and significant throughout the entire sample period, and if anything are larger in point-estimate in the most recent period through present-day.

Our paper relates to several emerging literatures. First, this paper contributes to the emerging literature that applies artificial intelligence and machine learning methods to core questions in finance. Cohen et al. (2020) use natural language processing to quantify similarities in corporate filings and show that these measures have strong implications for firms' future returns. Gu et al. (2020) demonstrate that machine learning methods improve predictive accuracy in estimating risk premia by flexibly capturing nonlinearities and high-dimensional interactions. Related work applies machine learning to enhance return prediction and portfolio performance (Freyberger et al. 2020; Kelly et al. 2019; Gu et al. 2021; Chen et al. 2024; Jensen et al. 2023; De Haan et al. 2025). Clayton and Coppola (2025) use a graph-based deep learning framework applied to large-scale portfolio holdings data, showing that structure-aware models can substantially improve predictions of intermediary trading behavior and systemic risk, particularly during stress episodes.

Our paper is also connected to studies that predict fund performance by exploiting fund and stock characteristics (Kaniel et al. 2023; Wu et al. 2021; DeMiguel et al. 2023). In contrast to this literature, we focus on predicting funds' trading directions and, more importantly, on evaluating whether an AI agent can systematically mimic the behavior of real-world asset managers.

This paper relates to two strands of research in portfolio management. The first examines whether mutual fund managers possess stock-picking ability, with mixed evidence: some studies find that active managers fail to outperform passive benchmarks once risk and costs are accounted for (Jensen 1968; Malkiel 1995; Carhart 1997), while others show that managers' trades or stock selections generate positive abnormal returns when evaluated using holdings-based or characteristics-based measures (Grinblatt and Titman 1989, 1993; Daniel et al. 1997; Wermers 1997). More recent work develops sharper tests of managerial skill, showing that managers undertake unobserved actions such as liquidity management that are not captured by standard measures (Kacperczyk et al. 2008), that the degree of active share predicts future performance (Cremers and Petajisto 2009), and that skilled managers create significant value when measured in dollar terms even if percentage-based alphas are small (Berk and Van Binsbergen 2015). Related studies emphasize determinants of fund performance more broadly, finding that funds focusing on specific industries outperform more diversified peers (Kacperczyk et al. 2007) and that flow-driven incentives shape the extent to which managers generate alpha (Guercio and Reuter 2014).

A newer strand applies machine learning to portfolio management by capturing the latent structure of securities. For example, Gabaix et al. (2024) estimate asset embeddings from institutional holdings using deep sequence models, predicting the placement of assets within an investor's ordered portfolio. Gabaix et al. (2025) apply firm embeddings derived from U.S. corporate bond holdings to extract high-dimensional representations of information used in portfolio choice, showing that these embeddings explain both the variation and the volatility of credit spreads. Our contribution bridges this literature by showing that machine learning can systematically mimic aspects of active investing, shedding light on the fundamental value added by active managers.

Finally, we also contribute to the burgeoning literature in labor economics that explores what tasks are most susceptible to innovation, and the resultant relative price and value changes across the labor force. Seminal work by Autor et al. (2003) shows that computerization substitutes for routine tasks while complementing non-routine analytic and interactive activities. Subsequent research documents the effects of technology for labor markets, including wage and employment polarization (Goos and Manning 2007; Autor and Dorn 2013; Michaels et al. 2014), aggregate employment (Autor et al. 2015; Acemoglu and Restrepo 2020), and firms' evolving demand for work skills (Bartel et al. 2007). More recent studies focus on the adoption of generative AI, showing how these tools reshape productivity and task allocation in specific occupations (Peng et al. 2023; Noy and Zhang 2023; Choi and Schwarcz 2024; Dell'Acqua et al. 2023; Walther and Dutordoir 2025). Brynjolfsson et al. (2025) find that generative AI raises productivity while compressing performance gaps across workers. Our paper extends this task-based perspective to finance by focusing on financial agents, providing empirical evidence on the fund and stock characteristics that are most susceptible to prediction and, hence, most exposed to displacement by AI.

The remainder of the paper is organized as follows. Section 2 outlines a theoretical framework for the asset management setting both pre- and post-AI introduction. Section 3 provides an overview of the AI-ML model we use, along with the data construction, variables, and empirical design. Section 4 provides summary statistics of the fund-, manager-, and firm-level data used. Section 5 examines the main predictability of fund-manager trading behavior, along with its dynamics across time and fund-categories. Section 6 explores the variation in trading behavior predictability across fund-, manager-, and firm-level characteristics, while Section 7 examines the fund performance of more and less predictable managers. Section 8 concludes.

2 Theoretical Framework

2.1 Model Setting

Combining Spence (1981) and Autor et al. (2003), we model mutual fund managers performing two types of tasks: routine tasks R and non-routine tasks N. Routine tasks are those predictable from our AI model. Non-routine tasks are those unpredictable by AI model.

Fund alpha is produced using a CES aggregator:

$$\alpha = A\left[aR^{\rho} + (1-a)N^{\rho}\right]^{1/\rho},\ a \in (0,1),\ 0 < \rho < 1, \tag{1}$$

where a is the weight on routine tasks , ρ determines the elasticity of substitution $\sigma = \frac{1}{1-\rho}$, and A is total factor productivity. Managers choose non-routine effort $e \geq 0$, which maps to non-routine input as

$$N = \theta_t^N e, \tag{2}$$

where $t \in \{L, H\}$ indexes low- and high-skill managers, and θ_t^N is type-dependent non-routine productivity. The cost of non-routine effort is

$$C_t(e) = \frac{e}{z_t}, \qquad z_H > z_L, \tag{3}$$

so high-type managers can produce non-routine output more efficiently. Routine tasks are scalable and relatively inexpensive, but entail a small convex cost that limits total activity. Markets are competitive and wages are determined by investor beliefs about expected alpha conditional on observable information.

2.2 Information Structure

Pre-AI Regime. Before AI-based decomposition, investors observe only total trading activity

$$T = R + N, \tag{4}$$

and cannot distinguish between routine and non-routine tasks. Compensation is therefore based on expected alpha conditional on total activity,

$$w(T) = \mathbb{E}[\alpha \mid T].$$

Because routine and non-routine tasks enter the observable signal additively, they are perfect substitutes from a signaling perspective.

Post-AI Regime. After the introduction of an AI model trained on historical trading data, routine trading becomes predictable to investors. Let

$$\hat{R} = \hat{r}(X_{t-1}), \tag{5}$$

denote predicted routine trading based on past information. Investors therefore observe the residual,

$$\tilde{N} = T - \hat{R} \ = \ \theta_t^N e, \tag{6}$$

which isolates the non-routine component of trading activity. As a result, routine tasks no longer affect investor beliefs or compensation.

Once AI removes the signaling value of routine activity, we treat routine input R as fixed at a baseline level. This captures the idea that routine tasks are performed mechanically and do not convey information about managerial skill in the post-AI regime. Fix R and define

$$C_0 = aR^\rho, \qquad k_t = (1-a)(\theta_t^N)^\rho.$$

Alpha can then be written as

$$\alpha_t(e) = A \left[C_0 + k_t e^{\rho}\right]^{1/\rho}. \tag{7}$$

Manager t chooses non-routine effort e to maximize

$$U_t = \alpha_t(e) - \frac{e}{z_t}. \tag{8}$$

2.3 Equilibrium Before and After AI

In the pre-AI regime, managers choose routine activity and non-routine effort subject to their respective costs, while compensation depends only on total activity.

Proposition 1 (Pooling Equilibrium Pre-AI). *In the pre-AI regime, equilibrium features pooling across types. Both high- and low-skill managers choose the same total activity level T^*, achieved primarily through routine tasks, while non-routine effort is low and does not separate types.*

Proof. Because investors observe only total activity, routine and non-routine tasks are equivalent as signals of skill. Routine tasks are cheaper to scale at the margin, allowing low-skill managers to mimic high-skill managers by expanding routine activity. Anticipating such mimicry, high-skill managers do not find it profitable to separate through costly non-routine effort. The equilibrium therefore features excessive routine activity and under-provision of non-routine effort. □

With AI-based decomposition, investors observe non-routine output $\tilde{N} = \theta_t^N e$, and routine activity no longer affects compensation. Each manager chooses e to solve

$$\max_{e \geq 0} \; A \left[C_0 + k_t e^{\rho}\right]^{1/\rho} - \frac{e}{z_t}.$$

The first-order condition is

$$A k_t e^{\rho-1} \left[C_0 + k_t e^{\rho}\right]^{\frac{1-\rho}{\rho}} = \frac{1}{z_t}. \tag{9}$$

Let $X = C_0 + k_t e^\rho$. Solving (9) yields

$$X_t^* = \frac{C_0}{1 - \left(\frac{1}{A k_t^{1/\rho} z_t}\right)^{\frac{\rho}{\rho-1}}},$$

and optimal effort is

$$e_t^* = \left[\frac{X_t^* - C_0}{k_t}\right]^{1/\rho}. \tag{10}$$

Proposition 2 (Separating Equilibrium Post-AI)**.** *After AI-based decomposition, equilibrium is separating, with*

$$e_H^* > e_L^*.$$

Proof. From (10), e_t^* is increasing in both $k_t = (1-a)(\theta_t^N)^\rho$ and z_t. Because the high type satisfies $\theta_H^N > \theta_L^N$ and $z_H > z_L$, we obtain $e_H^* > e_L^*$.

High-skill managers exert greater non-routine effort because they are both more productive and face lower marginal costs. Low-skill managers cannot profitably mimic this behavior, and non-routine output fully reveals type. □

In equilibrium, wages equal realized alpha,

$$w_t^* = A(X_t^*)^{1/\rho}, \text{ implying } w_H^* > w_L^*.$$

The model has several implications. Managers with higher skill exert more non-routine effort, making their trading behavior less predictable by AI models and generating higher alpha. AI sharpens skill revelation by stripping routine activity of its signaling value: before AI, low-skill managers pool with high-skill managers through routine trading, whereas after AI, compensation depends on costly non-routine output, leading to separation. This mechanism increases cross-sectional dispersion in performance and compensation. Finally, AI induces a reallocation of managerial effort away from routine execution and toward judgment-intensive, non-routine activities such as interpretation, strategic analysis, and complex decision-making.

In Appendix B, we examine fund level characteristics, manager level characteristics and competitive dynamics, along with their predictions for non-routine tasks. Further, we explore extensions with heterogenous investor types along with potential welfare dynamics surrounding the introduction of AI.

3 AI-ML Model and Data

3.1 Data and Sample Construction

Our analysis draws on a combination of fund- and security-level datasets assembled from multiple sources. We begin by constructing a panel of mutual fund equity holdings from Morningstar Direct, which provides comprehensive fund-security-level holdings (shares) and market value collected from SEC N-30D, N-CSR, N-Q, N-PORT filings, and fund prospectuses. These holdings are reported at either a quarterly or semiannual frequency, depending on the regulatory filing behavior of each fund. We restrict attention to actively managed U.S. equity mutual funds, excluding index, sector-specific, balanced, and international funds. The sample period spans 1990 to 2023.

We collect fund characteristics from both Morningstar and the CRSP Mutual Fund Database. The Morningstar data include metadata on fund inception dates, investment category, and manager information. We use the share-level characteristic from the oldest share of each fund. However, as some fund characteristics in Morningstar are only available for the most recent date, we complement it with CRSP, which provides monthly data on total net assets (TNA), fund returns, expense ratios, turnover, income yield, and other characteristics for each share class. We merge the CRSP and Morningstar datasets via a two-stage procedure: first on ticker, and then on CUSIP for any remaining unmatched cases. All funds are retained regardless of survivorship status, and we define each fund's active window based on its first and last appearance in the combined CRSP-Morningstar panel.

We augment the fund holdings panel with security-level and macroeconomic data to characterize the portfolios more comprehensively. Monthly security-level characteristics are obtained from the JKP Global Factor Database (Jensen et al., 2023), which provides a rich and standardized set of cross-sectional predictors widely used in empirical asset pricing research. To capture the macroeconomic environment faced by fund managers, we incorporate a set of aggregate variables from the Federal Reserve's FRED database. We include real and nominal short-term interest rates, the 10-year Treasury yield, the slope of the yield curve, and the BAA–Treasury credit spread, along with other macroeconomic indicators.

We begin by putting every fund–security holding on a monthly calendar, ensuring that each security–fund pair has one observation per month. If a month is missing for a fund–security pair, we carry forward its last known values to keep the time series continuous. We then restrict our fund–security holding data to quarter-ends (Mar/Jun/Sep/Dec). Funds must span at least seven calendar years and hold at least 10 securities per quarter; otherwise they are excluded. From that monthly grid, we keep just March, June, September, and December. These are the quarter-end snapshots we model. If a security briefly goes missing and then reappears, we allow a "recent non-zero" fill for up to six months.

For every quarter, we rank the fund's holdings by market value and assign each one an integer id: $id = 1$ is the largest holding, $id = 2$ the second-largest, and so on. We then lay the data onto a template with one row per quarter and one column slot per rank. If a fund holds, for example, 75 securities this quarter but the template expects 100 slots, ids 76–100 are empty on purpose–referred to as padding. This gives us a consistent width across time. Our final dataset includes 1,706 unique mutual funds and 5,434,702 fund–security–quarter observations over the sample period.

3.2 Features and Target Variable Construction

The prediction task is defined at the fund–security–quarter level. For each fund i and security position at time t, the target variable captures the direction of the subsequent quarter's share (sh) change. Specifically, we compute

$$\Delta sh_{i,t} = \frac{sh_{i,t+1} - sh_{i,t}}{sh_{i,t} + 1}, \qquad Y_{i,t} = \begin{cases} -1 & \Delta sh_{i,t} \leq -0.01, \\ 0 & |\Delta sh_{i,t}| < 0.01, \\ +1 & \Delta sh_{i,t} \geq 0.01. \end{cases}$$

Here, -1 denotes trim/sell, 0 denotes no material change, and $+1$ denotes add/buy. A $\pm 1\%$ band around zero counts as "no change." This is therefore a three-class classification at the fund–security level.

We construct a rich set of features from the merged dataset. At the fund level, we include lagged signals such as prior-quarter and multi-quarter returns, net cash flows aggregated at the quarterly horizon, and lagged characteristics including fund size as well as measures of value and momentum based on holdings or reported style metrics. At the security level, we incorporate lagged firm characteristics, such as logarithmic market capitalization, book-to-market ratios, and past return momentum, along with exposures to common risk factors. Factor predictors are drawn from a curated set that includes market, size, value, profitability, investment, and short-term reversal, merged at the permno–month level and aligned with quarterly observations. These are included in Appendix A.

Peer behaviors can affect the fund manager's decision. To capture peer effects, we calculate activity rates within the fund's Morningstar style category. These include the share-increase, share-decrease, and no-change frequencies of peer funds, along with their lags, thereby summarizing the contemporaneous behavior of managers with similar mandates. The macroeconomic backdrop is incorporated through quarterly lagged measures of the level and slope of the yield curve and of credit conditions, including term spreads (such as five-year minus one-year and ten-year minus three-month), default spreads (e.g., BAA–Treasury), and related short- and long-term yields, both nominal and real. These variables proxy for the broader financing and risk environment faced by portfolio managers.

Finally, we incorporate the within-fund context of each position. These features include the current market value of the position and lagged share counts tracked up to six quarters, as well as the dynamics of portfolio weights, and the size rank of the security within the fund each quarter. To ensure the panel structure remains rectangular, we also include an indicator for padded observations, which flags rows that arise from balanced-panel construction rather than actual reported holdings. All features are lagged so that they are observable at the time predictions are made.

3.3 Model, Empirical Design, and Evaluation

The machine learning model we use for the study is Long Short-Term Memory (LSTM), which was first introduced by Hochreiter and Schmidhuber (1997). LSTM networks are a type of Recurrent Neural Network (RNN) designed to capture long-term dependencies in sequential data. Unlike standard RNNs, which suffer from vanishing or exploding gradients when modeling long sequences, LSTMs address this problem through a gating mechanism. This architecture is particularly appropriate for our application, as fund trading decisions may exhibit persistent patterns over time, and effective prediction requires the model to retain and exploit these long-term dependencies.

