Do Trades and Holdings of Market Participants Contain Information About Stocks? A Machine-Learning Approach

Victor DeMiguel Li Guo Bo Sang Zhe Zhang*

March 27, 2025

Abstract

We use machine learning to capture nonlinearities and interactions in the relation between trades and holdings of multiple market participants and future stock returns. Our predictor yields a long-short portfolio with significant out-of-sample alpha, forecasts firm fundamentals, and assigns stocks on the right side of most anomalies. Predictability is stronger for smaller or illiquid stocks and stocks with lower analyst coverage or higher idiosyncratic volatility. A factor model based on our predictor achieves higher Sharpe ratio than existing models. Our findings suggest that incorporating nonlinear interactions between trades and holdings of various participants reveals valuable information for price discovery.

`JEL classification:` G10, G11, G23.
`Keywords:` Return predictability; Machine learning; Nonlinearities and interactions; Institutional investors; Trades and holdings; Anomalies.

*DeMiguel: London Business School, avmiguel@london.edu. Guo: School of Economics, Fudan University; Shanghai Institute of International Finance and Economics, guo_li@fudan.edu.cn. Sang: School of Business, University of Bristol, bo.sang@bristol.ac.uk. Zhang: Lee Kong Chian School of Business, Singapore Management University, joezhang@smu.edu.sg. We are grateful for valuable comments from Zareei Abalfazl, Suleyman Basak, Ivan Blanco, Maxime Bonelli, Svetlana Bryzgalova, Manav Chaudhary (discussant), Mena El Hefnawy, Julian Franks, Francisco Gomes, Marco Grotteria, Ron Kaniel, Paul Karehnke, Bryan Kelly (discussant), Antonia Kirilova, Igor Kuznetsov, Martin Lettau, Narayan Naik, Tina Oreski, Jean Pauphillet, Jeffrey Pontiff, Alvaro Remesal, Matthew Ringgenberg, Andre P. Santos, Dacheng Xiu, Hanbin Yang, Qian Yang (discussant), Ruixun Zhang, and seminar participants at the Utah Winter Finance Conference, Northern Finance Association Annual Meeting, Midwest Finance Association Annual Meeting, CUNEF University, ESCP Business School, London Business School, and Peking University.

1 Introduction

If market participants possess value-relevant information about stocks, a signal that combines their collective trades and holdings should predict future stock returns. Although there is an extensive literature on the informativeness of the trades and holdings of various market participants, most studies focus on one particular type of investor, such as mutual funds, hedge funds, short sellers, or the firms themselves.[1] Several papers also study the role of interactions between pairs of investors—through learning or competition—in the informational content of their trades.[2] However, these interactions may extend beyond pairs of investors. Also, different investors may hold various types of information and be informed at different times. Even the trades of investors who make systematically poor bets can reveal information about future stock returns. In this paper, we examine the combined informational role of multiple types of market participants. Specifically, we use machine learning to construct a return predictor that aggregates the information in the trades and holdings of various market participants and assess its ability to predict stock returns out of sample.

Like McLean et al. (2025), we consider nine market participants, including six types of institutional investors (mutual funds, hedge funds, banks, insurance companies, wealth management firms, and other institutions) as well as firms, short sellers, and retail investors. We aggregate trade and holdings by type of market participant so that we can identify the effect of interactions between different types of investors. This also facilitates the comparison of our results with those in the literature on the informational content of market participant trades. Moreover, following Gompers and Metrick (2001) and Yan and Zhang (2009), we decompose current quarterly holdings into lagged quarterly holdings plus current quarterly trades. If a specific type of market participant is informed about a change in the expected return of a stock,

[1] For studies on the relation between mutual fund *holdings* and performance see, for example, Daniel et al. (1997) and Wermers (2000), and for studies focusing on the *trades* of specific types of investors see, for instance, Diether et al. (2009), Aggarwal and Jorion (2010), Boehmer et al. (2010), Baker et al. (2010), Kaniel et al. (2012), Kelley and Tetlock (2013), Pástor et al. (2017), and Boehmer et al. (2021). McLean et al. (2025) study the relation between the trades of nine market participants and the return forecasts obtained using firm characteristics, by focusing on the individual effect of each type of participant.

[2] For example, Sias and Whidbee (2010) study interactions between insiders and institutional investors, Massa et al. (2015) between insiders and shortsellers, and Jiao et al. (2016) between shortsellers and the long positions of hedge funds.

its current quarterly trade in that stock should strongly predict future returns. Conversely, to the extent that lagged holdings proxy for current holdings, the predictability of lagged holdings could reflect the effect of persistent demand on future stock returns. Persistent holdings might also reflect information about the stock long-term performance.

To construct the return predictor, we use several nonlinear machine-learning methods that can capture nonlinearities and interactions in the relation between the trades and holdings of multiple market participants and future stock returns. Following Gu et al. (2020), we consider random forest, gradient boosting regression trees, and artificial neural networks.[3] For comparison, we also employ six linear models: ordinary least squares plus five linear machine-learning methods, including principal component regression, partial least squares, adaptive least absolute shrinkage and selection operator, ridge regression, and elastic net. These linear machine-learning methods have been recently used in asset pricing (Chinco et al., 2019; Gu et al., 2020; Dong et al., 2022; Leippold et al., 2022) and can mitigate the overfitting resulting from multicollinearity among predictors.