LSTMs are composed of a set of recurrently connected subnets, called memory cells, containing three multiplicative units, the forget gate, input gate, and output gate. The basic architecture of a memory block is shown in Figure 1, where C_{t-1} and C_t represent the previous and current state of the memory cell. h_{t-1} and h_t are current and previous hidden states, and X_t is the current signal. The forget gate decides which historical information to discard from the previous cell state. It receives the previous hidden state h_{t-1} and current signal X_t and processes them using a sigmoid activation function which generates a forget coefficient $f(t) \in [0,1]$. If $f(t)$ is close to 0, the past information is largely forgotten.

$$f(t) = \varphi\left(w_{fx} \cdot X_t + w_{fh} \cdot h(t-1) + b_f\right), \tag{11}$$

where $\varphi(\cdot)$ is the sigmoid function, w_{fx}, w_{fh} are the weight matrices, and b_f is the bias term. The input gate $i(t)$ controls how much new information to add, which includes the following components:

$$i(t) = \varphi\left(w_{ix} \cdot X_t + w_{ih} \cdot h(t-1) + b_i\right), \tag{12}$$

$$\widetilde{C}(t) = \tanh\big(w_{cx} \cdot X_t + w_{ch} \cdot h(t-1) + b_c\big). \tag{13}$$

The current state of the memory cell is determined by the previous state and the newly computed candidate state $\widetilde{C}(t)$:

$$C(t) = f(t) \odot C(t-1) + i(t) \odot \widetilde{C}(t), \tag{14}$$

where $\tanh(\cdot)$ represents the hyperbolic tangent activation function. The output gate $o(t)$ predicts the final outcome after quantifying the current information from the cell, which is calculated using:

$$o(t) = \varphi\left(w_{ox} \cdot X_t + w_{oh} \cdot h(t-1) + b_o\right). \tag{15}$$

In addition, the hidden state $h(t)$ is updated using:

$$h(t) = o(t) \odot \tanh\big(C(t)\big). \tag{16}$$

We construct the empirical design around rolling panels constructed from fund holdings. For each eligible fund, we form overlapping windows of 28 quarters. Within a given window, we retain up to N securities with N is determined by the realized cross-sectional size of the fund's holdings in that window. From each window, we generate fixed-length sequences of eight consecutive quarters. Each sample is initially represented as a three-dimensional tensor $X \in \mathbb{R}^{T \times N \times F}$, where the sequence length is $T = 8$, N is the number of distinct securities retained in the window, and F is the number of security-level features. At each quarter t, securities are ordered in descending order of market value within the fund. To feed the data into a standard single-layer LSTM, we vectorize the cross-section at each time step by concatenating the ordered $N \times F$ feature matrix into a single NF-dimensional vector. The resulting sequence $(x_1, \ldots, x_T) \in \mathbb{R}^{T \times NF}$ is used as the LSTM input. This construction preserves the economic cross-section of fund holdings while ensuring that the LSTM input dimension is fixed and comparable across samples.

Each 28-quarter window is then split chronologically into training and test subsamples, with the first 20 quarters reserved for training and the final eight quarters for testing (see Figure 2). No random shuffling is employed, thereby preserving the time-series ordering of the data. We explicitly log the implied train-test cutoff quarter for reproducibility. All inputs are implicitly standardized by construction: the feature design relies on relative rankings within fixed-width panels. While explicit normalization can be layered on top, it was not required for our baseline feature set.

We estimate a single-layer LSTM network on these sequences. The hidden dimension scales with the size of the realized cross-section to balance model flexibility and overfitting risk. To mitigate overfitting, we apply both dropout and recurrent dropout, each with a rate of 0.25. The output of the recurrent layer is passed through a dense transformation, reshaped and mapped to a probability grid of dimension $(N \times 3)$ via a softmax activation. This design yields, for each security in the retained cross-section, a probability distribution across three mutually exclusive states.

We include two masks that encode feasibility. First, we implement a time-step mask. At any quarter in which the cross-sectional panel is purely synthetic, that is, all securities in the retained output set are padding indicators ($sh\_past = 0$ for every identifier), we mask the entire time step from recurrence. This prevents the network from updating parameters based on artificial or economically irrelevant observations.

Second, we impose an output feasibility mask at the security level. For a given identifier, let the predicted softmax vector be $\mathbf{p} = (p_{-1}, p_0, p_{+1})$, corresponding to sell, hold, and buy states. If the security lacks continuous presence over the full eight-quarter input horizon (i.e., if any period exhibits $sh\_past = 0$), we enforce feasibility by zeroing out the sell probability p_{-1} and renormalizing the distribution over the remaining states $\{0, +1\}$. A looser variant of this mask requires only that the security be held in the final quarter of the input horizon, thereby accommodating entry within the window while still prohibiting spurious sell signals.

Our estimation objective is a weighted categorical cross-entropy loss:

$$\mathcal{L} = -\sum_{c \in \{-1,0,+1\}} w_c \, y_c \, \log p_c,$$

where y_c is the true class indicator, p_c is the predicted probability, and w_c denotes the class weight. By default, $w_c = 1$ for all classes. To enhance robustness to class imbalance and to emphasize hard-to-classify observations, we also implement an optional focal variant in which the loss is multiplied by $(1 - p_c)^\gamma$, with $\gamma = 1$. Optimization proceeds using the Adam algorithm with early stopping based on validation loss. The training horizon is capped at 50 epochs, ensuring convergence while preventing overfitting.

To benchmark model performance, we construct both a naive baseline and a set of model-based precision measures. We first define a naive precision benchmark. For each training window, the naive classifier predicts the max (frequency) class observed across all securities and time steps. This baseline provides a conservative reference point, ensuring that measured gains are not artifacts of class imbalance.

We then compute model precisions for both the naive and LSTM-based predictors. We adjust our precision metrics by restricting attention to feasible cases. Specifically, we restrict attention to test outputs where the feasibility mask permits a sell decision, i.e., securities with continuous presence over the input horizon.

4 Summary Statistics

We describe the data and sample in Section 3. In Table I, we show the summary statistics at the fund level. From Panel A of Table I, managers in our sample have an average tenure of over 13 years (Manager Tenure (Mean)) as we require a minimum of 7 years to train and test the managers' behavior. The Manager Tenure (Max) variable takes the maximum of this tenure for each team-managed fund, and thus is slightly longer at an average of roughly 15 years. Managers on average manage 2.83 funds across 2 different investment style categories.

From Panel B of Table I, the average fund age is roughly 14 years, and has 3.6 managers per fund. The average fund faces 75 fund competitors within the category, and has an average of 4.79 funds in the fund management company.

Table II then shows additional fund-specific characteristics of the sample. The average naive prediction precision of a fund manager's trading behavior in the sample is 0.52, meaning that we can predict (buy, sell, hold) decisions at the security-quarter level at 52% accuracy. The naive prediction model simply looks at the most common behavioral action of the manager (max frequency, also termed the Majority Class or Zero Rule Classifier) during the training period. In contrast, our AI training model described in Section 3 achieves an average prediction precision of 71%, a substantial economic and statistically large increase of 36.5% ($p < 0.0001$). The average fund size is roughly 700 million dollars (maximum of over 38 billion dollars), with 5.29 million dollars of average flows each quarter. Moreover, 95% are listed as open to new investment each quarter.

In Table III, we examine the average position characteristics at the firm-quarter-fund level. From Panel A, for instance, the average position size is roughly \$6 million, and we examine the top 100 positions held by these active equity fund managers. From Panel B, the average ME (firm size) over our sample period of held firms is roughly \$30 billion, with a median size of \$6 billion.

5 Main Results

5.1 Full Sample Results

The central aim of the study is to explore the proportion – and *which* proportion – of fund manager's behavior can be predicted using the fund manager's past behaviors, and their reactions to market conditions, and changing firm characteristics. Using modern AI-ML techniques and tools, drawing systematic and repeated components of these behaviors is becoming ever-more tractable.

In Figure 3, we illustrate using the simple and straightforward model we outline in Section 3 the potential to do exactly this. In particular, the shaded red distribution uses the Max Frequency estimator – also termed the Majority Class or Zero Rule Classifier. This estimator uses the most frequent behavior by the manager in the training period (sell, hold, or buy) and applies this as the naive estimated behavior for all of her future trades in her testing period. This results in a fairly wide distribution of estimated precisions, with a mean precision of estimated behavior of 0.52, or 52% of the correctly predicted trades. In contrast, applying our simple AI model that learns from the manager's past behavior, the average prediction accuracy jumps to 71%. This difference represents a 36.5% increase in prediction accuracy, and is highly significant ($p < 0.0001$). Further, as can be seen in Figure 3, the entire distribution of prediction accuracy tightens and shifts to the right. Thus, a substantially larger percentage of the sample of fund managers can be predicted more accurately, with significantly fewer managers being predicted with less accuracy. In fact, while 44.38% of the managers fall below the 50% prediction accuracy threshold for the naive prediction method, only 11.11% fall below this 50% threshold using our simple AI model.

5.2 Cross-sectional and Time-series Dynamics of Precision Score

Figure 4 then graphs the time-series of this prediction precision over the sample period. From Panel A, graphing the prediction precision of our AI model over time, the precision is fairly stable around the 71% mean, exhibiting a slightly upward trend toward the end of the sample period, moving toward present-day. In addition, quantitative tools and factor investing became more common, which again could drive their predictability (using loadings on these factors), increasing potential predictability. Panel B then graphs this corresponding time series for the naive prediction model. From Panel B, we again see a fairly stable precision over time around the mean of 52%, without an increase toward the end of the sample period.

In Figure 5, we plot prediction precision across Morningstar fund style categories. Across all nine style groups, the prediction precision of our model is significantly higher than that of the naive prediction, and retains a similar shape. The mass of the distribution has shifted significantly rightward, toward higher accuracy levels. That said, differences do emerge: large-cap and mid-cap categories generally have higher precision than small-cap. This is true across the growth, Value, and blend categories. The single highest prediction precision is seen in the mid-cap blend category, with our model achieving a 75% prediction accuracy, on average. Figure 6 then similarly separates our fund universe into broad size categorizations - small, mid, and large - including specialized funds (e.g., sector or geographic focused funds) in the other category. Consistent with Figure 5, Figure 6 shows that large and mid-sized funds exhibit higher median prediction precision relative to small and other funds, and their distributions are more concentrated at the upper end of the precision scale. By contrast, small and especially other funds display wider dispersion and lower central precision tendency, indicating greater heterogeneity and average lower precision among these groups.

Figure 7 then plots the time series of prediction precision of fund categories over the sample period. Across time – and in particular, post-2000 – large and mid-sized fund categories achieve higher prediction precision relative to small and other funds. All groups also exhibit improvement over time, in particular post-2016, consistent with the general time pattern observed in Figure 4.

6 Dynamics of Precision Score at the Fund and Security Level

Section 6 explores the prediction performance of our baseline AI model described in Section 3. The average prediction precision of 71% of managers' trading decisions (buy, hold, and sell) was relatively stable across time and fund categories, and substantially larger than the naive prediction method. In this section, we turn to examining the manager, fund, and security characteristics associated with the highest (and lowest) predictive accuracy across and within funds.

6.1 Fund and Manager Level

We begin by examining fund and manager level characteristics. The results in Table IV provide an examination of fund manager characteristics' association with the prediction accuracy of fund-level investment behavior. A consistent finding across specifications is that managerial tenure, along with when managers are stretched across more funds, is associated with a significantly higher ability to predict subsequent trading behavior. In contrast, managerial ownership appears to work in the opposite direction, with higher managerial ownership of the fund leading to less predictive ability of their trading behavior.

The specifications in Table IV regress future prediction precision at the fund-quarter level on a host of manager level characteristics. Year and fund fixed effects are included to ensure that the results are not driven by temporal shocks or persistent fund-level heterogeneity. For instance, controlling for temporal-level shocks or variation across all funds (e.g., the financial crisis of 2008), along with potential stable idiosyncratic components of funds' universes (e.g., large, well-covered technology firms versus smaller, less well-covered oil prospecting firms). Standard errors are clustered at the quarter level.

Starting with Manager Ownership, the coefficient is negative and highly significant, indicating that funds in which managers have larger personal stakes in the form of dollar ownership have trading behavior that is more difficult to predict. One interpretation is that when managers hold significant ownership, they may be more incentive-aligned and thus take active decisions to optimize portfolio choice and trading dynamics. Turning to Manager Tenure, both the maximum tenure (measured in quarters across the longest-serving manager) and the average tenure across managers show positive and statistically significant relations. This implies that more experienced managers, whether considered individually or collectively, are associated with easier to predict trading behavior. A plausible interpretation is that long-tenured managers develop stable decision-making patterns and investment philosophies, making their trading easier to anticipate.

The results for Manager #Funds and Manager #Styles further highlight the potential importance of managerial scope. Managers who simultaneously oversee more funds or adopt multiple investment styles have trading behavior that is significantly easier to predict. Managing multiple funds or strategies (with a common constraint on time) may require a degree of standardization, and reliance on common investment processes or models, which again creates regularity in decision-making that the model can capture.

Taken together, these results suggest that fund behaviors are most predictable among seasoned managers with broad responsibilities, while in contexts where managers hold significant personal ownership stakes, this reduces predictability.

Table V analyzes how fund characteristics influence prediction precision, using an analogous empirical specification to Table IV. The broad findings provide evidence suggesting that the ability to predict fund-level managerial behavior increases with fund age, while size of the team – along with competition both within and across fund families – reduces predictive ability.

In more detail, Fund Age, measured by the number of quarters since inception, is associated with significantly higher prediction precision. This suggests that as funds mature, their strategies and trading behavior become more predictable, perhaps as they become more stable and consistent over time, making them easier for the model to capture. In contrast, both forms of competition, within category and within management company, are negatively associated with prediction precision. The negative coefficient on Within Category Competition indicates that funds facing a higher number of peers within the same investment category are harder to predict. Competitive pressure may lead managers to adjust their trading in response to rivals' strategies, causing behavior that is less regular and harder for the model to capture. The stronger negative effect of Within Management Company Competition suggests that when many funds coexist within the same parent company, prediction becomes even more challenging. Internal dynamics, or overlapping mandates, may add complexity to game-theoretic decision-making, thereby reducing the regularity of trading patterns.

Finally, the results for Number of Managers point to a clear pattern: funds managed by larger teams exhibit lower prediction precision. While multiple managers may bring diverse perspectives, this diversity can lead to less consistent and predictable aggregate decision-making. Differences in judgment, coordination challenges, or rotation of responsibilities among managers all may contribute to greater variation in trading outcomes, which weakens predictability from the model's perspective.

Lastly, Table VI examines portfolio and fee aspects of funds and their relationships with prediction precision, with analogous empirical specifications to Tables IV and V, including year and fund fixed effects. Column 1 shows a strong negative relationship between fund size and prediction precision. The coefficient on total net asset (−2.1384, significant at the 1% level) implies that larger funds are systematically more difficult to predict. As funds grow in size, they may be required to execute more complex trades and trades at different scales, both of which could add variability to their portfolio decisions. Column 2 demonstrates that funds with higher turnover ratios are less predictable. The coefficient (−0.0149, $t = -9.712$) is highly significant and indicates that each incremental increase in turnover is associated with a measurable decline in prediction precision. High turnover suggests more active trading, frequent rebalancing, and shorter investment horizons, which could generate more irregularity and heterogeneity in trading patterns.

Column 3 finds that fund flows negatively affect prediction precision. The coefficient on Fund Flow is −4.5290 (significant at 5%), suggesting that large inflows or outflows into the fund in a given quarter reduce predictability. Investor-driven capital movements force managers to accommodate cash injections or redemptions in ways that may not perfectly align with underlying strategy, resulting in less consistent trading behavior, thus causing reduced predictability. Column 4 shows that higher management fees are associated with reduced prediction precision (−0.0078, significant at the 5% level). This could indicate that managers who are more active, charge higher fees for this activity (in equilibrium) to investors. In contrast, there is no significant association between the total expense ratio of the fund and prediction precision, from Column 5. These jointly suggest that only the active management component of fees is associated with predictability, while general administrative or operational costs have a less meaningful relationship with trading regularity.

Column 6 indicates that funds open to new investors are less predictable (−0.0060, significant at the 1% level). It may be that when funds remain open, they are more exposed to unpredictable investor inflows, which may compel managers to adjust holdings in irregular ways. This contrasts with closed funds, where managers are less exposed to frequent disruptions from capital movements. The final specification in Column 7 finds that funds with sales restrictions are more predictable. The coefficient (0.0578, $t = 4.874$) suggests that restrictions – such as minimum holding periods or redemption fees - may stabilize investor allocation behavior by discouraging short-term withdrawals, again reducing large short-run fluctuations that would induce more variable trading.

Taken together, the results suggest that smaller funds with lower turnover, stable flows, limited investor access, and sales restrictions have more predictable trading patterns that can be anticipated. In contrast, larger funds, those with high turnover, heavy flows, and higher management fees, may employ more complex or variable strategies that are more difficult to forecast.

From a broader perspective, these findings suggest that fund characteristics shape not only performance and risk, as commonly studied, but also are associated with the degree to which managers' behaviors in these funds can be captured by AI and machine learning models.

6.2 Security Level

In this section, we move on to explore whether certain characteristics of underlying firms themselves are associated with being more or less predictable in mutual fund manager's trading behavior in these firms.

We begin in Table VII, exploring prediction precision of fund managers at the fund-quarter-security level holding level. All specifications include Quarter-Year FE and Firm FE, and standard errors are two-way clustered by firm and quarter, so results are robust to both cross-sectional and temporal correlation. Column 1 indicates that Portfolio Position – measured as the largest holding (1) to the smallest (100) – has a strong positive association with prediction precision. The coefficient of 0.0020 is highly significant, and suggests that securities held as smaller relative positions in the fund (and so a higher Portfolio Position) are more predictable. As these have a smaller relative contribution to fund-level return, one might expect managers to allocate fewer variable rules and adjustments to these positions, resulting in more predictable behavior among them. Column 2 then examines firm size, measured by market equity. The coefficient is negative and significant, implying that larger firms are slightly harder to predict once their securities are held in fund portfolios. This is consistent with Column 1, in that these are often larger positions held within funds, and more actively and heterogeneously traded and rebalanced.

Column 3 adds the number of distinct funds holding a given security in a given quarter. The coefficient of –0.0002 ($t = 4.659$) implies that securities held by more funds are on average less predictable. When a stock is widely held across many funds, trading activity may reflect heterogeneity (e.g., style rebalancing, liquidity needs, benchmark-tracking differences) across funds, introducing complexity that lowers prediction accuracy at the security level. By contrast, securities held by fewer funds may potentially be those that align with fund-specific strategies, which the model can capture more reliably, along with this measure likely being correlated with firm size (as larger stocks on average have increased breadth of institutional ownership). Column 4 introduces Firm Age, measured as the number of months since the firm's first appearance in CRSP. The coefficient is positive and significant, suggesting that older firms are slightly more predictable. This could reflect mature firms being larger, along with having potentially greater stability and predictability of firm-level action, compared to newer firms with more growth uncertainty, or rapid valuation shifts.

Table VIII then explores a number of additional firm-level fundamental characteristics. These include standard dimensions of, capital structure, profitability, growth, cash flows, and equity issuance, exploring which types of firms are associated with more predictable security-level trading outcomes. In Column 1, leverage has a strong positive relationship with prediction precision, suggesting that firms with higher debt-to-market equity ratios are more predictable. One possible explanation for this is that highly leveraged firms may be monitored closely, reducing uncertainty in trading behavior surrounding them. In Column 2, the coefficient on Asset Turnover (net sales relative to assets) is negative and significant, suggesting that firms that more quickly turnover assets into sales may be more actively and variably traded, and so more difficult to predict. Of the remaining firm fundamental proxies, a number are significantly associated with fund manager's predictability of their trading. Fund manager trading behavior is significantly more predictable for Value firms (Column 7), high Free Cash Flow firms (Column 6), and high Net Equity Payout firms (Column 8), as all could be more stable and so predictable in optimal portfolio composition and position. In contrast, firms with earnings surprises (Column 9), show less predictability, perhaps due to the inability of market participants to anticipate future cash flow and earnings realizations.

In sum, the results in Table VIII provide suggestive evidence that predictability is higher for stable, mature firms with stronger balance sheets and predictable payout policies. In contrast, predictability appears lower for growth, high R&D, and shock-prone firms, subject to larger earnings surprises.

Lastly, Table IX explores the relationship of past firm returns and realized return volatility with fund manager predictability of behavior in the firms. These capture return-based factors such as short-term reversals, momentum, and idiosyncratic volatility.

From Table IX, the only return-based variables that are significantly associated with fund manager predictability in behavior are Past 1 Month Returns and Idiosyncratic Volatility. From Column 1, the coefficient on Past 1 Month Returns is negative and significant (−0.0239, t = -3.089). This suggests that stocks experiencing recent short-term negative returns are somewhat easier to predict manager trading behavior around. From Column 4, stocks that have higher idiosyncratic volatility (measured as the standard deviation of residuals from a CAPM regression estimated over the prior 36 months), have a marginally significantly higher predictability in fund manager trading behavior surrounding them.

7 Implications for Fund Return Predictability

While prior sections have explored the associations of fund, manager, and firm-level characteristics with the ability to predict fund manager's trading behavior, in this section we ask whether managers whose behavior is easier or more difficult to predict exhibit differing levels of performance. To test this, we split our manager universe each quarter by their ranking on predictability of trading behavior in the testing period. We split managers into quintiles, and measure the future performance of each quintile.

The results are in Table X, where Q1 is those managers with the least prediction precision up through Q5 – managers whose trades are easiest to predict. The table reports benchmark-adjusted abnormal returns (relative to funds in the same Morningstar category) over horizons of one to four quarters. The table highlights both individual quintile performance and the spread between the highest-precision (Q5) and lowest-precision (Q1) quintiles.

Managers whose trading behavior is the most difficult to predict (Q1) consistently generate positive and significant abnormal returns over all horizons. At the one-quarter horizon, Q1 funds earn 0.14% ($t = 2.05$), which rises steadily to a cumulative abnormal return of 0.36% ($t = 2.88$) by quarter four. This indicates that funds with the least predictable trading behavior outperform their category peers, an effect that does not reverse, but strengthens over longer horizons.

In stark contrast, Q5 fund managers – those whose trading is easiest to predict – consistently underperform peer funds. Their cumulative abnormal returns are slightly negative in the short run (−0.09% at one quarter, $t = 1.52$), becoming increasingly negative and statistically significant over longer horizons (−0.22% at two quarters, −0.31% at three quarters, −0.42% at four quarters, $t = 2.36$).

Prediction precision appears to be a relevant attribute that acts across the entire fund manager universe – not only acting upon the extremes of the distribution. From Table X, returns are nearly monotonic in prediction precision at each of the four horizons. Turning to the Q1–Q5 Spread, which reports the difference in excess returns between the least predictable and most predictable funds, this difference is significant at each horizon and increasing across horizons. In particular, from 0.23% ($t = 2.13$) at one quarter to 0.79% ($t = 3.05$) at four quarters. Figure 8 illustrates these return regularities across quintiles for each of the various horizons.

To assess the robustness of these findings to different risk adjustment methodologies, Table XI reports the cumulative four-quarter excess return and alphas generated by quintile portfolios sorted on prediction precision, using multiple factor models. Consistent with the excess return analysis in Table X, we observe a strong negative relationship between trading predictability and risk-adjusted performance. Managers in the lowest-precision quintile (Q1) generate significant alphas across all specifications: 0.37% under the Fama-French three-factor model, 0.31% under the four-factor model, and 0.37% under the five-factor model, all statistically significant at conventional levels. Conversely, managers in the highest-precision quintile (Q5) consistently underperform, with alphas ranging from −0.41% to −0.48%, again significant at the 1% level. The Q1–Q5 spread ranges from 0.78% to 0.82% across factor models (all t-statistics exceed 2.5), demonstrating that the superior performance of unpredictable managers is not merely compensation for systematic risk exposure, but rather reflects superior stock-picking ability or market-timing skill that persists after controlling for standard asset pricing factors.

Table X, Table XI, and Figure 8 provide strong and consistent evidence that funds whose trading is highly predictable underperform their peers, while those that are least predictable deliver superior returns. This result is both statistically and economically meaningful, particularly at horizons of three to four quarters, where the return spreads become sizable. As potential drivers of these dynamics, predictable funds could be those that follow more mechanical strategies, that may be relatively easier for models to learn, but also potentially for markets to anticipate, learn, and quickly price, reducing their ability to generate alpha. By contrast, unpredictable funds may reflect managers who adapt dynamically or exploit informational advantages, leading to stronger subsequent performance.

7.1 Correctly vs. Incorrectly Predicted Positions - Portfolio Level

From Table II and Figure 3, we find that, on average, we are able to predict 71% of a fund manager's trades each quarter. The question then is whether the relative out-performance of unpredictable fund managers found in Table X is driven by the 29% of her portfolio that we cannot predict, on average, or the 71% that we are able to. In order to test this, for each fund-quarter observation, we split each manager's stocks into two portfolios: an "Incorrect Predictions" portfolio that contains stocks where the fund's predicted direction was wrong, and "Correct Predictions," containing stocks where the fund's predicted direction was accurate. We take the average of quarterly returns each quarter across all managers' "Correct Predictions" and "Incorrect Predictions" portfolios and report them in Table XII.

From Table XII, the incorrect positions strongly outperform correct positions. The results show that the Incorrect Predictions portfolio (holdings harder to predict) generates a mean annualized return of 7.98% compared to 7.07% for Correct Predictions (holdings easier to predict), yielding a difference of 91 basis points. The t-statistic of 12.41 indicates this difference is highly statistically significant at the 1% level. Table XII also provides robustness evidence at the holding level by examining both excess returns and factor-adjusted alphas. We decompose each fund's portfolio into two distinct components based on prediction accuracy: stocks where our model correctly predicted the fund's direction versus those where it was incorrect. For both excess returns and the three-, four-, and five-factor models, holdings where the fund's trading direction was incorrectly predicted substantially outperform those with correct predictions. Panel A presents fund-quarter-level analysis using all fund-quarter observations, where each fund-quarter contributes one observation for both correct and incorrect prediction portfolios.

Figure 9 then plots the cumulative returns to both the Correct Predictions Portfolio and Incorrect Predictions Portfolio over the entire sample period. Panel A shows this separately for both portfolios, while Panel B graphs the cumulative (L-S) difference between the two over time. From Figure 9, the Correct Predictions Portfolio outperforms the Incorrect Predictions Portfolio during nearly the entire sample period, with the cumulative difference rising relatively stably throughout the period.

7.2 Correctly vs. Incorrectly Predicted Positions - Security Level

Table XIII then explores this question in a slightly different specification. Namely, it tests whether the signals across all managers can be aggregated to arrive at a security-time level measure with implications for return realizations. In particular, Table XIII reports the performance of quintile stock portfolios formed on stocks' percentage of correct directional prediction by fund managers that held that stock. Each quarter, stocks are ranked by the average prediction accuracy of all funds holding that stock and sorted into five equal-weighted portfolios. Quintile 1 (Q1) captures stocks where, on average, the behavior of fund managers is least predictable, while Quintile 5 (Q5) contains stocks where, on average, the behavior of fund managers is most predictable. From Table XIII, stocks in which the behavior of fund managers are least predictable (Q1) strongly outperform stocks in which the behavior of fund managers are most predictable (Q5) in the coming quarter. The long-short (Q1–Q5) difference between the two generate an average return of 424 basis points (t = 5.74) annualized over the next year. Moreover, across all quintile portfolios, the returns are monotonically decreasing across the predictability from Quintile 1 through Quintile 5.

For robustness, we also examine excess returns and factor-adjusted alphas across the quintile portfolios. Table XIII also reports performance metrics across multiple return specifications, including excess returns and three-, four-, and five-factor Fama-French alphas. The results demonstrate that the outperformance of stocks with least predictable manager behavior (Q1) relative to most predictable behavior (Q5) is robust across all specifications. The long-short (Q1–Q5) portfolio generates substantial and statistically significant alphas, for instance with a Fama-French five-factor annualized alpha of 386 basis points ($t = 4.96$).

Figure 10 shows the cumulative returns of each quintile over the entire sample period. Panel A shows that the cumulative difference between portfolios is monotonic, with the gap in cumulative returns across the portfolios increasing over time. Panel B then plots the cumulative long-short difference (Q1–Q5) between the two levels of predictability over time. Again, this cumulative outperformance appears across the entire sample period, with the difference rising stably throughout the sample period. Figure 10 also demonstrates the robustness and persistence of our main findings across the entire sample period and multiple risk-adjustment specifications. The figure displays cumulative performance trajectories for quintile portfolios formed on fund manager's prediction accuracy, with results presented across four return specifications: excess returns, FF3 alpha, FF4 alpha, and FF5 alpha. Across all four specifications, stocks where fund managers are least accurate in predicting directional returns (Q1) consistently and substantially outperform stocks for which managers are most accurate (Q5), with this performance ranking remaining remarkably stable throughout our sample period.

7.3 Fama-MacBeth Regression Framework with Active Shares Control

To further examine the robustness of our findings using an alternative empirical framework, we employ Fama-MacBeth regressions to assess the relationship between prediction precision and subsequent abnormal returns while controlling for additional fund characteristics. Specifically, we include Active Share as a control variable, which Cremers and Petajisto (2009) define as the fraction of a fund's portfolio level holdings that differ from its benchmark index. This metric captures the degree to which a manager's portfolio deviates from its benchmark and serves as a control as it may reflect managerial skill or conviction independent of predictability.

Table XIV presents the results of these cross-sectional regressions, where the dependent variable is cumulative abnormal returns measured from quarter 0 through 4. Standard errors are adjusted following the Newey-West procedure to account for potential autocorrelation and heteroskedasticity in the time series of cross-sectional regression estimates. The inclusion of Active Share as a control allows us to isolate the effect of Prediction Precision regarding active trading on abnormal returns from effects driven by portfolio-level deviations from the benchmark.

The coefficient on Prediction Precision remains economically and statistically significant across all specifications, even after the inclusion of Active Share. In Model (1), the coefficient on Prediction Precision is −0.8554 and statistically significant at the 5% level. Upon adding Active Share as a control in Model (2), the coefficient on Prediction Precision increases slightly in magnitude to −1.0074 ($t = -2.38, p < 0.05$), while Active Share itself enters with a positive and significant coefficient of 0.0041 ($t = 5.13, p < 0.01$). Critically, the magnitude and significance of the Prediction Precision regarding active trading decisions remains largely unchanged, demonstrating that our findings are not driven by differences in portfolio-level deviations from benchmarks. Rather, the predictability of trading behavior itself emerges as a distinct and robust predictor of abnormal returns, orthogonal to the level of Active Share.

8 Conclusion

We construct a new measure and a benchmark to assess manager performance, namely the simple measure of behaviors that can (and cannot) be easily predicted and replicated. We use a straightforward AI model to learn and replicate the trading behavior of managers. Using simple measures of managers' past behaviors, along with fund-, manager-, and firm-level characteristics, we train the model to predict future manager trading behavior. On average, our model is able to predict 71% of active equity managers' trades (buy, hold, and sell). This predictability is relatively stable across the entire sample period and across fund manager styles. Moreover, this predictability of behavior is strongly associated with future performance: managers whose behavior is most predictable significantly underperform their peers, while those who are least predictable exhibit strong future outperformance. Further, this outperformance is concentrated in the securities that are least predictable, both within and across fund managers.

Stepping back, the results have broad implications. In theory, compensation to agents should not be tied to dimensions of behavior or human capital that can be replicated at low cost. Consequently, equilibrium outcomes – such as wage levels, hiring practices, and job tenure – may be materially influenced. However, financial agents also have a role and best-response in this dynamic game. For example, faced with more powerful detection and replication tools, agents could take measures to obfuscate and change behavior to reduce the ability of these tools to extract predictable behaviors. In equilibrium, the new labor wage- and duration-contract, and its dynamics, must take into account these burgeoning tools and implications.

Lastly, while we have focused on the economic agents of active mutual fund managers in this paper, future research should address equivalent implications across the agent spectrum. These range from other financial agents such as sell-side analysts, to other firm-level decision-makers such as CEOs, who synthesize both public and private information held within and across these organizations and markets.

References

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. Journal of Political Economy 128(6), 2188–2244.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. American Economic Review 103(5), 1553–1597.