We use a rolling-window approach to evaluate the out-of-sample performance of the linear and nonlinear predictors. For each model and five-year estimation window, we use stock-month panel data, including both trades and lagged holdings of the nine market participants, to predict next-month excess returns. Then, for each stock we average the predicted returns across the six linear models to construct a linear composite predictor (LCP) and across the nonlinear models to construct a nonlinear composite predictor (NLCP).[4] Each month, we sort our sample of stocks into decile portfolios based on the LCP or NLCP, construct long-short portfolios that buy the high and short the low decile, and evaluate their out-of-sample monthly alphas with respect to the five Fama and French (2015) factors plus the momentum factor (FF5+MOM) and the five Hou et al. (2021) factors plus the momentum factor (q5+MOM).

Although our work focuses on the information in the combined trades of multiple market participants, as a benchmark we also evaluate the performance of the univariate decile portfo-

[3]The literature has documented the presence of nonlinear relations between predictors and future asset returns (Kelly et al., 2019; Gu et al., 2020; Bianchi et al., 2021; Binsbergen et al., 2022; Leippold et al., 2022). Given the complex interactions among multiple types of market participants, it is reasonable to expect similar nonlinear relations between the trades and holdings of different market participants and future stock returns.

[4]To simplify the exposition, in the main body of the manuscript we discuss the performance of only the linear and nonlinear *composite* predictors. However, Table B.3 in Appendix B shows that our findings are robust to considering the individual performance of each of the 12 linear and nonlinear models.

lios constructed by sorting stocks based on the trades (or holdings) of a single type of investor. We find that the univariate long-short portfolios obtained from the trades of several types of investors have significant alphas. For instance, while the long-short portfolios based on the trades of hedge funds, other institutions, or retail traders have significantly positive alpha, those based on the trades of banks or mutual funds have significantly negative alpha. We also find that the only univariate long-short portfolio based on holdings that achieves a significant alpha is that based on short seller holdings, consistent with the literature that documents the informational content of short interest (Desai et al., 2002; Asquith et al., 2005; Diether et al., 2009; Karpoff and Lou, 2010; Boehmer et al., 2010).

The results for the linear composite predictor (LCP) show that there are benefits from linearly combining the information in the trades and holdings of multiple market participants. In particular, we find that a long-short portfolio obtained by buying the high and shorting the low decile portfolio of stocks sorted by the LCP has significant monthly alphas of 1.78% and 1.68% with respect to the FF5+MOM and q5+MOM models. Moreover, these long-short portfolio alphas are driven by both the long and short legs, each with significant but oppositely signed alphas. The long-short alphas are also significantly higher than those of the univariate long-short portfolios, with the exception of the portfolio based on short seller holdings, for which although the alpha of the LCP long-short portfolio is more than 10% higher, the difference is not significant.

The results for the nonlinear composite predictor (NLCP) reveal even more significant gains from capturing nonlinearities and interactions in the relation between the trades and holdings of multiple market participants and future stock returns. For instance, the long-short portfolio of stocks based on the NLCP has monthly alphas of 2.74% and 2.60% with respect to the FF5+MOM and q5+MOM models, which are 96 and 92 basis points higher than those of the long-short portfolio based on the LCP, with the differences being significant at the 1% confidence level. The NLCP long-short portfolio significantly outperforms also the univariate long-short portfolios, including that based on short seller holdings.

We find that the NLCP long-short portfolios continue to generate significantly positive alpha for up to 16 months after portfolio formation. This suggests that the return predictability is not driven by price pressure and that there is information about stocks in the collective

trades and holdings of market participants. The performance of the NLCP long-short portfolio is robust to considering transaction costs of 60 basis points. Fama-Macbeth regressions show that the positive relation between next-month returns and the NLCP remains statistically significant after controlling for common firm characteristics. The effect in Fama-MacBeth regressions is also economically significant, with a spread between the expected monthly return of stocks in the high and low NLCP deciles of 1.54%.

To assess the relative significance of trades and holdings for the composite predictors, we construct versions of the LCP and NLCP based on only trades or only holdings. When we include these predictors individually in Fama-MacBeth regressions, we find that both trading- and holdings-based LCPs and NLCPs significantly predict future stock returns. However, when we include all predictors in a single Fama-MacBeth regression, only the NLCPs remain significant. Specifically, the time-series mean coefficient estimate for the trading-based NLCP is 0.32, significant at the 5% level, and that for the holdings-based NLCP is 0.46, significant at the 1% level. In contrast, the coefficients for the trading-based and holdings-based LCPs are not significant after controlling for the NLCPs based on trades and holdings.

Is the predicting power of the machine-learning models driven by the (nonlinear) interactions among the trades and holdings of all types of market participants, or primarily by a few? To answer this question, we investigate the importance of the different predicting variables and their interactions for the random forest model using the SHAP values of Lundberg and Lee (2017), which measure the contribution of each variable to the prediction for each observation. We find that no single prediting variable drives the predictions of random forest, with 13 out of the 16 predicting variables having an average absolute SHAP value above 2.5 basis points. While certain investors—such as mutual funds, insurance companies, and wealth management firms—are more important, no single type of investors dominates the predictions. Consistent with the results of the Fama-MacBeth regressions, we find that both trades and holdings are important. We also find that random forest exploits the interactions between the lagged holdings of different types of investors as well as between the trades of some investors and the lagged holdings of others. Short seller holdings and other institutional holdings, in particular, are instrumental for the interactions revealed by random forest, with one or the other appearing in 16 out of the 20 most important interactions.

If the NLCP reflects valuable information that is not yet absorbed by the market, we would expect its ability to predict returns to be stronger for firms with a more uncertain information environment, where opportunities for informed investors are greater, but arbitrage costs are also relatively high. Consistent with this hypothesis, we find that the predictive power of the NLCP is stronger for stocks with lower market capitalization, higher idiosyncratic volatility, lower liquidity, lower analyst coverage, and higher analyst disagreement, although the effects are similar regardless of firm age.

If the NLCP contains value-relevant information, we would expect it to predict also future firm fundamentals. The existing literature provides mixed evidence on whether the trades of various types of market participants predict firm fundamentals (Baker et al., 2010; Boehmer et al., 2020). We find that the NLCP forecasts key fundamentals for the next quarter, including gross margin (GM), return on assets (ROA), return of equity (ROE), and cash flow (CF), even after controlling for recent fundamentals and common firm characteristics. The results for the LCP are qualitatively similar. Given that the NLCP can predict firm fundamentals, it should also predict earnings announcement returns, which capture the combined effect of reported fundamentals on stock returns. Indeed, we find that the spread between the expected earnings announcement return of stocks in the high and low NLCP deciles is 0.56%, significant at the 1% level. The ability of LCP and NLCP to predict future firm fundamentals and earnings announcement returns provides support to the hypothesis that the collective trades and holdings of market participants contain information about stocks.

Prior studies suggest that market participants such as institutional investors trade on the wrong side of anomalies; see, for instance, Hirshleifer et al. (2011), Edelen et al. (2016), Patton and Weller (2020), and McLean et al. (2025). We sort our sample stocks into decile portfolios based on the LCP and NLCP and examine firm-characteristic values of these portfolios with respect to each anomaly, as well as the characteristic spreads between the high and low deciles. Out of the 102 stock return anomalies documented in Green et al. (2017), we select the 12 anomalies whose returns are statistically significant for our sample period. Both the LCP and NLCP assign stocks on the right side of the 12 anomalies, with the exception of short-term reversal and analyst coverage.

Given that the LCP and NLCP are aligned with most anomalies, it is interesting to

examine whether a factor model based on the LCP or NLCP can explain the cross-section of anomaly returns. To do this, we construct an LCP or NLCP factor by taking a long position in the high decile and a short position in the low decile of stocks sorted by the LCP or NLCP. We then form a two-factor model with the market factor and the LCP or NLCP factor. We find that the average absolute alpha for the 12 significant anomalies with respect to the LCP and NLCP factor models is 0.46 and 0.50, with average absolute t-statistics of 1.19 and 1.16, comparable to the FF5+MOM and q5+MOM models. That is, the two-factor model with only the market and our LCP or NLCP factor performs similar to the prominent FF5+MOM and q5+MOM models in terms of its ability to explain anomaly returns. This is a notable result because the LCP and NLCP factors are constructed using only the trades and holdings of market participants, ignoring any anomaly information.

Barillas and Shanken (2018) show that an economically meaningful criterion to compare factor models is the Sharpe ratio generated by their factors. We find that the NLCP factor model produces an annualized Sharpe ratio of 3.09, which is significantly higher than those generated by the FF5+MOM and q5+MOM models (1.13 and 1.71). Thus, the Sharpe ratio results confirm that the NLCP factor model contains information that helps to span the *overall* investment opportunity set better than existing factor models. In other words, the NLCP identifies information important for asset pricing that is not captured by either existing factor models or the set of anomalies we consider.

Our manuscript is closely related to the literature that studies whether market participant trades contain information about future stock returns. Most of the existing literature focuses on a particular type of investors (Diether et al., 2009; Aggarwal and Jorion, 2010; Boehmer et al., 2010; Baker et al., 2010; Kaniel et al., 2012; Kelley and Tetlock, 2013; Cao et al., 2018; Boehmer et al., 2021) or the interactions between pairs of investors (Sias and Whidbee, 2010; Massa et al., 2015; Jiao et al., 2016). Da et al. (2024) find that the net trading by outside arbitrageurs (hedge funds and short sellers) and that by firm insiders both independently predict stock returns, and the two types of investors hold different types of information. Gao and Zhang (2023) adopt a Bayesian approach within the intermediary asset-pricing theory framework (He and Krishnamurthy, 2013) and show that using a prior based on information from institutional holdings significantly improves return predictability. McLean et al. (2025)

study the trades of nine market participants and their relation with anomalies and future returns, by focusing on the individual effect of each type of participant. A distinctive feature of our work is that we highlight the importance of considering the trades and holdings of multiple types of investors *jointly* and using machine learning to capture nonlinearities and interactions in their relation to future stock returns.

Our work is also related to the growing literature that uses machine learning in asset pricing and investment. Most of the existing work focuses on using machine learning to study the cross section and time series of stock returns, using firm characteristics (Gu et al., 2020; Kozak et al., 2020; Bryzgalova et al., 2021; Chatigny et al., 2022; Leippold et al., 2022). Several recent papers use machine learning to study mutual-fund performance (Li and Rossi, 2021; DeMiguel et al., 2023; Kaniel et al., 2023). We show that machine learning successfully extracts value-relevant information from the combined trades and holdings of multiple market participants, and the predictor based on these signals strongly predicts future stock returns after controlling for historical returns and firm characteristics.

Koijen and Yogo (2019) develop an equilibrium asset-pricing model based on the asset demand of market participants. They use individual institutional holdings to study the model's implications for asset market movements and volatility. Their research has spurred a growing literature on demand system asset pricing (Egan et al., 2024; Davis et al., 2024; Noh et al., 2024; Huebner, 2024; Da, 2022). Importantly, the asset demands or demand shocks in these models are exogenous and unrelated to firm fundamentals, with demand elasticity determining contemporaneous equilibrium prices. In contrast, we show that the NLCP, which captures the (nonlinear) interactions of the trades and holdings of multiple market participants, contains information about firm fundamentals and predicts future stock returns. We also aggregate trade and holdings data at the market-participant type level, which allows us to examine which types of market participants have information about stocks and how they interact in financial markets.