Autor, D. H., D. Dorn, and G. H. Hanson (2015). Untangling trade and technology: Evidence from local labour markets. The Economic Journal 125(584), 621–646.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics 118(4), 1279–1333.

Bartel, A., C. Ichniowski, and K. Shaw (2007). How does information technology affect productivity? plant-level comparisons of product innovation, process improvement, and worker skills. The Quarterly Journal of Economics 122(4), 1721–1758.

Berk, J. B. and J. H. Van Binsbergen (2015). Measuring skill in the mutual fund industry. Journal of Financial Economics 118(1), 1–20.

Brynjolfsson, E., D. Li, and L. Raymond (2025). Generative ai at work. The Quarterly Journal of Economics 140(2), 889–942.

Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of Finance 52(1), 57–82.

Chen, L., M. Pelger, and J. Zhu (2024). Deep learning in asset pricing. Management Science 70(2), 714–750.

Choi, J. H. and D. Schwarcz (2024). Ai assistance in legal analysis: An empirical study. Journal of Legal Education 73, 384.

Clayton, C. and A. Coppola (2025). Financial regulation and ai: A faustian bargain? arXiv preprint arXiv:2507.18747.

Cohen, L., C. Malloy, and Q. Nguyen (2020). Lazy prices. The Journal of Finance 75(3), 1371–1415.

Cremers, K. M. and A. Petajisto (2009). How active is your fund manager? a new measure that predicts performance. The Review of Financial Studies 22(9), 3329–3365.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997). Measuring mutual fund performance with characteristic-based benchmarks. The Journal of Finance 52(3), 1035–1058.

De Haan, E., C. Lee, M. Liu, and S. Noh (2025). The shadow price of "public" information. Working Paper.

Dell'Acqua, F., E. McFowland III, E. R. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, and K. R. Lakhani (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper (24-013).

DeMiguel, V., J. Gil-Bazo, F. J. Nogales, and A. A. Santos (2023). Machine learning and fund characteristics help to select mutual funds with positive alpha. Journal of Financial Economics 150(3), 103737.

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. The Review of Financial Studies 33(5), 2326–2377.

Gabaix, X., R. S. Koijen, R. Richmond, and M. Yogo (2024). Asset embeddings. Available at SSRN.

Gabaix, X., R. S. Koijen, R. Richmond, and M. Yogo (2025). Upgrading credit pricing and risk assessment through embeddings. Working Paper.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. The Review of Economics and Statistics 89(1), 118–133.

Grinblatt, M. and S. Titman (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. Journal of Business, 393–416.

Grinblatt, M. and S. Titman (1993). Performance measurement without benchmarks: An examination of mutual fund returns. Journal of Business, 47–68.

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

Gu, S., B. Kelly, and D. Xiu (2021). Autoencoder asset pricing models. Journal of Econometrics 222(1), 429–450.

Guercio, D. D. and J. Reuter (2014). Mutual fund performance and the incentive to generate alpha. The Journal of Finance 69(4), 1673–1704.

Hochreiter, S. and J. Schmidhuber (1997). Long short-term memory. Neural Computation 9(8), 1735–1780.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. The Journal of Finance 23(2), 389–416.

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023). Is there a replication crisis in finance? The Journal of Finance 78(5), 2465–2518.

Kacperczyk, M., C. Sialm, and L. Zheng (2008). Unobserved actions of mutual funds. The Review of Financial Studies 21(6), 2379–2416.

Kacperczyk, M. T., C. Sialm, and L. Zheng (2007). Industry concentration and mutual fund performance. Journal of Investment Management 5(1), 50–64.

Kaniel, R., Z. Lin, M. Pelger, and S. Van Nieuwerburgh (2023). Machine-learning the skill of mutual fund managers. Journal of Financial Economics 150(1), 94–138.

Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134(3), 501–524.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. The Journal of Finance 50(2), 549–572.

Michaels, G., A. Natraj, and J. Van Reenen (2014). Has ict polarized skill demand? evidence from eleven countries over twenty-five years. Review of Economics and Statistics 96(1), 60–77.

Noy, S. and W. Zhang (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science 381(6654), 187–192.

Peng, S., E. Kalliamvakou, P. Cihon, and M. Demirer (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.

Spence, M. (1981). Signaling, screening, and information. In Studies in Labor Markets, pp. 319–358. University of Chicago Press.

Walther, T. and M. Dutordoir (2025). Certainly! generative ai and its impact on academic writing (in finance). Working Paper.

Wermers, R. (1997). Momentum investment strategies of mutual funds, performance persistence, and survivorship bias. Unpublished Working Paper, University of Colorado.

Wu, W., J. Chen, Z. Yang, and M. L. Tindall (2021). A cross-sectional machine learning approach for hedge fund return prediction and selection. Management Science 67(7), 4577–4601.

Figure 1. Architecture of a Long Short-Term Memory (LSTM) Cell

This figure illustrates the architecture of a *Long Short-Term Memory (LSTM)* cell. The cell consists of three main components: the forget gate, the input gate, and the output gate. The forget gate determines how much of the previous cell state (c_{t-1}) is retained, based on the sigmoid activation function (f_t). The input gate regulates how much new information, derived from the current input (X_t) and prior hidden state (h_{t-1}), is incorporated into the cell state via the sigmoid and tanh functions (i_t, $\tilde{c}_t$). The output gate controls how much of the updated cell state (c_t) influences the current hidden state (h_t), again using sigmoid and tanh activations (o_t). The diagram uses symbols to denote operations: circles with a cross represent multiplication, circles with a plus represent addition, and labeled blocks represent nonlinear activations (sigmoid or tanh). Together, these components enable the LSTM cell to preserve long-term dependencies while mitigating vanishing or exploding gradient problems during training.

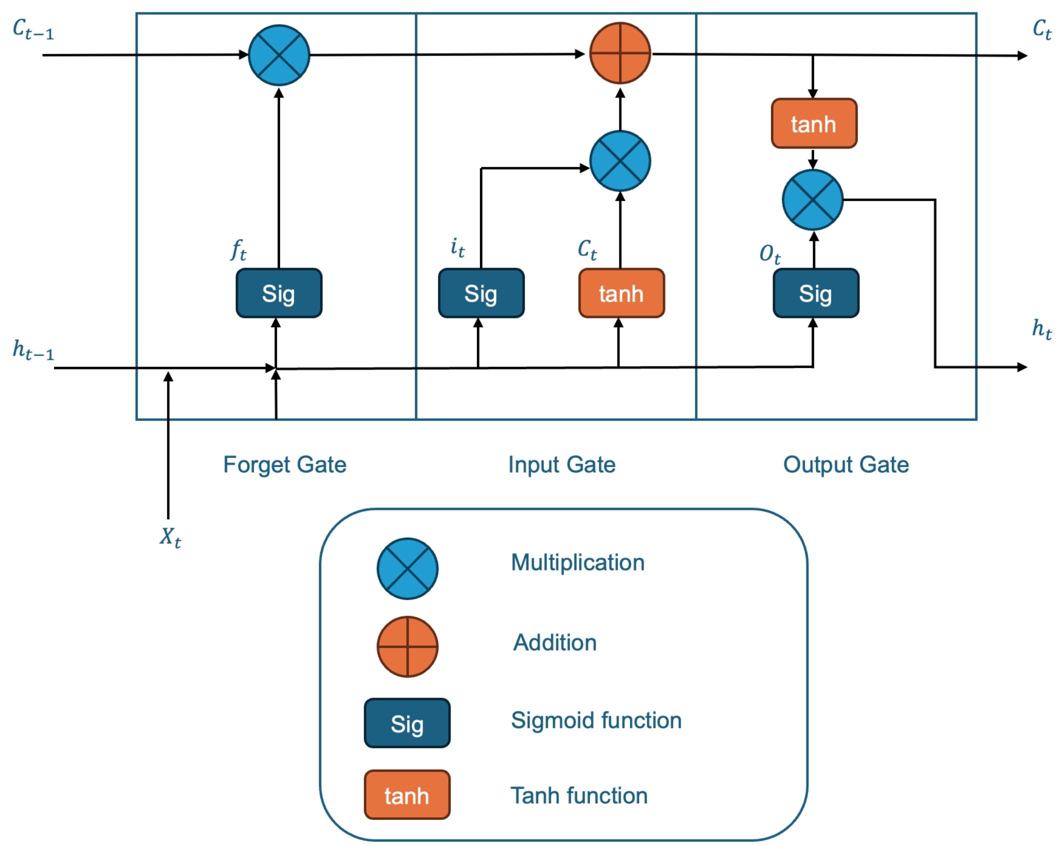


Figure 2. The training and testing time periods

This figure illustrates the rolling training and testing periods used in the analysis. Each observation window consists of a five-year training period followed by a two-year testing period, with both windows advancing forward one quarter at a time. For example, Observation 1 uses data from 2000Q1–2004Q4 for training and 2005Q1–2006Q4 for testing. Observation 2 shifts forward by one quarter, training on 2000Q2–2005Q1 and testing on 2005Q2–2007Q1, while Observation 3 trains on 2000Q3–2005Q2 and tests on 2005Q3–2007Q2. This rolling-window design ensures that predictions are always generated out-of-sample, preventing look-ahead bias and capturing the time-varying nature of fund and market characteristics.

Figure 3. Performance Precision

This figure compares the distribution of model-based *Weighted Average Precision* with the *Weighted Average Naive Precision*. The vertical dashed lines indicate the mean values.

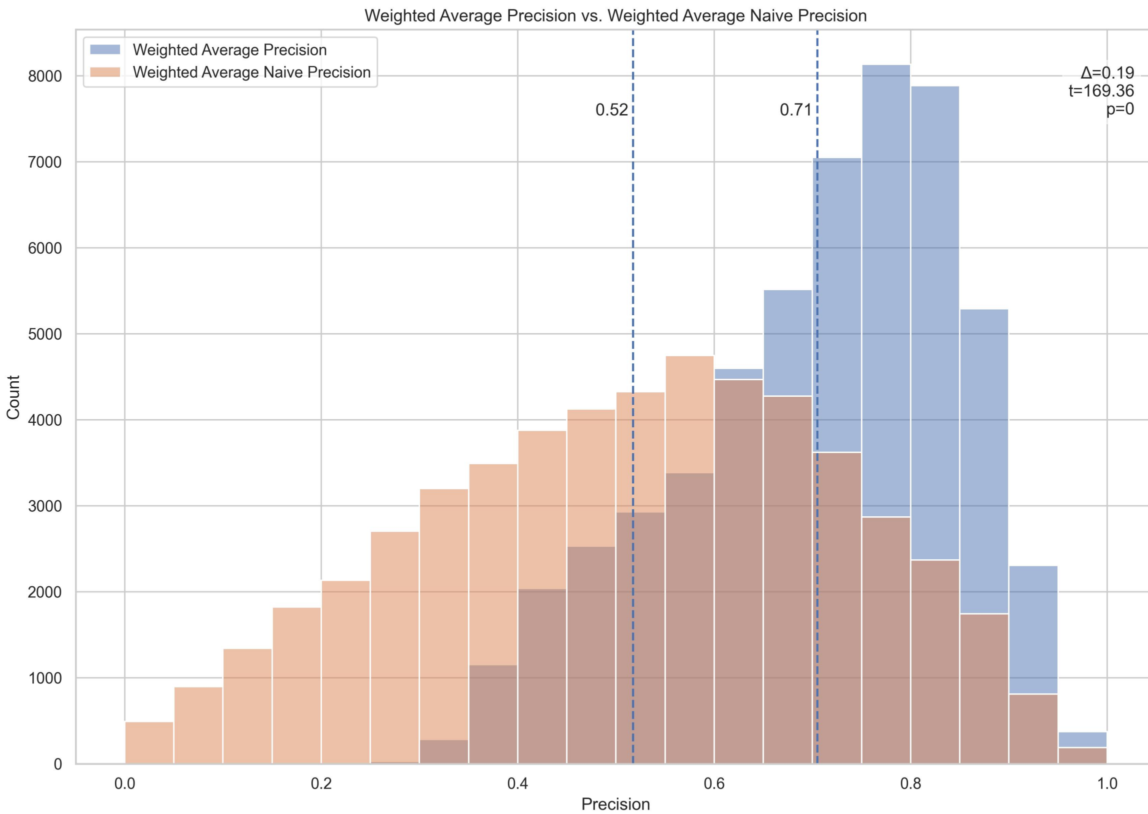
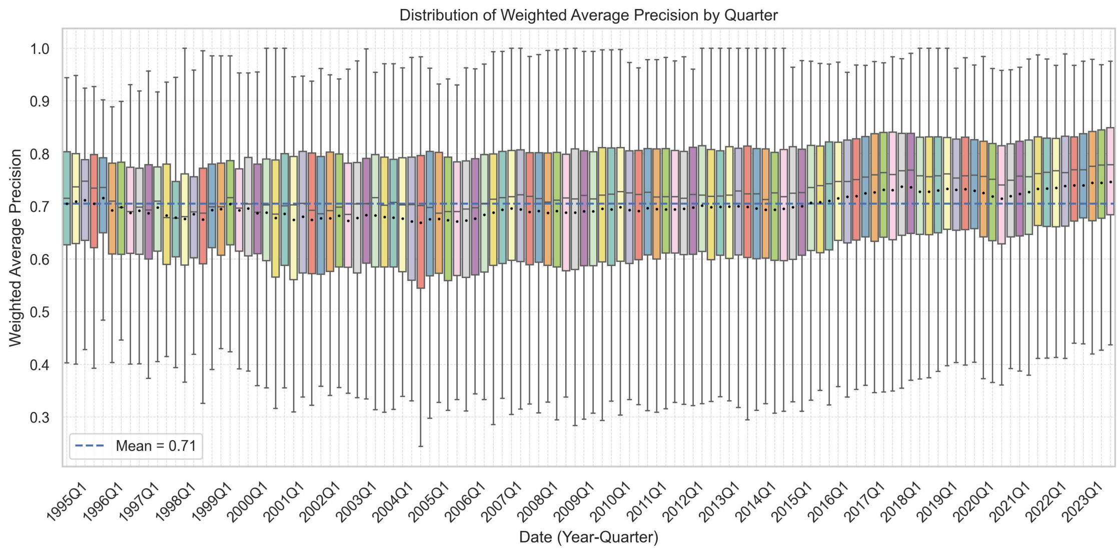


Figure 4. Performance Precision Over Time

This figure shows the distribution of *Prediction Precision* over time. The top Panel A reports the quarterly distribution of weighted average precision from 1995Q1 to 2023Q1, while the bottom Panel B shows the corresponding distribution for the naive benchmark. Each box plot summarizes the cross-sectional distribution of funds in a given quarter, with the dashed horizontal line representing the time-series mean.

Panel A



Panel B

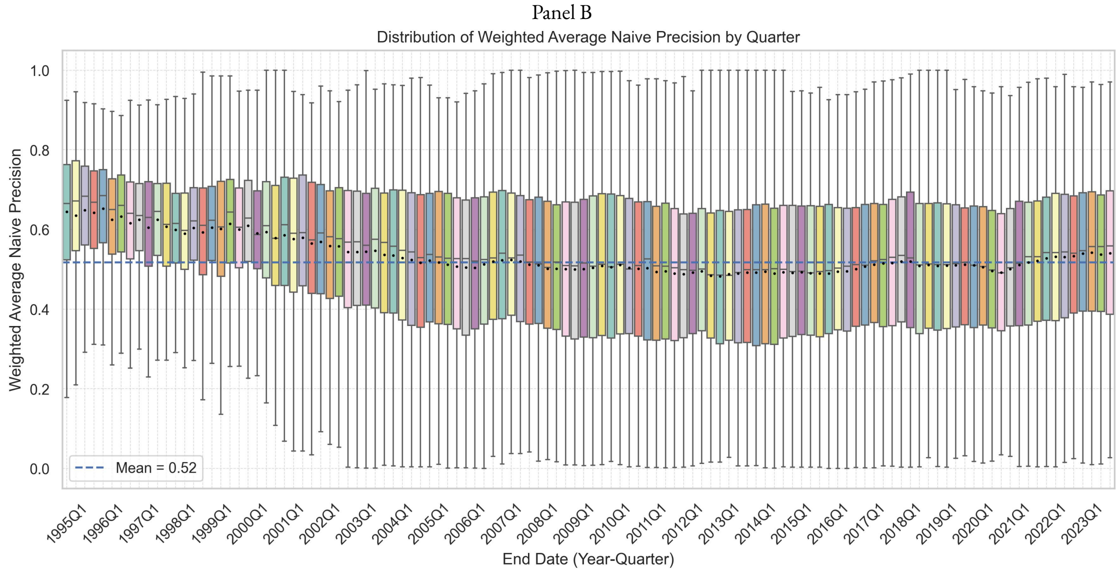


Figure 5. Performance Precision By Morningstar Style Category

This figure presents the distribution of *Prediction Precision* and the *Naive Prediction Precision* across Morningstar fund style categories. Each panel shows a histogram of accuracy (x-axis) and fund counts (y-axis) for Large, Mid, and Small funds split into Growth, Value, and Blend categories. Vertical dashed lines denote the mean precision for each category.