The remainder of this manuscript is organized as follows. Section 2 describes the data, methodology, and descriptive statistics. Section 3 evaluates the performance of the composite return predictors. Section 4 studies how the predictability of the composite predictors varies with the firm's information environment, and explores whether the composite predic-

tors forecast firm fundamentals and earnings announcement returns. Section 5 examines the relation between composite predictors and anomalies and the performance of factor models constructed from these predictors. Section 6 concludes.

2 Data, methodology, and descriptive statistics

This section discusses our data and methodology. Section 2.1 describes the data sources and how we construct the predicting variables for the machine-learning models. Section 2.2 gives an overview of the machine-learning methods we employ. Section 2.3 explains how we evaluate the out-of-sample performance of the linear and nonlinear composite predictors. Finally, Section 2.4 provides descriptive statistics.

2.1 Data sources and predicting variables

The predicting variables we use to train the machine-learning models are the quarterly trades and lagged quarterly holdings of the nine market participants considered by McLean et al. (2025), which include six types of institutional investors (mutual funds, hedge funds, banks, insurance companies, wealth management firms, and other institutions) as well as firms, short sellers, and retail investors.[5] For firms and retail investors, we consider only their trades because we lack relevant information about their holdings. Thus, we have 16 predicting variables, including nine trading and seven holding variables.

We obtain institutional holdings data from *Thomson/Refinitiv S12* and *13F* filings to construct the quarterly holdings of the aforementioned six types of institutional investors.[6] We normalize quarterly institutional holdings by dividing them by the total number of shares outstanding. We estimate the trades of institutional investors as the change in their standardized holdings from the previous quarter.

[5]Brian Bushee provides a classification of institutional investors consisent with that of McLean et al. (2025) at https://accounting-faculty.wharton.upenn.edu/bushee/.

[6]Consistent with McLean et al. (2025), we associate holdings data with the date at which the investor held the positions, rather than the date at which they were reported, which could be up to 45 days later. This is because our research question is whether investor trades contain information at the time they are executed. Nevertheless, Table 6 shows that the NLCP portfolios continue producing significantly positive alpha up to 16 months after portfolio formation, which suggests that one could implement profitable investment strategies in real time based on publicly available holdings data.

To measure retail trades, we use the approach proposed by Boehmer et al., 2021 to identify daily off-exchange marketable orders from the *Trade and Quote* (*TAQ*) dataset and apply the algorithm recently proposed by Barber et al. (2024) to sign each trade. We then estimate aggregate retail trades using the monthly retail order imbalance.[7]

We construct the rest of the predicting variables using the same methodology as McLean et al. (2025). We calculate short interest ratio as quarter-end short interest from *Compustat*, divided by the number of shares outstanding. Following McLean et al. (2025), we sign the short interest ratio so that higher values correspond to less shorting. We term the signed short interest ratio as short seller holdings to be in line with the other holdings variables. We then estimate short seller trades as the quarterly change in signed short interest ratio.

We estimate firm trades as the difference between quarter-end share repurchases and share issues, divided by the number of outstanding shares. We sign this variable so that decreases (increases) in shares outstanding correspond to positive (negative) values of firm trading. Panel A of Table 1 contains the definition of the 16 predicting variables and Appendix A.1 gives further details of their construction.

[Insert Table 1 here]

To construct our sample from the *CRSP* dataset, we select common stocks (with share codes of 10 or 11) listed on the NYSE, AMEX, or Nasdaq (with exchange codes of 1, 2, or 3), excluding stocks with prices below \$1 as well as those of financial and utility firms. We also acquire accounting variables from *Compustat* and analyst forecast and recommendation data from the *Institutional Brokers' Estimate System (I/B/E/S)*. Following the methodology outlined by Green et al. (2017), we replicate their 102 stock anomalies. Our sample period spans January 2008 to December 2020.

[7] McLean et al. (2025) use the same dataset and approach to construct retail trades. However, in addition to the retail order imbalance, they also consider a measure of retail trades normalized using the number of shares outstanding. We focus on the monthly retail order imbalance to be consistent with Boehmer et al. (2021) and Barber et al. (2024). While we term the resulting measure "retail investor trades," our findings do not depend solely on its capturing the trades of retail investors. It may also reflect the collective effect of trades by investors, including possibly some institutional investors, who benefit from price discounts in the off-exchange market. Moreover, Table B.4 in Appendix B shows that our findings are robust to excluding retail investor trades from our sample.

2.2 Machine-learning methods

We use the quarterly trades and lagged quarterly holdings constructed in Section 2.1 as predicting variables for the linear and nonlinear machine-learning models. We closely follow the approach in existing papers that apply machine learning to stock analysis (Gu et al., 2020; Kozak et al., 2020).

We consider six linear models: *ordinary least squares* (OLS) and five linear machine-learning methods. In particular, we consider the two dimension-reduction methods used by Gu et al. (2020): *principle components regression* (PCR) and *partial least squares* (PLS). We also consider three popular regularization methods described in Hastie et al. (2009): the adaptive version of *least absolute shrinkage and selection operator* (ALasso), *ridge regression* (Ridge), and *elastic net* (ENet).

Gu et al. (2020) show that accounting for nonlinearities and interactions substantially improves the accuracy of return predictive models. Similarly, DeMiguel et al. (2023) find that nonlinear machine-learning models help to select a portfolio of actively managed mutual funds with significant out-of-sample alphas. Following these papers, we consider three sets of nonlinear machine-learning models: *gradient boosted regression trees* (GBRT), *random forest* (RF), and *artificial neural networks* (ANN). GBRT integrates multiway predictor interactions and utilizes a boosting algorithm that recursively combines forecasts from regression trees to enhance performance. Similarly, RF aggregates forecasts from regression trees using a bootstrap aggregation method, which reduces correlation among the regression trees and increases prediction stability. As a more complex machine-learning model, ANN introduces "feed-forward" networks that increase flexibility and capture intricate interactions among predictors. We consider neural networks with one to four hidden layers linking the predicting and dependent variables (ANN1, ANN2, ANN3, ANN4). Thus, we consider at total of six nonlinear machine-learning models.

2.3 Out-of-sample evaluation of composite predictors

We evaluate the out-of-sample performance of the linear and nonlinear composite predictors using a rolling-window approach. Following Feng et al. (2020) and Gu et al. (2020),

we first normalize the 16 predicting variables described in Section 2.1 (the quarterly trades and lagged quarterly holdings) by applying a cross-sectional transformation so that they are uniformly distributed between minus one and one.[8] Then, for the first estimation window spanning January 2008 to December 2012, we train each linear and nonlinear prediction model using stock-month structured panel data to predict next month's stock excess return, using the normalized variables as predictors. Next, we feed the normalized predicting variables for December 2012 into each of the trained models to forecast the excess return of each stock in January 2013. We then roll the estimation window one month forward and repeat this process.

At the end of this procedure, we have stock return forecasts for the 12 models for each month in our out-of-sample period from January 2013 to December 2020. To construct the linear composite return predictor (LCP) for each stock and out-of-sample month, we compute the average out-of-sample predictor across the six linear models: OLS, PCR, PLS, ALasso, Ridge, and ENet. Similarly, the nonlinear composite return predictor (NLCP) is the average across the six nonlinear model predictors: GBRT, RF, ANN1, ANN2, ANN3, ANN4. Panel B of Table 1 gives a definition of the linear and nonlinear composite predictors and Appendix A.2 gives a detailed description of the 12 predicting models we consider.

To simplify the exposition, in the main body of the manuscript, we discuss the performance of only the linear and nonlinear *composite* predictors. However, Table B.3 in Appendix B reports the individual performance of each of the 12 linear and nonlinear models. We find that the six linear models perform similarly, with the exception of principal component regression, which underperforms.[9] In addition, we find that the six nonlinear machine-learning models also perform similarly to each other and outperform the linear models.

[8] We replace missing values of the normalized predicting variables with their cross-sectional mean (zero), provided that there is data for at least one predicting variable. If there is no data for any predicting variable, then we drop the observation from the panel.

[9] To understand why the OLS model performs on par with the linear machine-learning methods, note that we consider a relatively small number of 16 predicting variables, so OLS is unlikely to suffer from overfitting, and thus the regularization term in the linear machine-learning models is unlikely to have a substantial effect. The underperformance of the principal component regression approach is explained by our considering a small number of only up to three principal components.

2.4 Descriptive statistics

Table 2 provides summary statistics (mean, median, standard deviation, and 10th and 90th percentiles) estimated at the stock-month panel-data level for the holdings and trades of the nine types of market participants, and for the linear and nonlinear composite return predictors.

[Insert Table 2 here]

Panel A of Table 2 shows that mutual funds and other institutions (non-categorized 13F institutions) together hold around 40% of the shares outstanding, consistent with McLean et al. (2025). In contrast, the holdings from insurance companies and wealth management firms are much smaller, around 2–3% of shares outstanding each. The mean short seller holdings (short interest) is substantial at −15.46% of shares outstanding. The 10th and 90th percentiles show that there is wide variation in holdings across our panel data. For instance, mutual-fund holdings range from close to zero to just below 40%. Panel B of Table 2 shows that there is also variation across our panel data in the trades of the different types of investors. In particular, we observe that the 10th percentile of the trading variables is negative and the 90th percentile is positive for every type of investor. Panel C of Table 2 shows that the average linear composite return predictor (LCP) is 0.63%, while the average nonlinear composite return predictor (NLCP) is 1.22%. The NLCP shows a greater variation compared to the LCP, with a standard deviation of 0.83%.

3 Out-of-sample performance of return predictors

In this section, we evaluate the out-of-sample performance of the portfolios of stocks sorted by the return predictions obtained from the trades and holdings of the nine types of market participants. As a benchmark, Section 3.1 reports the performance of the univariate portfolios formed by sorting stocks based on the trades (or holdings) of a single type of market participants. Section 3.2 reports the performance of the portfolios based on the linear and nonlinear composite predictors obtained from the combined trades and holdings of all market participants. Section 3.3 uses Fama and MacBeth (1973) regressions to evaluate the

performance of the composite predictors after controlling for common firm characteristics, and Section 3.4 to examine the significance of trades and holdings for the performance of the composite predictors. Finally, Section 3.5 studies the importance of the 16 predicting variables and their interactions for the predictions of the nonlinear machine-learning models.

3.1 Univariate portfolios

Although our work focuses on the information in the *combined* trades and holdings of multiple market participants, as a benchmark we also evaluate the performance of the univariate portfolios obtained by sorting stocks based on the trades (or holdings) of a single type of market participant. Panels A and B of Table 3 report the performance of the univariate portfolios constructed using trades and lagged holdings, respectively. For each panel, the first column reports the name of the sorting variable, the second to seventh columns report the alpha of the low, second, third, eight, ninth, and high decile portfolios, respectively, and the eighth column reports the alpha for the long-short portfolio, which goes long stocks in the high decile and short stocks in the low decile. At the end of each month, we rank all sample stocks based on either the trades or lagged holdings of each type of market participant and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile portfolio, as well as the long-short portfolio, and report their alphas with respect to the five Fama and French (2015) and momentum factors. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively.[10] To facilitate the comparison, we evaluate performance in the period from 2013 to 2020, which coincides with the out-of-sample period for the evaluation of the composite return predictors.