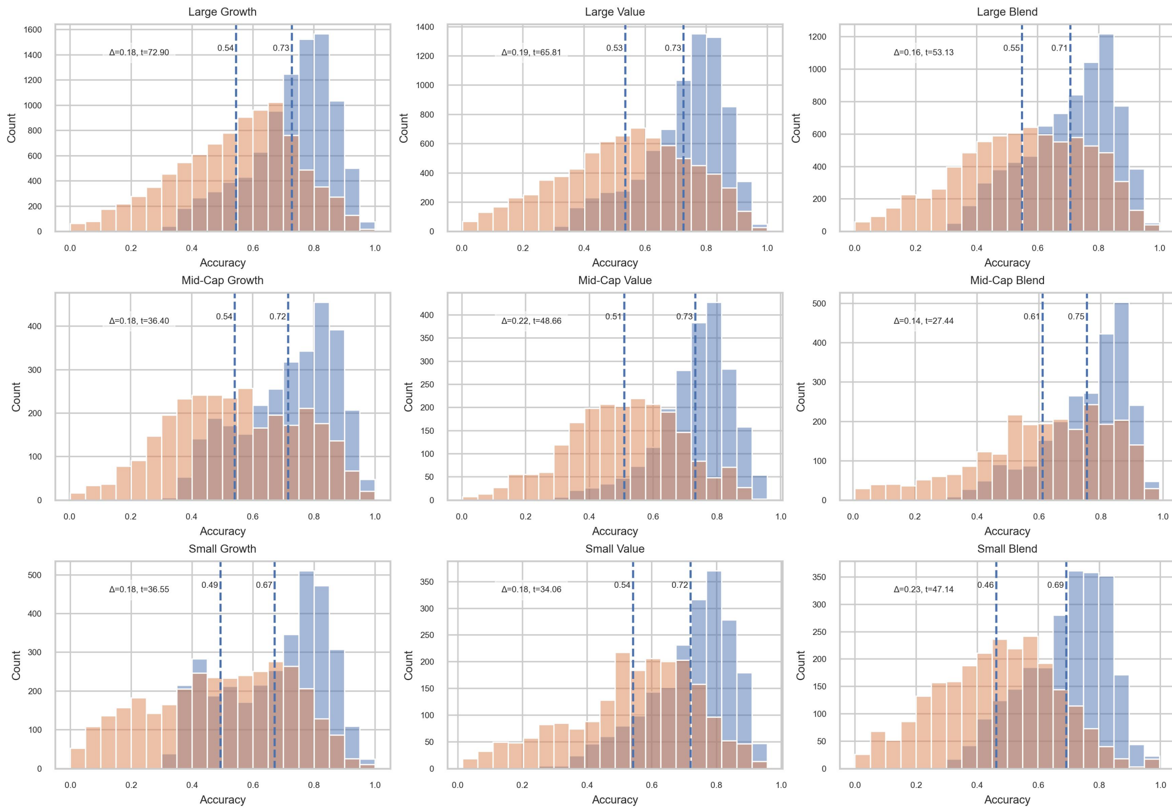


Figure 6. Performance Precision By Fund Group

This figure plots the distribution of *Prediction Precision* across fund size groups. The x-axis reports weighted average precision, while the y-axis shows the density of funds. Separate histograms are displayed for Large, Mid, Small, and Other funds, with vertical dashed lines indicating the mean precision for each group.

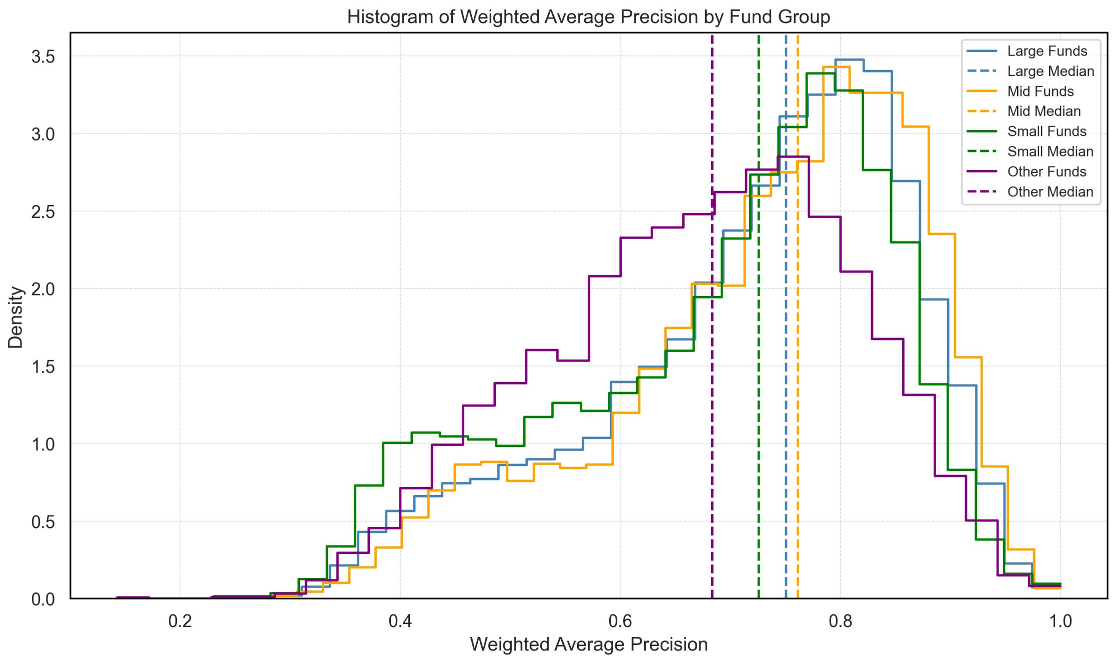


Figure 7. Performance Precision By Fund Group Over Time

This figure plots the evolution of *Prediction Precision* over time by fund size group. Funds are classified each quarter into four categories: Large Funds, Mid Funds, Small Funds, and Other Funds. The y-axis reports the average adjusted precision of predictions, while the x-axis denotes calendar time from 1995Q1 to 2023Q1.

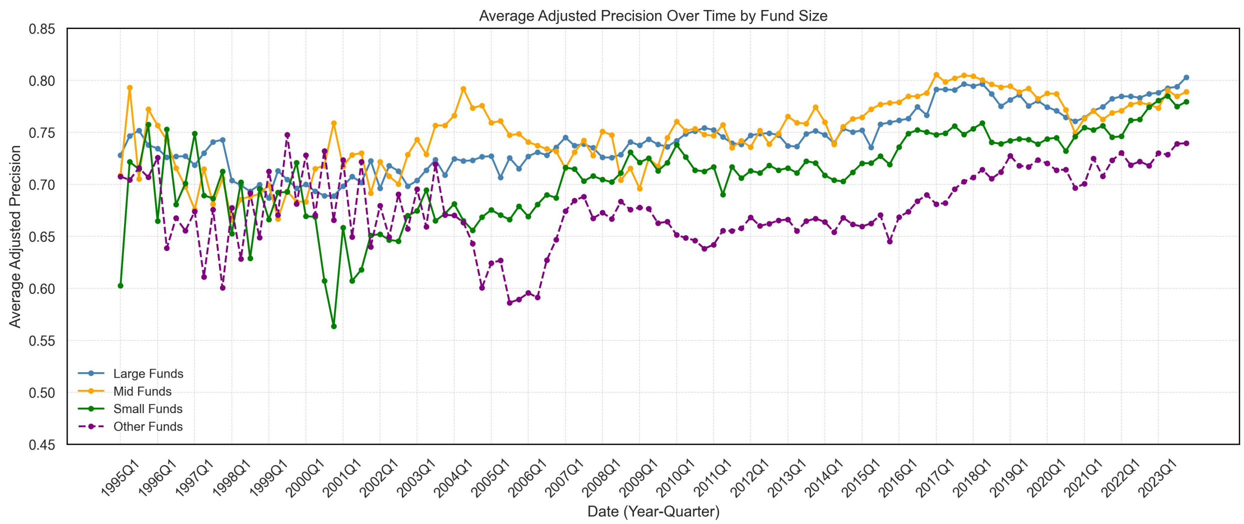


Figure 8. Performance Precision And Future Fund Performance

This figure illustrates the relationship between *Prediction Precision* and subsequent fund performance. Funds are sorted into quintiles based on lagged weighted-average (WA) prediction precision, with Q1 denoting the lowest-precision quintile and Q5 the highest. The panels report average cumulative returns over 1-, 2-, 3-, and 4-quarter horizons following portfolio formation.

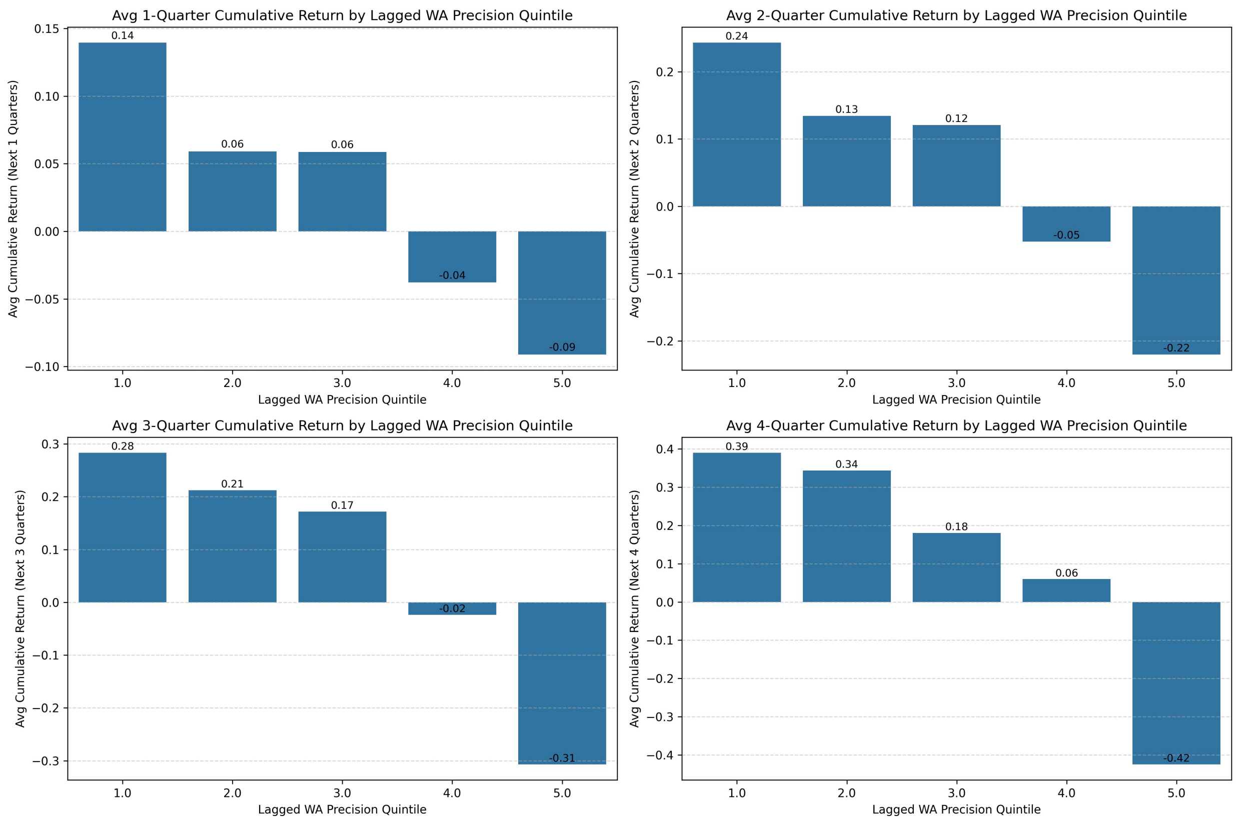
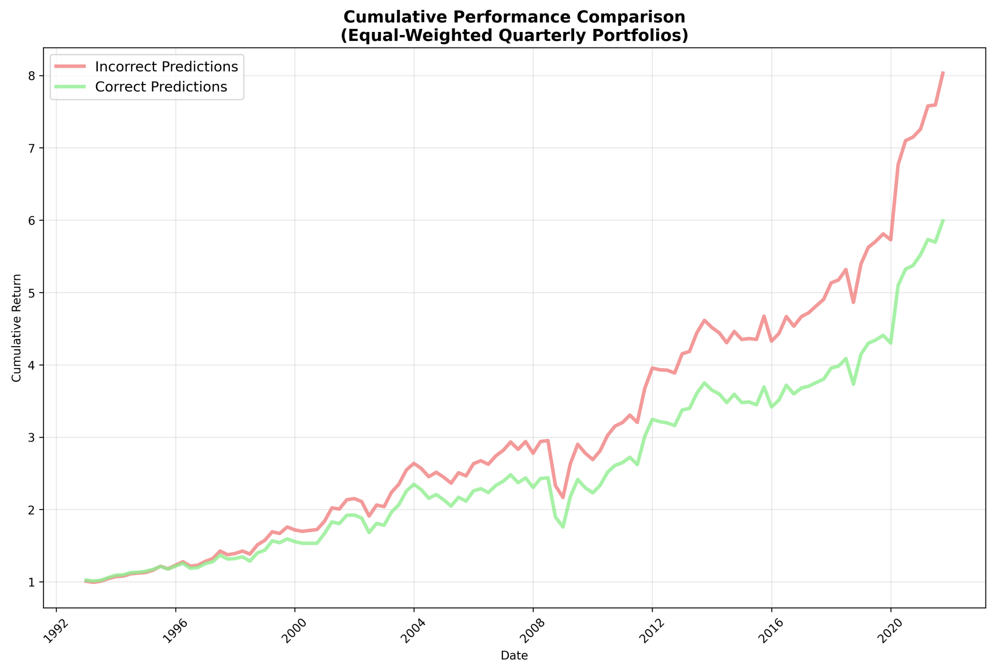


Figure 9. Portfolio Returns of Correct vs. Incorrect Prediction Components of Funds

This Figure show returns of correct vs. incorrect prediction portions of funds. For each fund-quarter observation, stocks held by each fund are classified into two portfolios: "Incorrect Predictions" portfolio that contains stocks where the fund's predicted direction was wrong, while "Correct Predictions" contains stocks where the fund's predicted direction was accurate. The upper panel displays cumulative returns over time for two equal-weighted component portfolios. The lower panel quantifies the performance differential between the two portfolios by displaying the cumulative gap, calculated as the cumulative return of Incorrect Predictions minus Correct Predictions.



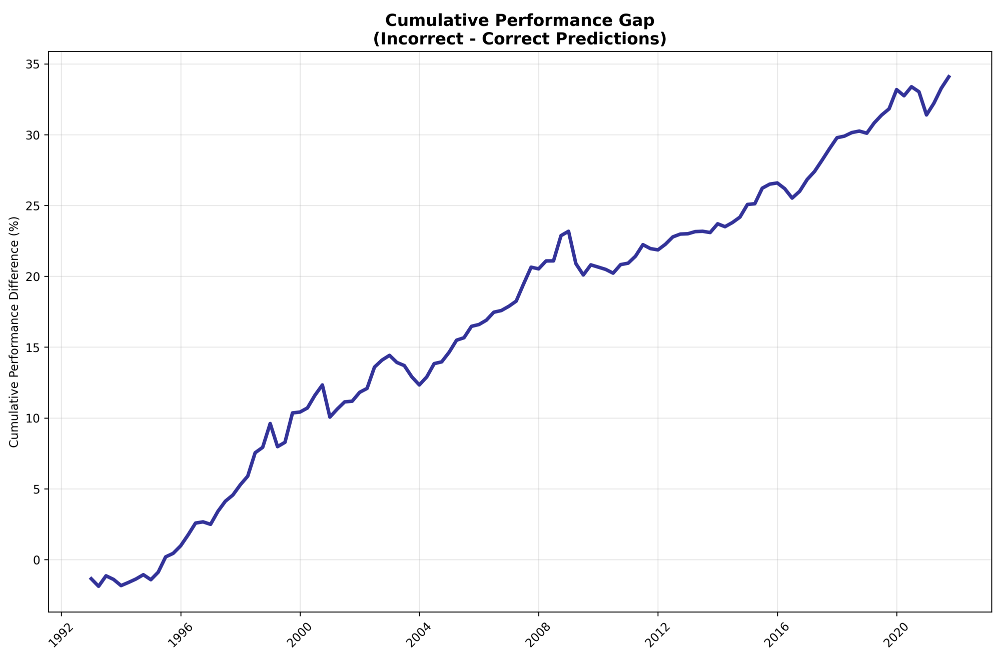


Figure 10. Portfolio Returns of Stocks Sorted on Percent of Correct Prediction

This Figure show two complementary views of quintile portfolio performance based on fund manager percent correct prediction. For each quarter, stocks are ranked by the average prediction accuracy of all funds holding that stock and sorted into five equal-weighted portfolios. Quintile 1 (Q1) captures stocks where fund managers have the lowest percent correct prediction (i.e., funds are least accurate in predicting these stocks' return directions), while Quintile 5 (Q5) contains stocks where fund managers demonstrate the highest percent correct prediction accuracy. The top panel displays the cumulative returns over time for each quintile, showing the long-term performance evolution of portfolios formed on prediction accuracy. The second panel shows the cumulative return of Quintile 1 (Q1) minus Quintile 5 (Q5). The bottom panel shows the average quarterly returns for each quintile with statistical significance indicators. Each bar displays the mean quarterly return, with t-statistics and significance levels shown in parentheses below (*** p<0.01, ** p<0.05, * p<0.10).

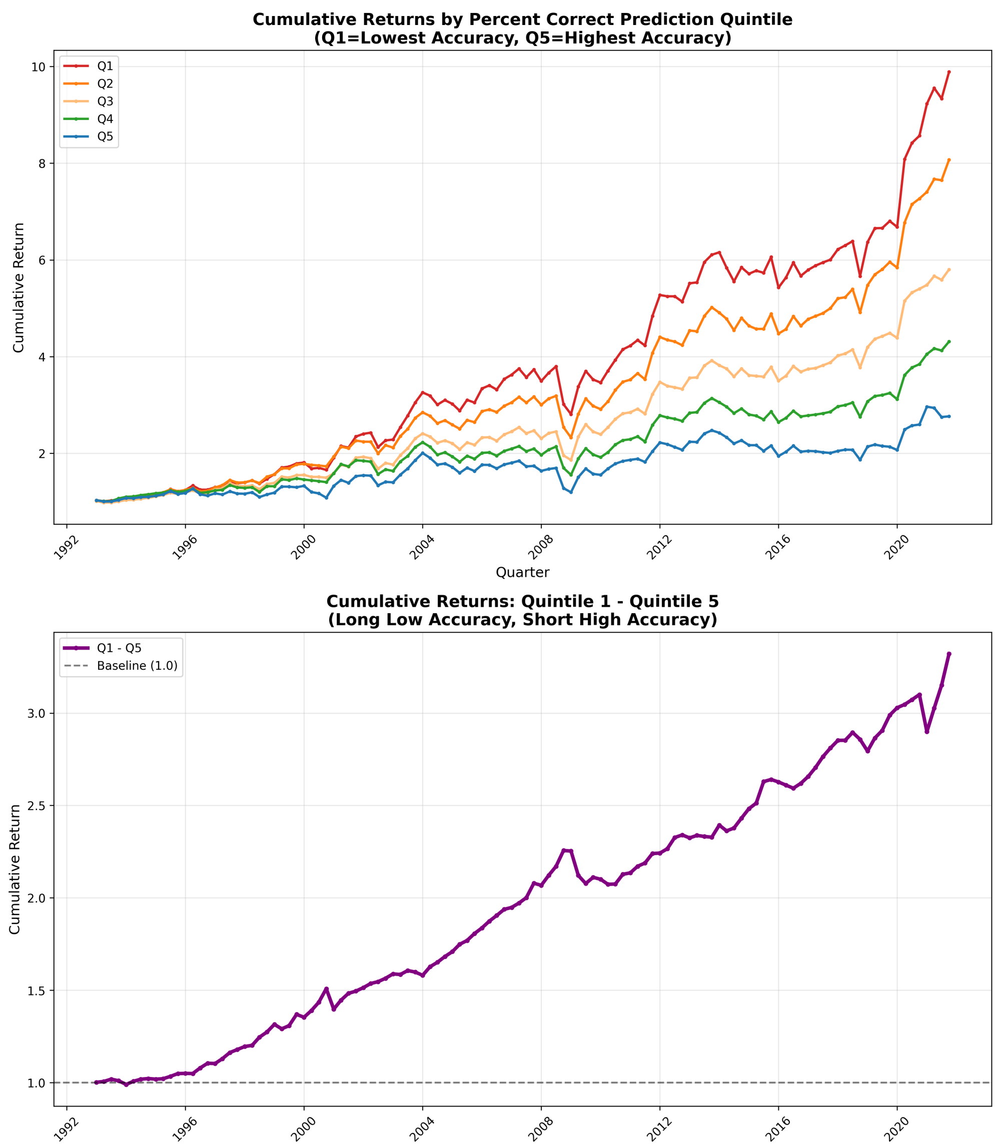


Figure 11. Portfolio Returns of Stocks Sorted on Percent of Correct Prediction

This Figure presents cumulative performance trajectories for quintile portfolios formed on funds' Prediction Precision across the entire sample period, displayed across multiple risk-adjusted return specifications. For each quarter, stocks are ranked by the average prediction accuracy of all funds holding that stock and sorted into five equal-weighted portfolios. Quintile 1 (Q1) captures stocks where fund managers are least accurate in predicting directional returns, while Quintile 5 (Q5) contains stocks where fund managers demonstrate the highest prediction accuracy. The figure displays four panels, each representing a different return metric: cumulative excess returns, cumulative FF3 alpha (adjusting for market, size, and value factors), cumulative FF4 alpha (adding momentum), and cumulative FF5 alpha (adding profitability and investment factors).

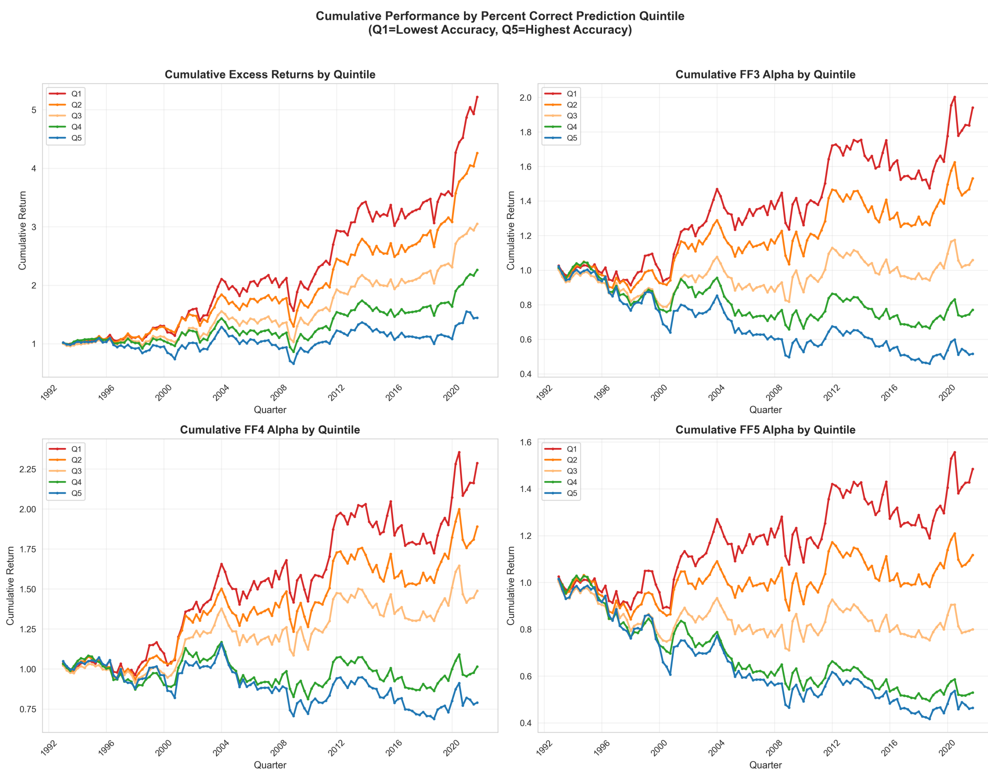


Table I. Summary Statistics – Morningstar Fund Characteristics

Panel A presents summary statistics on manager characteristics obtained from Morningstar. *Manager Ownership* represents the average dollar ownership across all managers. *Manager Tenure (Mean)* and *Manager Tenure (Max)* report, respectively, the average and maximum tenure of managers, measured in quarters. *Manager Funds* is the average number of funds managed per manager. *Manager Styles* denotes the average number of distinct investment styles overseen by each manager. Panel B presents summary statistics on fund characteristics obtained from Morningstar. *Fund Age* is the number of quarters since the inception of the fund. *Within Category Competition* measures the number of funds operating within the same investment category as the focal fund in a given quarter. *Within Management Company Competition* denotes the number of other funds managed by the same management company during the quarter. *Number of Managers* is the count of individual portfolio managers responsible for managing the fund. The unit of observations in both quarters are the fund-quarter.

Panel A: Manager Characteristics

	Mean	Std. Dev.	Min	Median	Max
Manager Ownership ($10,000)	79.37	75.08	0.00	100.00	320.00
Manager Tenure (Mean)	52.49	19.17	28.00	48.00	115.80
Manager Tenure (Max)	61.25	28.16	28.00	53.00	155.00
Manager # Funds	2.83	4.35	1.00	2.00	29.00
Manager # Styles	2.00	2.06	1.00	1.00	12.00
Observations	53878				

Panel B: Fund Characteristics

	Mean	Std. Dev.	Min	Median	Max
Fund Age	57.20	25.57	28.00	50.00	165.00
Within Category Competition	75.90	79.83	1.00	35.00	218.00
Within Management Company Competition	4.79	6.82	1.00	2.00	36.00
Number of Managers	3.60	2.81	1.00	3.00	31.00
Observations	53878				

Table II. Summary Statistics – CRSP Fund Characteristics

This Table presents the summary statistics on manager characteristics obtained from CRSP. *Prediction Precision* is the weighted average correct predictions of the direction of the subsequent quarter's share change generated by the model. *Naive Prediction Precision* is the weighted average correct predictions of the direction of the subsequent quarter's share change generated by the naive benchmark model. *Total Net Asset* is the total market value of fund assets (in millions of dollars). *Fund Flow* is the net quarterly flow into the fund. *Income Yield* is the annualized dividend yield of the fund. *NAV/52 week high NAV* is the ratio of the current net asset value (NAV) to the highest NAV in the prior 52 weeks. *Actual 12b1* is the annual distribution (12b-1) fee charged by the fund. *Management Fee* is the annual management fee charged by the fund. *Expense Ratio* is the ratio of total annual expenses to average net assets. *Turnover Ratio* is the proportion of the fund's portfolio replaced during the year. *Open to Investment* is an indicator equal to one if the fund is open to new investors, and zero otherwise. *Has Sales Restrictions* is an indicator equal to one if the fund imposes restrictions on sales, and zero otherwise.

	Mean	Std. Dev.	Min	Median	Max
Prediction Precision	0.71	0.14	0.14	0.73	1.00
Naive Prediction Precision	0.52	0.21	0.00	0.53	1.00
Total Net Asset	667.87	1391.91	0.30	178.50	8987.90
Fund Flow	5.29	108.74	-464.72	0.20	548.62
Income Yield	0.01	0.01	0.00	0.01	0.72
NAV/52 week high NAV	0.89	0.12	0.03	0.94	1.00
Actual 12b1	0.00	0.00	0.00	0.00	0.01
Management Fee	0.78	0.29	0.04	0.77	1.95
Expense Ratio	1.26	0.49	0.37	1.19	3.26
Turnover Ratio	0.71	0.85	0.03	0.46	5.84
Open to Investment	0.95	0.21	0.00	1.00	1.00
Has Sales Restrictions	0.01	0.08	0.00	0.00	1.00
Observations	53489				

Table III. Summary Statistics – Firm Level

This table reports summary statistics of firms. *Position Size* is the dollar value of a fund's holding in a given firm. *Portfolio Position* is the order of a fund's holding of the firm within its overall portfolio, with higher values indicating smaller relative positions. *Market Equity* is the market value of the firm's equity (in millions of dollars). *#Funds Position Held In* is the number of distinct funds that hold a position in the firm in a given quarter. *Firm Age* is the number of months since the firm's first appearance in CRSP. *Total Debt/ME* is the ratio of total debt to market equity. *Asset Turnover* is sales divided by total assets. *Sales scaled/BE* is sales scaled by book equity. *R&D/Sales* is research and development expenditures scaled by sales. *Gross Profit/ME* is gross profits scaled by market equity. *Free Cash Flow/ME* is free cash flow scaled by market equity. *Book Equity/ME* is the ratio of book equity to market equity. *Net Equity Payout/ME* is net equity payouts (repurchases minus issuances plus dividends) scaled by market equity. *Earnings Surprise* is the difference between reported quarterly earnings and the most recent consensus analyst forecast, scaled by the stock price. *Short Term Reversal* is the past one-month stock return (negative values indicate recent losers). *Momentum 1–6 Months* is the cumulative stock return from months 1 through 6 prior to portfolio formation. *Momentum 1–12 Months* is the cumulative stock return from months 1 through 12 prior to portfolio formation. *CAPM Idiosyncratic Vol.* is the standard deviation of residuals from a CAPM regression estimated over the prior 36 months. The unit of observations is fund-firm–quarter.