[Insert Table 3 here]

Panel A of Table 3 shows that the univariate long-short portfolios obtained from the trades of several types of investors achieve significant, although not always positive, alphas. For instance, the alphas of the long-short portfolios based on trades are significant at the 1%

[10] We use the R-package "NeweyWest" to select the bandwidth using the approach of Newey and West (1994).

level for banks and other 13F institutions, at the 5% level for hedge funds, mutual funds, and retail investors, and at the 10% level for insurance companies and wealth management firms. However, the sign of the alphas demonstrates that only the trades of hedge funds, other 13F institutions, and retail investors (those from off-exchange markets) predict abnormal returns on the right side. This is consistent with the existing literature that highlights the predictive power of hedge fund trades (Kosowski et al., 2007; Aggarwal and Jorion, 2010; Cao et al., 2018) and that of retail trades that goes beyond mere liquidity provision (Kaniel et al., 2012; Kelley and Tetlock, 2013; Boehmer et al., 2021).[11] Although the trades of other (uncategorized) 13F institutions have not received much attention in the existing literature, McLean et al. (2025) suggest that some of these trades may involve proprietary trading, which could explain their ability to predict abnormal returns on the right side. On the other hand, the monthly alpha of the long-short portfolio obtained from mutual fund trades is -0.39%, negative and significant at the 5% level, suggesting that mutual fund trades perform poorly. This is consistent with evidence in the existing literature that actively managed mutual funds (on average) underperform their benchmarks after fees (Jensen, 1968; Elton et al., 1996; Fama and French, 2010; Ferreira et al., 2013). Finally, Panel A shows that the trades of insurance companies and wealth management firms also negatively predict abnormal returns significantly at the 10% level.[12]

Panel B of Table 3 reports the alphas of univariate portfolios of stocks sorted by lagged holdings. The only univariate long-short portfolio based on holdings that achieves a significant alpha is that based on short seller holdings, which has a substantial monthly alpha of 1.51%. As explained in Section 2.1, we use short interest as a proxy for short seller holdings and we change its sign so that a higher value corresponds to less shorting. The high decile of stocks

[11] In contrast, earlier work showed that retail investors as a group can be uninformed and irrational (Barber and Odean, 2000; Barber et al., 2009)

[12] The results in Panel A of Table 3 are consistent with some of those in table 8 of McLean et al. (2025). For instance, both tables show that bank trades negatively predict returns and retail trades (retail order imbalance) positively predict returns. However, there are also some differences. For example, while McLean et al. (2025) find that short seller trades and firm trades positively predict returns, we find that the trades of hedge funds and other institutions positively predict abnormal returns. Also, while McLean et al. (2025) find that institutional trades (other than those of banks) generally do not predict returns individually, we find that the trades of mutual funds, insurance companies, and wealth management firms negatively predict abnormal returns. The differences between the results in McLean et al. (2025) and ours are explained by two main factors. First, while the sample data of McLean et al. (2025) spans 2006 to 2021, ours spans 2013 to 2020. Second, while McLean et al. (2025) aggregate trades annually, we aggregate them quarterly.

sorted by short seller holdings has a significant monthly alpha of 1.04% and the low decile a significantly negative alpha of -0.47%. These findings align with the existing literature, which demonstrates that low levels of short interest signal positive information about future stock returns, while high levels indicate negative information (Desai et al., 2002; Asquith et al., 2005; Diether et al., 2009; Karpoff and Lou, 2010; Boehmer et al., 2010).[13]

3.2 Composite return predictors

In this section, we evaluate the out-of-sample performance of the portfolios of stocks sorted by the linear (LCP) and nonlinear (NLCP) composite predictors and study the impact of transaction costs on their performance. We also examine the persistence over time of the performance of the NLCP-based portfolios.

Table 4 reports the alpha of the portfolios of stocks sorted by the LCP and NLCP. The first column reports the label for each decile portfolio and the long-short portfolio, the second and third columns report the alpha of each portfolio based on the LCP with respect to the five Fama and French (2015) and momentum factors (FF5+MOM) and the five Hou et al. (2021) and momentum factors (q5+MOM), the fourth and fifth columns for the portfolios based on the NLCP, and the sixth and seventh columns for the difference between the returns of the portfolios based on NLCP and LCP (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile portfolio, as well as on the long-short portfolio, and report their alphas. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

[13]Tables B.1 and B.2 of Appendix B report the performance of the univariate portfolios in terms of excess returns and alpha with respect to the five Hou et al. (2021) and momentum factors, respectively, and lead to similar conclusions. Consistent with panel B of table 10 in McLean et al. (2025), Table B.1 of Appendix B shows that the monthly *excess return* of the univariate long-short portfolio based on short seller holdings is 0.71%, substantially lower than its monthly alpha of 1.51%. The difference between the excess return and alpha of this portfolio is explained by its negative exposure ($\beta = -0.61$) to the market factor, which earns a substantial average excess return of 1.26% in our sample.

[Insert Table 4 here]

The portfolios based on the linear composite predictor (LCP) achieve good out-of-sample performance. For instance, the second and third columns of Table 4 show that the alphas of the LCP decile portfolios increase, although not always monotonically, from the low (L) to the high (H) decile. Moreover, the LCP long-short portfolio (H-L) generates a significant monthly FF5+MOM alpha of 1.78% (t-stat = 10.03) and q5+MOM alpha of 1.68% (t-stat = 6.01). The alphas of the LCP long-short portfolio are also significantly higher than those of the univariate long-short portfolios, with the exception of the portfolio based on short seller holdings, for which although the alpha of the LCP long-short portfolio is more than 10% higher, the difference is not significant. Overall, we find that there are benefits from linearly combining the information in the trades and holdings of multiple market participants.

The portfolios based on the nonlinear composite predictor (NLCP) perform even better than those based on the LCP. For instance, the fourth and fifth columns of Table 4 show that the NLCP low and high decile portfolios achieve significant monthly FF5+MOM alphas of -1.49% (t-stat = -4.11) and 1.25% (t-stat = 4.87), which are more than 50% higher in absolute value than those of the LCP low and high decile portfolios. Also, the NCLP long-short portfolio generates significant monthly FF5+MOM alpha of 2.74% (t-stat = 9.80) and q5+MOM alpha of 2.60% (t-stat = 7.60), which are also more than 50% higher than those of the LCP long-short portfolio, with the difference being significant at the 1% confidence level as shown in the sixth and seventh columns of Table 4. Moreover, we find that the alphas of the NLCP long-short portfolio are also significantly higher than those of the univariate long-short portfolios, including that based on short seller holdings. These results highlight the benefits from exploiting nonlinearities and interactions in the relation between the trades and holdings and future stock returns, consistent with the recent literature that emphasizes the importance of nonlinearities and interactions in the prediction of asset returns (Kelly et al., 2019; Gu et al., 2020; Bianchi et al., 2021; Bryzgalova et al., 2021; Binsbergen et al., 2022).[14]

[14]Table B.3 in Appendix B reports the performance of the long-short portfolios obtained based on each of the twelve predictive models individually. The alphas for all methods are statistically significant at the 1% level. Note also that while the performance of the LCP is on par with that of the six linear models, the performance of the NLCP is significantly better than that of the six individual nonlinear machine-learning

Like Gu et al. (2020), our statistical objective functions minimize equally weighted forecast errors, and thus, we focus our discussion primarily on the results for equal-weighted portfolios. However, in Section 4.1, which examines the return predictability of our composite predictors under different information environments, we also evaluate the performance of the LCP and NLCP portfolios across sub-samples of stocks with different firm sizes. We find that the LCP and NLCP long-short portfolio alphas remain generally significant across small, medium, and large stocks, for both equal- and value-weighted portfolio returns, with larger magnitudes for small and medium stocks.

The favorable out-of-sample performance of the LCP and NLCP long-short portfolios is also robust to considering proportional transaction costs of 60 basis points. Table 5 reports the alpha of the LCP and NLCP long-short portfolios, as well as the alpha for the difference between the returns of the NLCP and LCP long-short portfolios in the presence of proportional transaction costs of 30 and 60 basis points. The first column reports the level of transaction costs, the second and third columns report the alpha of the LCP long-short portfolio with respect to the five Fama and French (2015) and momentum factors (FF5+MOM) and the five Hou et al. (2021) and momentum factors (q5+MOM), the fourth and fifth columns for the NLCP long-short portfolio, and the sixth and seventh columns for the difference between the returns of the NLCP and LCP long-short portfolios (NLCP minus LCP). Although the alphas of the LCP and NLCP long-short portfolios decrease with transaction costs, they remain statistically significant at the 1% level even in the presence of transaction costs of 60 basis points. Moreover, the table shows that the monthly alpha of the NLCP long-short portfolio is around 60 basis points higher than that of the LCP long-short portfolio in the presence of transaction costs of 60 basis points, with the difference being significant at the 5% level for the FF5+MOM model and at the 10% level for the q5+MOM model.[15]

[Insert Table 5 here]

models. To understand this, we compute the cross-sectional correlation matrix for the 12 individual-model predictors, and then compute its time-series average. We find that while the average correlation across the six linear predictions is 76%, the average correlation across the six nonlinear predictors is only 61%. This suggests that the six nonlinear predictors capture different information, and thus explains why combining the six nonlinear predictors improves performance.

[15]We find that the monthly turnover of the LCP and NLCP long-short portfolios is 42.14% and 70.19%.

We now examine the persistence of the performance of the NLCP-based portfolios. Table 6 reports the monthly alpha of the long-short portfolios of stocks sorted by LCP and NLCP, for specific months after portfolio formation, ranging from the first month to the 24th month after portfolio formation. The first column reports the number of the month following portfolio formation. The second and third columns report the FF5+MOM and q5+MOM alphas of the LCP long-short portfolio, the fourth and fifth columns for the NLCP long-short portfolio, and the sixth and seventh columns for the difference between the returns of the NLCP and LCP long-short portfolios (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted excess returns on the long-short portfolio, for up to 24 months after portfolio formation, and report the alphas for each month separately.

[Insert Table 6 here]

Table 6 shows that the LCP long-short portfolio continues generating monthly alphas above 1% for up to nine months after portfolio formation, which are significant at the 1% level with respect to the FF5+MOM model and at the 5% level with respect to the q5+MOM model. The monthly alphas of the NLCP long-short portfolio are even more persistent, remaining above 1% for up to 15 months after portfolio formation and significant at the 5% level. The monthly alpha of the NLCP long-short portfolio is higher than that of the LCP long-short portfolio for all 24 months after portfolio formation, with the difference remaining statistically significant for the first two months. Moreover, the difference is statistically significant at the 10% level for 10 out of 24 months after portfolio formation for the FF5+MOM model and 14 out of 24 months for the q5+MOM model. Overall, we find that the out-of-sample performance of the NLCP long-short portfolios is robust to evaluating its alpha for up to 15 months after portfolio formation.