Table III. Summary Statistics – Firm Level (continued)

Panel A: Fund-Firm Holding Characteristics

	Mean	Std. Dev.	Min	Median	Max
Portfolio Position Size	5.8e+06	3.0e+07	0.00	0.00	4.2e+09
Portfolio Position	43.70	24.15	1.00	42.00	100.00
Observations	5434702				

Panel B: Firm Fundamental Characteristics

	Mean	Std. Dev.	Min	Median	Max
Market Equity	29988.67	84275.90	0.00	6118.08	2.5e+06
#Funds Position Held In	55.89	46.09	1.00	44.00	263.00
Firm Age	400.25	268.23	0.00	322.00	1150.00
Total Debt/ME	0.48	0.97	0.00	0.19	6.95
Asset Turnover	1.25	420.98	0.00	0.69	4.6e+05
Sales scaled/BE	2.08	2.71	0.09	1.31	18.93
R&D/Sales	0.12	0.31	0.00	0.04	2.55
Gross Profit/ME	0.30	0.22	-0.14	0.26	1.02
Free Cash Flow/ME	0.04	0.11	-0.48	0.04	0.46
Book Equity/ME	0.52	0.44	0.02	0.39	2.58
Net Equity Payout/ME	0.04	0.05	0.00	0.03	0.27
Earnings Surprise	0.20	9.88	-271.69	0.17	3463.89
Observations	5434702				

Panel C: Firm Return Characteristics

	Mean	Std. Dev.	Min	Median	Max
Short Term Reversal	0.02	0.14	-1.00	0.01	17.60
Momentum 1-6 Months	0.08	7.62	-1.00	0.05	8899.00
Momentum 1-12 Months	0.17	1.64	-1.00	0.11	1259.00
CAPM Idiosyncratic Vol.	0.02	0.02	0.00	0.02	14.22
Observations	4807535				

Table IV. Manager Characteristics and Prediction Precision

This table shows the relationship between *Manager Characteristics* and *Prediction Precision* at the fund level. *Manager Ownership* (in $10000) is the average dollar ownership of fund managers in the corresponding fund. *Manager Tenure (Max)* is the maximum tenure, measured in quarters, among all managers of a given fund, while *Manager Tenure (Mean)* is the average tenure across managers. *Manager #Funds* is the average number of funds simultaneously managed by a given manager. *Manager #Styles* is the average number of distinct investment styles overseen by each manager. *Prediction Precision* is the fraction of correct predictions produced by the model at the fund level. *Quarter-Year FE* and *Fund FE* indicate the inclusion of quarter-year fixed effects and fund fixed effects, respectively. Standard errors are clustered at the fund level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Prediction Precision				
Manager Ownership	-0.0003***				
	(10.131)				
Manager Tenure (Max)		0.0007***			
		(4.253)			
Manager Tenure (Mean)			0.0006***		
			(9.296)		
Manager #Funds				0.0007***	
				(3.424)	
Manager #Styles					0.0023***
					(4.629)
Constant	0.6328***	0.6651***	0.6523***	0.6587***	0.6572***
	(111.445)	(320.586)	(338.059)	(387.955)	(382.550)
Observations	16,208	53,878	53,878	53,878	53,878
R-squared	0.723	0.673	0.673	0.673	0.673
Quarter-Year FE	YES	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES	YES

Table V. Fund Characteristics and Prediction Precision

This table shows the relationship between *Fund Characteristics* and *Prediction Precision* at the fund level. *Fund Age* is the number of quarters since the fund's inception. *Within Category Competition* measures the number of other funds operating within the same Morningstar investment category as the focal fund in a given quarter. *Within Management Company Competition* denotes the number of other funds managed by the same management company during the quarter. *Number of Managers* is the count of individual portfolio managers responsible for managing the fund. *Prediction Precision* is the fraction of correct predictions produced by the model at the fund level. *Quarter-Year FE* and *Fund FE* indicate the inclusion of quarter fixed effects and fund fixed effects, respectively. Standard errors are clustered at the quarter level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	Prediction Precision			
Fund Age	0.0013***			
	(61.440)			
Within Category Competition		-0.0001***		
		(3.058)		
Within Management Company Competition			-0.0023***	
			(3.260)	
Number of Managers				-0.0014***
				(2.705)
Constant	0.6788***	0.6531***	0.6471***	0.6592***
	(471.378)	(500.897)	(419.093)	(388.559)
Observations	53,878	52,536	40,803	53,878
R-squared	0.673	0.671	0.672	0.673
Quarter-Year FE	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES

Table VI. Fund Characteristics and Prediction Precision - Fund CRSP Characteristics

This table shows the relationship between *Fund CRSP Characteristics* and *Prediction Precision* at the fund level. *Total Net Asset* is the total market value of the fund's assets (in millions of dollars). *Fund Flow* is the net quarterly flow into the fund, scaled by total net assets. *Actual 12b1* is the annual distribution (12b-1) fee charged by the fund. *Management Fee* is the annual management fee charged by the fund. *Expense Ratio* is the ratio of total annual expenses to average net assets. *Turnover Ratio* is the proportion of the fund's portfolio replaced during the year. *Open to Investment* is an indicator equal to one if the fund is open to new investors, and zero otherwise. *Has Sales Restrictions* is an indicator equal to one if the fund imposes restrictions on sales (e.g., minimum holding period or redemption fees), and zero otherwise. *Prediction Precision* is the fraction of correct predictions produced by the model at the fund level. Standard errors are clustered at the quarter level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table VI. Fund Characteristics and Prediction Precision - Fund CRSP Characteristics (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES				Prediction Precision			
Total Net Asset	-2.1384***						
	(4.812)						
Turnover Ratio		-0.0149***					
		(9.712)					
Fund Flow			-4.5290**				
			(2.182)				
Management Fee				-0.0078**			
				(2.603)			
Expense Ratio					-0.0046		
					(1.432)		
Open to Investment						-0.0060***	
						(2.748)	
Has Sales Restrictions							0.0578***
							(4.874)
Constant	0.6612***	0.6737***	0.6620***	0.6709***	0.6678***	0.5908***	0.6472***
	(403.580)	(337.763)	(410.220)	(175.306)	(137.071)	(134.766)	(326.832)
Observations	50,983	46,043	50,962	45,864	46,231	48,177	46,527
R-squared	0.665	0.684	0.664	0.686	0.682	0.686	0.697
Year FE	YES	YES	YES	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES	YES	YES	YES

Table VII. Prediction Precision and Firm Characteristics

This table shows the relationship between *Firm Characteristics* and *Prediction Precision* at the firm level. *Portfolio Position* is the position of the security's market value within the fund's portfolio, with higher values indicating lower relative holdings. *Market Equity* is the market value of the firm's equity (in millions of dollars). *# Funds Position Held In* is the number of distinct funds holding the security in a given quarter. *Firm Age* is the number of months since the firm's first appearance in CRSP. *Prediction Precision* is the fraction of correct predictions produced by the model at the security level. *Quarter-Year FE* indicate the inclusion of quarter fixed effects. Standard errors are two-way clustered at the firm level and quarter level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES		Prediction Precision		
Portfolio Position	0.0020***			
	(52.407)			
Market Equity		-0.0000***		
		(3.375)		
#Funds Position Held In			-0.0002***	
			(4.659)	
Firm Age				0.0000***
				(5.114)
Constant	-0.3202***	-0.2566***	-0.2563***	-0.2677***
	(62.793)	(52.432)	(50.564)	(50.839)
Observations	5,958,755	5,316,273	5,958,755	5,327,875
R-squared	0.253	0.252	0.244	0.252
Quarter-Year FE	YES	YES	YES	YES

Table VIII. Prediction Precision and Security Characteristics

This table shows the relationship between *Firm Fundamental Characteristics* and *Prediction Precision* at the security level. *Total Debt/ME* is the ratio of total debt to market equity. *Asset Turnover* is the ratio of sales to total assets. *Sales scaled/BE* is sales scaled by book equity. *R&D/Sales* is research and development expenditures scaled by sales. *Gross Profit/ME* is gross profits scaled by market equity. *Free Cash Flow/ME* is free cash flow scaled by market equity. *Book Equity/ME* is the ratio of book equity to market equity. *Net Equity Payout/ME* is net equity payouts (repurchases minus issuances plus dividends) scaled by market equity. *Earnings Surprise* is the difference between reported quarterly earnings and the most recent consensus analyst forecast, scaled by the stock price. *Prediction Precision* is the fraction of correct predictions produced by the model at the security level. Standard errors are two-way clustered at the firm and quarter level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table VIII. Prediction Precision and Security Characteristics (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Prediction Precision								
Total Debt/ME	0.0083***								
	(4.739)								
Asset Turnover		-0.0000***							
		(66.758)							
Sales/BE			0.0002						
			(0.581)						
R&D/Sales				-0.0092***					
				(2.668)					
Gross Profit/ME					0.0005				
					(0.110)				
Free Cash Flow/ME						0.0235***			
						(3.373)			
Book Equity/ME							0.0365***		
							(16.053)		
Net Equity Payout/ME								0.2564***	
								(16.216)	
Earnings Surprise									-0.0002**
									(1.992)
Constant	0.1305***	0.1349***	0.1327***	0.1365***	0.1348***	0.1319***	0.1198***	0.1257***	0.1303***
	(50.853)	(52.026)	(48.556)	(40.811)	(44.928)	(50.260)	(42.654)	(46.769)	(49.185)
Observations	5,323,840	5,320,129	5,188,708	2,995,288	5,318,193	5,234,659	5,209,346	4,983,976	5,119,510
R-squared	0.223	0.223	0.223	0.223	0.223	0.223	0.224	0.223	0.224
Quarter-Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table IX. Prediction Precision and Firms Return Characteristics

This table shows the relationship between *Firm Return Characteristics* and *Prediction Precision* at the security level. *Past 1 Month Return* is the past one-month stock return (negative values indicate recent losers). *Momentum 1–6 Months* is the cumulative stock return from months 1 through 6 prior to portfolio formation. *Momentum 1–12 Months* is the cumulative stock return from months 1 through 12 prior to portfolio formation. *CAPM Idiosyncratic Vol.* is the standard deviation of residuals from a CAPM regression estimated over the prior 36 months. *Prediction Precision* is the fraction of correct predictions produced by the model at the security level. Standard errors are two-way clustered at the firm and quarter level. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	Prediction Precision			
Past 1 Month Return	-0.0239***			
	(3.089)			
Momentum 1-6 Months		-0.0000		
		(0.928)		
Momentum 1-12 Months			-0.0019	
			(1.341)	
CAPM Idiosyncratic Vol.				0.1857*
				(1.722)
Constant	0.1354***	0.1345***	0.1349***	0.1312***
	(51.770)	(51.528)	(50.869)	(36.589)
Observations	5,286,960	5,257,523	5,213,302	5,248,059
R-squared	0.223	0.223	0.224	0.224
Quarter-Year FE	YES	YES	YES	YES

Table X. Prediction Precision and Future Fund Performance

This table shows the relationship between *Prediction Precision* and subsequent fund performance. The sample is sorted into quintiles based on predicted precision, and future fund returns are reported for each quintile. $CRET_{0,1}$ through $CRET_{0,4}$ denote cumulative excess fund returns, computed as the fund returns minus the average returns of all funds within the same Morningstar category, over horizons of one to four quarters ahead, respectively. The column labeled *Q1–Q5* reports the difference in returns between the lowest-precision quintile (Q1) and the highest-precision quintile (Q5). t-statistics are shown in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Q1	Q2	Q3	Q4	Q5	Q1–Q5
$CRET_{[0,1]}$	0.14**	0.06	0.06	-0.04	-0.09	0.23**
	(2.05)	(0.82)	(1.12)	(0.64)	(1.52)	(2.13)
$CRET_{[0,2]}$	0.25***	0.13	0.12	-0.05	-0.22**	0.47***
	(2.58)	(1.31)	(1.22)	(0.56)	(2.07)	(2.87)
$CRET_{[0,3]}$	0.33***	0.21*	0.17	-0.02	-0.31**	0.64***
	(3.18)	(1.76)	(1.39)	(0.21)	(2.47)	(3.34)
$CRET_{[0,4]}$	0.36***	0.34**	0.18*	0.06	-0.42**	0.79***
	(2.88)	(2.31)	(1.66)	(0.44)	(2.36)	(3.05)

Table XI. Prediction Precision and Future Fund Performance

This table shows the relationship between *Prediction Precision* and subsequent fund performance. The sample is sorted into quintiles based on predicted precision, and future fund excess 4-quarter cumulative returns, and 4-quarter cumulative factor alphas. FF3 includes Mkt–RF, SMB, HML; FF4 includes Mkt–RF, SMB, HML, MOM; and FF5 includes Mkt–RF, SMB, HML, RMW, CMA. The column labeled *Q1–Q5* reports the difference in returns between the lowest-precision quintile (Q1) and the highest-precision quintile (Q5). t-statistics are shown in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Q1	Q2	Q3	Q4	Q5	Q1–Q5
Excess Return	0.365***	0.319**	0.156	0.035	-0.449**	0.815***
	(2.86)	(2.15)	(1.43)	(0.26)	(2.49)	(3.05)
FF3 Alpha	0.374***	0.285*	0.185	0.091	-0.439**	0.812***
	(2.80)	(1.85)	(1.62)	(0.64)	(2.33)	(2.85)
FF4 Alpha	0.308**	0.313*	0.281**	0.119	-0.475**	0.783***
	(2.24)	(1.95)	(2.45)	(0.80)	(2.43)	(2.66)
FF5 Alpha	0.370***	0.105	0.249**	0.079	-0.409**	0.779**
	(2.58)	(0.66)	(2.04)	(0.51)	(2.03)	(2.53)

Table XII. Portfolio Returns of Incorrect vs. Correct Predictions of Fund Portfolios

This Table reports the performance comparison between two components of fund portfolios based on prediction accuracy at the stock level. For each fund-quarter observation, stocks held by each fund are classified into two portfolios: "Harder to Predict" portfolio that contains stocks where the fund's predicted direction was wrong, while "Easier to Predict" contains stocks where the fund's predicted direction was accurate. The estimates of the returns and alphas are annualized by multiplying by four. The table presents mean annualized returns, excess returns, factor alphas, and t-statistics. FF3 includes Mkt−RF, SMB, HML; FF4 includes Mkt−RF, SMB, HML, MOM; and FF5 includes Mkt−RF, SMB, HML, RMW, CMA. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Harder to Predict	Easier to Predict	Difference
Return	7.981***	7.069***	0.912***
	(61.59)	(54.62)	(12.41)
Excess Return	6.490***	5.579***	0.912***
	(49.92)	(42.92)	(12.41)
FF3 Alpha	1.897***	0.800***	1.097***
	(17.12)	(7.44)	(14.37)
FF4 Alpha	2.177***	1.276***	0.902***
	(19.29)	(11.68)	(11.60)
FF5 Alpha	0.624***	-0.459***	1.083***
	(5.21)	(3.95)	(13.03)

Table XIII. Portfolio Returns of Stocks Sorted on Percent of Correct Prediction

This Table reports the performance of quintile stock portfolios formed on stocks' percent of correct directional prediction by fund managers that held that stock. Each quarter, stocks are ranked by the average prediction accuracy of all funds holding that stock and sorted into five equal-weighted portfolios. Quintile 1 (Q1) captures stocks where, on average, the behavior of fund managers are least predictable, while Quintile 5 (Q5) contains stocks where, on average, the behavior of fund managers are most predictable. The Q1–Q5 column presents the quarterly performance differential between the lowest and highest accuracy quintiles. The estimates of the returns and alphas are annualized by multiplying by four. FF3 includes Mkt–RF, SMB, HML; FF4 includes Mkt–RF, SMB, HML, MOM; and FF5 includes Mkt–RF, SMB, HML, RMW, CMA. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Q1 (Lowest)	Q2	Q3	Q4	Q5 (Highest)	Q1–Q5
Return	8.707***	7.884***	6.745***	5.765***	4.467*	4.240***
	(3.85)	(3.79)	(3.17)	(2.61)	(1.73)	(5.74)
Excess Return	6.476***	5.653***	4.513**	3.533	2.236	4.240***
	(2.85)	(2.70)	(2.11)	(1.59)	(0.86)	(5.74)
FF3	2.689	1.803	0.526	-0.543	-1.736	4.425***
	(1.53)	(1.12)	(0.33)	(0.33)	(0.85)	(6.08)
FF4	3.254*	2.524	1.686	0.392	-0.303	3.557***
	(1.78)	(1.52)	(1.03)	(0.23)	(0.15)	(5.12)
FF5	1.753	0.706	-0.446	-1.842	-2.109	3.862***
	(0.92)	(0.41)	(0.26)	(1.03)	(0.95)	(4.96)

Table XIV. Prediction Precision and Active Shares

This Table reports Fama-Macbeth regression results examining the relationship between mutual fund managers' trade predictability and subsequent abnormal returns, measured as cumulatively quarter 0 through quarter 4. Active Share is defined in Cremers and Petajisto's (2009) as the fraction of a mutual fund's portfolio holdings that differ from its benchmark index. Standard errors are adjusted following the Newey-West procedure to account for autocorrelation and heteroskedasticity. t-statistics are shown in parentheses below coefficient estimates. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
VARIABLES		$CRET_{[0,4]}$	
Prediction Precision	-0.8554**	-1.0074**	-1.0261***
	(2.06)	(2.38)	(2.70)
Active Share		0.0041***	0.0032***
		(5.13)	(4.57)
$RET_{[-1,0]}$			0.1311**
			(2.09)
$CRET_{[-4,0]}$			0.0628**
			(2.12)
Cons	0.6736**	0.5527**	0.5648**
	(2.39)	(2.01)	(2.25)
Observations	40585	40585	40585
R-squared	0.00533	0.00854	0.106

Appendix for

"Mimicking Finance"

A Variable Descriptions

Variable	Description
Actual 12b1	Annual distribution (12b-1) fee charged by the fund.
Asset Turnover	Ratio of sales to total assets.
Book Equity/ME	Ratio of book equity to market equity.
CAPM Idiosyncratic Vol.	Standard deviation of residuals from a CAPM regression estimated over the prior 36 months.
$CRET_{0,1}$	Cumulative excess fund return one quarter ahead.
$CRET_{0,2}$	Cumulative excess fund return two quarters ahead.
$CRET_{0,3}$	Cumulative excess fund return three quarters ahead.
$CRET_{0,4}$	Cumulative excess fund return four quarters ahead.
Earnings Surprise	Difference between reported quarterly earnings and the most recent consensus analyst forecast, scaled by stock price.
Expense Ratio	Ratio of total annual expenses to average net assets.
Firm Age	Number of months since the firm's first appearance in CRSP.
Free Cash Flow/ME	Free cash flow scaled by market equity.
Fund Age	Number of quarters since the fund's inception.
Fund Flow	Net quarterly flow into the fund, scaled by total net assets.
Gross Profit/ME	Gross profits scaled by market equity.
Has Sales Restrictions	Indicator equal to one if the fund imposes restrictions on sales (e.g., redemption fees), zero otherwise.
Income Yield	Annualized dividend yield of the fund.
Management Fee	Annual management fee charged by the fund.
Manager # Funds	Average number of funds simultaneously managed by a given manager.
Manager # Styles	Average number of distinct investment styles overseen by a manager.
Manager Ownership	Average dollar ownership of fund managers in the corresponding fund.
Manager Tenure (Max)	Maximum tenure, measured in quarters, among all managers of a given fund.
Manager Tenure (Mean)	Average tenure of managers in a given fund, measured in quarters.
Market Equity	Market value of the firm's equity (in millions of dollars).
Momentum 1–6 Months	Cumulative stock return from months 1 through 6 prior to portfolio formation.
Momentum 1–12 Months	Cumulative stock return from months 1 through 12 prior to portfolio formation.
MV Rank	Rank of a security's market value within a fund's portfolio; higher values indicate larger relative positions.

Variable	Description
Naive Prediction Precision	Fraction of correct predictions from a naive benchmark model.
NAV/52 week high NAV	Ratio of the current net asset value (NAV) to the highest NAV in the prior 52 weeks.
Net Equity Payout/ME	Net equity payouts (repurchases minus issuances plus dividends) scaled by market equity.
Number Funds	Number of distinct funds holding the security in a given quarter.
Number of Managers	Count of individual portfolio managers responsible for managing the fund.
Open to Investment	Indicator equal to one if the fund is open to new investors, zero otherwise.
Prediction Precision	Fraction of correct predictions produced by the model (fund or security level).
R&D/Sales	Research and development expenditures scaled by sales.
Sales scaled/BE	Sales scaled by book equity.
Short Term Reversal	Past one-month stock return (negative values indicate recent losers).
Total Debt/ME	Ratio of total debt to market equity.
Total Net Asset	Total market value of a fund's assets (in millions of dollars).
Turnover Ratio	Proportion of a fund's portfolio replaced during the year.
Within Category Competition	Number of other funds operating in the same Morningstar investment category in a given quarter.
Within Management Company Competition	Number of other funds managed by the same management company during the quarter.

Variable	Description
Fund-level signals	
Lagged Fund Returns	Prior-quarter and multi-quarter returns of the fund portfolio.
Lagged Fund Flows	Net cash inflow/outflow at the fund level, aggregated to the quarter.
Lagged Fund Characteristics	Size (e.g., AUM), value tilt, and momentum tilt of the fund, constructed from holdings or reported style metrics.
Security-level signals	
Lagged Security Characteristics	Equity size (e.g., log market capitalization), value (e.g., B/M ratio), and momentum (past returns over standard lookbacks).
Risk/Factor Exposures	Predictors from a curated factor list (market, size, value, profitability, investment, short-term reversal), merged at the permno–month level and aligned to quarters.
Peer/Category behavior	
Category Activity Rates	Share-increase, share-decrease, and no-change rates for the fund's Morningstar category (plus their lags), summarizing peer managers' trading behavior within the style bucket.
Macroeconomic backdrop	
Lagged Macroeconomic Factors	Level and slope of the yield curve and credit conditions, including term spreads (e.g., 5y–1y, 10y–3m), default spreads (e.g., BAA–Treasury), and related short/long yields (real and nominal).
Position & within-fund context	
Market Value and Shares	Current market value (mv) and lagged shares (sh) up to six quarters.
Weight Dynamics	Lagged and pass-through values used to compute portfolio weight changes.
Rank Inside Fund (id)	Size rank of the security within the fund each quarter (1 = largest).
Padding Indicator (mask)	Flags placeholder rows created to keep the panel rectangular.

B Theoretical Framework: Extension

B.1 Additional Predictions on Predictability

Let Pred_i denote the predictability of fund i's trades by the AI model. In the model,

$$\frac{\partial \mathrm{Pred}_i}{\partial e_i} < 0,$$

because non-routine effort e_i generates novel trades not captured by historical patterns. We now derive additional predictions linking predictability to fund and manager characteristics.

Prediction 1 (Fund Size and Predictability)**.** *Larger funds possess more resources, specialized teams, and access to superior information and technology. In the model, this increases the effective efficiency of non-routine effort:*

$$z_t = z_t(S), \qquad \frac{\partial z_t}{\partial S} > 0.$$

Since e_t^* is increasing in z_t, we obtain

$$\frac{\partial e_t^*}{\partial S} > 0 \quad \Rightarrow \quad \frac{\partial \mathrm{Pred}_i}{\partial S} < 0.$$

Larger funds exert more non-routine effort and exhibit less predictable trading.

Prediction 2 (Manager Tenure and Predictability)**.** *Early-career managers face strong career concerns and signaling incentives. Let the marginal benefit of effort decline with tenure τ,*

$$\frac{\partial MB(e_t)}{\partial \tau} < 0,$$

where tenure can be interpreted as a reduced-form shifter to the payoff from alpha (e.g., declining career concerns or reputation incentives). In equilibrium,

$$\frac{\partial e_t^*}{\partial \tau} < 0 \quad \Rightarrow \quad \frac{\partial Pred_i}{\partial \tau} > 0.$$

Younger managers trade less predictably, long-tenured managers trade more predictably.

Prediction 3 (Fund Competition and Predictability)**.** *Let κ denote the intensity of competition among active funds. Higher competition increases the value of differentiation, raising the effective payoff to alpha:*

$$U_t(e) = B(\kappa)\,\alpha_t(e) - \frac{e}{z_t}, \qquad \frac{\partial B}{\partial \kappa} > 0.$$

Then,

$$\frac{\partial e_t^*}{\partial \kappa} > 0 \quad \Rightarrow \quad \frac{\partial \text{Pred}_i}{\partial \kappa} < 0.$$

Funds in more competitive segments exhibit lower predictability and more novel trading.

Prediction 4 (Manager Ownership and Predictability). *Let ω measure the strength of performance-based incentives (manager ownership, incentive fees). Manager utility becomes:*

$$U_t(e) = \omega\, \alpha_t(e) - \frac{e}{z_t}.$$

Since e_t^* increases in ω,

$$\frac{\partial e_t^*}{\partial \omega} > 0 \quad \Rightarrow \quad \frac{\partial \text{Pred}_i}{\partial \omega} < 0.$$

Managers with higher ownership stake trade less predictably and produce more novel alpha.

Prediction 5 (Number of Managers and Predictability). *A larger management team can bring complementary skills, broader networks, and more diverse information sources. In the model, this raises the effective efficiency or productivity of non-routine effort:*

$$z_t = z_t(M), \qquad \frac{\partial z_t}{\partial M} > 0 \quad \textit{or} \quad k_t = k_t(M),\ \frac{\partial k_t}{\partial M} > 0.$$

Since e_t^* is increasing in both z_t and k_t, we have

$$\frac{\partial e_t^*}{\partial M} > 0 \quad \Rightarrow \quad \frac{\partial \text{Pred}_i}{\partial M} < 0.$$

Funds with more managers allocate more effort to non-routine, judgment-intensive activities and therefore have less predictable trades.

B.2 Social Welfare with Information Production Externalities

So far, welfare has been measured at the fund level: managers generate alpha for their own investors. However, active managers also produce information that improves price efficiency and benefits all market participants. We now extend the analysis to incorporate such information-production externalities.

Let $\phi(e_t)$ denote the contribution of a type-t manager's non-routine effort e_t to market-wide information production (or price efficiency). Assume:

$$\phi'(e_t) > 0, \qquad \phi''(e_t) \leq 0,$$

so that additional non-routine effort raises, but at a decreasing rate, the informativeness of prices.

Let $\Psi(\cdot)$ be the aggregate benefit of information production for the economy, with

$$\Psi'(x) > 0, \qquad \Psi''(x) \leq 0.$$

Total information production is

$$I = \lambda\phi(e_H^*) + (1-\lambda)\phi(e_L^*),$$

where λ is the fraction of high-skill managers.

Define social welfare as:

$$W^{\text{Social}} = W^{\text{Total}} + \Psi(I),$$

where W^{Total} is the total private welfare derived earlier (sum of managers' utilities and investors' net alpha).

Proposition 3 (AI and Social Welfare with Information Externalities)**.** *The introduction of AI-based decomposition increases social welfare:*

$$W^{\textit{Social}}_{\textit{Post-AI}} > W^{\textit{Social}}_{\textit{Pre-AI}}.$$

Proof. From the post-AI separating equilibrium we know:

$$e_t^{\text{Post-AI}} > e_t^{\text{Pre-AI}}, \quad \forall t,$$

since pre-AI pooling suppresses non-routine effort. Let total private welfare be defined as

$$W^{\text{Total}} \equiv \sum_t \big(\phi(e_t) - c(e_t)\big),$$

where $\phi(\cdot)$ is increasing and $c(\cdot)$ is increasing and convex. Define $g(e) \equiv \phi(e) - c(e)$. Under these assumptions, g is increasing over the equilibrium effort range. Therefore,

$$g(e_t^{\text{Post-AI}}) > g(e_t^{\text{Pre-AI}}) \quad \forall t,$$

which implies

$$W^{\text{Total}}_{\text{Post-AI}} > W^{\text{Total}}_{\text{Pre-AI}}.$$

Thus

$$\phi(e_t^{\text{Post-AI}}) > \phi(e_t^{\text{Pre-AI}}), \quad \Rightarrow \quad I_{\text{Post-AI}} > I_{\text{Pre-AI}}.$$

Because $\Psi'(\cdot) > 0$, we have

$$\Psi(I_{\text{Post-AI}}) > \Psi(I_{\text{Pre-AI}}).$$

Combining the increase in total private welfare with the increase in information externalities yields

$$W_{\text{Post-AI}}^{\text{Social}} = W_{\text{Post-AI}}^{\text{Total}} + \Psi(I_{\text{Post-AI}}) > W_{\text{Pre-AI}}^{\text{Total}} + \Psi(I_{\text{Pre-AI}}) = W_{\text{Pre-AI}}^{\text{Social}}.$$

□

Interpretation. AI-based decomposition not only improves the allocation of effort inside funds, it also increases market-wide information production by encouraging non-routine analysis. This improves price efficiency and social welfare beyond the gains captured by fund investors alone. This said, fee-adjustments by managers could partially (to completely) offset this if adopted.

B.3 Welfare Under Heterogeneous Fund Investors

We now allow for heterogeneity among fund investors. Suppose there are two types of investors, and their types are not perfectly observable:

- Type-A investors, who can interpret the AI-based decomposition and condition capital flows and fees on non-routine performance.
- Type-B investors, who cannot fully process the AI signal and continue to respond primarily to coarse activity measures.

Let $\mu \in [0, 1]$ be the mass of type-A investors and $(1 - \mu)$ the mass of type-B investors. Let w_t^A and w_t^B denote equilibrium compensation for a type-t manager from type-A and type-B investors, respectively. Total compensation is:

$$w_t = \mu w_t^A + (1 - \mu) w_t^B.$$

Type-A investors use the AI signal $\tilde{N} = \theta_t^N e_t$ and thus implement the separating wage schedule

$$w_t^A = \alpha_t(e_t^*),$$

as in the baseline post-AI analysis. Type-B investors base compensation on coarser observable statistics, e.g., partial use of T or historical performance, and may not achieve full separation.

Define investor-specific welfare:

$$W^{I,A} = \mu\big(\alpha - w^A\big), \qquad W^{I,B} = (1-\mu)\big(\alpha - w^B\big),$$

and manager welfare as before:

$$W^M = \sum_t \pi_t U_t^*,$$

where π_t is the population share of type t.

Total welfare is:

$$W^{\text{Het}} = W^{I,A} + W^{I,B} + W^M.$$

Proposition 4 (Heterogeneous Investors and Welfare)**.** *As μ increases (a larger share of sophisticated, AI-using investors), the economy converges to the benchmark post-AI separating allocation and total welfare increases:*

$$\frac{\partial W^{Het}}{\partial \mu} > 0.$$

Proof. As μ rises, a greater fraction of capital is allocated by investors who correctly interpret $\tilde{N}$ and implement the separating wage schedule. This increases incentives for non-routine effort, pushing e_t toward the post-AI benchmark e_t^*.

Manager utilities U_t increase because effort moves closer to the privately optimal level, and aggregate alpha becomes more strongly tied to skill. Type-A investor welfare is by construction maximized under correct pricing of skill; as μ rises, the mispricing distortions associated with type-B investors shrink in measure.

Thus total welfare rises monotonically in μ and converges to the post-AI benchmark when $\mu \to 1$. □

Interpretation. The benefits of AI-based decomposition are strongest when investors can understand and act on the information it provides. Heterogeneity in investor sophistication leads to an intermediate regime between pure pre-AI pooling and full post-AI separation.

B.4 Comparative Statics of Welfare

We now summarize how welfare responds to key technological and preference parameters: the routine-task weight a, the CES parameter ρ, effort efficiency z_t, and non-routine productivity θ_t^N.

Recall that welfare for type t is:

$$W_t^* = A\left[C_0 + k_t(e_t^*)^\rho\right]^{1/\rho} - \frac{e_t^*}{z_t}, \qquad k_t = (1-a)(\theta_t^N)^\rho.$$

Effect of a (Importance of Routine Tasks)

As a increases, routine tasks receive more weight in the CES aggregator:

$$k_t = (1-a)(\theta_t^N)^\rho \downarrow .$$

This reduces the marginal return to non-routine effort, leading to lower e_t^* and lower welfare W_t^*.

Prediction 6 (Routine Weight and Welfare)**.** *Welfare is decreasing in the routine task weight a:*

$$\frac{\partial W_t^*}{\partial a} < 0.$$

Environments where routine tasks dominate (high a) are less conducive to welfare-improving non-routine analysis.

Effect of ρ (Elasticity of Substitution)

The parameter ρ determines the elasticity of substitution $\sigma = 1/(1-\rho)$ between routine and non-routine tasks.

- When ρ is higher (tasks more easily substitutable), routine input can more easily replace non-routine input, weakening incentives to exert non-routine effort.
- When ρ is lower (tasks more complementary), non-routine input is more valuable at the margin, raising e_t^* and W_t^*.

Prediction 7 (Complementarity and Welfare). *Welfare is higher when routine and non-routine tasks are stronger complements:*

$$\frac{\partial W_t^*}{\partial \rho} < 0.$$

Lower ρ (stronger complementarity) increases the value of non-routine effort and raises welfare.

Effect of z_t (Effort Efficiency)

Effort efficiency z_t measures how costly it is for type t to supply non-routine effort.

Prediction 8 (Effort Efficiency and Welfare). *Welfare is increasing in z_t:*

$$\frac{\partial W_t^*}{\partial z_t} > 0.$$

Higher z_t reduces the marginal cost of effort, raises optimal effort e_t^*, and increases both alpha and manager utility.

Effect of θ_t^N (Non-Routine Productivity)

Non-routine productivity θ_t^N determines how effectively a manager converts effort into non-routine input $N = \theta_t^N e$.

Since

$$k_t = (1-a)(\theta_t^N)^\rho$$

and e_t^* is increasing in k_t, alpha and welfare both increase in θ_t^N.

Prediction 9 (Non-Routine Productivity and Welfare). *Welfare is increasing in non-routine productivity:*

$$\frac{\partial W_t^*}{\partial \theta_t^N} > 0.$$

Managers who are more productive at non-routine tasks generate higher alpha and greater welfare.

Summary. Welfare is higher when:

- routine tasks are less important (a is low),
- routine and non-routine tasks are more complementary (low ρ),
- managers are more efficient in exerting effort (high z_t),

- and more productive in non-routine tasks (high θ_t^N).

These comparative statics provide clear empirical guidance on where AI-based decomposition is likely to yield the largest welfare gains.