3.3 Controlling for firm characteristics

Although the LCP and NLCP portfolios have significant alpha with respect to the FF5+MOM and q5+MOM factor models, their predictive power may still be subsumed by common firm

characteristics not included in these models. To address this, we use Fama-MacBeth regressions to control for firm characteristics known to predict stock returns. Because Table 4 indicates that the composite predictors predict returns most significantly for the high and low deciles, we create indicator variables for these predictors. Each month, we sort sample stocks into decile portfolios based on their LCP or NLCP values. The indicator variables, I_LCP and I_NLCP, are set to one if the corresponding LCP or NLCP falls within the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of stock returns on these indicator variables controlling for eight firm characteristics: firm size (SIZE), book-to-market ratio (BM), momentum (MOM), short-term reversal (STR), asset growth (AG), gross profitability (GP), turnover (TO), and idiosyncratic volatility (IVOL). Panel C of Table 1 gives a definition of the eight firm characteristics. We winsorize characteristics at the 1st and 99th percentiles and standardize them to have zero mean and unit standard deviation.

[Insert Table 7 here]

The second and third columns of Table 7 report the results for the Fama-MacBeth regressions that include the indicator variables individually. We observe that the spread in expected monthly return between stocks in the high and low deciles of LCP is 0.65%, while the spread for NLCP is more than double, at 1.54%. Note that this spread is double the slope coefficient of the indicator variables, as these variables range from minus one to one. The fourth column reports the regression results when we include both indicator variables. The slope coefficient for the NLCP indicator variable remains largely unchanged, but the coefficient for the LCP indicator variable is close to zero and no longer significant.

The results in Tables 3–7 overall highlight the importance of aggregating the trades and holdings from multiple market participants using machine learning. The NLCP, which contains information extracted from nonlinear interactions among predictors, is a particularly strong return predictor.

3.4 Trading-based and holdings-based composite predictors

Both the linear and nonlinear composite predictors rely on the trades and lagged holdings of market participants. As discussed in the introduction, both trades and lagged holdings of investors may predict stock returns. Trades may reflect more recent or short-term information. Lagged holdings could indicate investors' persistent demand for the stock and hence reflect price pressure (Gompers and Metrick, 2001). They may also embed long-term information about the firm that affects future returns. In this section, we assess the relative significance of trades and lagged holdings for the composite return predictors. To do this, we generate linear and nonlinear composite predictors based on either only the combined trades or only the combined holdings of all market participants.

Table 8 reports the results for the Fama and MacBeth (1973) regressions of monthly stock returns on indicator variables obtained using the composite predictors based on only the trades (LCP_Trading and NLCP_Trading) or only the lagged holdings (LCP_Holding and NLCP_Holding) of all market participants. The indicator variables are equal to one if the corresponding composite predictor is within the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of stock returns on the indicator variables and the firm characteristics defined in Section 2.4. The first column reports the symbol of each explanatory variable, the second, third, fourth, and fifth columns report the results for the regressions that include individually the LCP_Trading, LCP_Holding, NLCP_Trading, or NLCP_Holding indicator variables, respectively, and the sixth column for the regression that includes all four indicator variables.

[Insert Table 8 here]

The second to fifth columns of Table 8 show that (individually) both the trading- and holdings-based predictors, linear or nonlinear, significantly predict monthly stock returns. The coefficient on the NLCP_Holdings indicator variable is the largest, followed by those of the NLCP_Trading, LCP_Holdings, and LCP_Trading indicator variables, in this order. The coefficients are also economically significant: the spread in expected monthly return between a stock in the high and low deciles is 1.16% for NLCP_Holdings, 0.85% for NLCP_Trading,

0.65% for LCP_Holdings, and 0.43% for LCP_Trading. The sixth column shows that, when we include all four indicator variables in Fama-MacBeth regressions, only the trading- and holdings-based NLCP exhibit significant predictability. The coefficients of the linear predictors are no longer statistically significant. The holdings component contributes about 59.2% to the NLCP predictability, while the trading component contributes 40.8%. These results confirm that when put together, NLCP dominates LCP, highlighting the importance of capturing the nonlinear interactions among the trades and holdings of various market participants.

3.5 Variable and interaction importance

In this section, we investigate the relative importance of the 16 predicting variables and their interactions for the predictions of the nonlinear machine-learning models. To do this, we estimate the SHAP values of Lundberg and Lee (2017), which have been used in finance by Bali et al. (2023) and DeMiguel et al. (2023). SHapley Additive exPlanations (SHAP) is a method based on cooperative game theory and used to estimate the contribution of each variable to the prediction for each observation. SHAP is an additive method because aggregating SHAP values across variables, one recovers the difference between the prediction for an individual observation and the average prediction across all observations.[16] To reduce the heavy computational workload associated with computing SHAP values, we focus our analysis on the random-forest model.

[Insert Figure 1 here]

Figure 1 illustrates the importance of each of the 16 predicting variables for random forest. We estimate variable importance by averaging the absolute SHAP values of each variable across all observations within the last estimation window, which spans December 2015 to November 2020. The SHAP values are expressed as monthly returns in decimal form,

[16]The SHAP method is model-agnostic, applicable to any type of data, and provides additive interpretation (contribution of each predicting variable to the prediction) of machine-learning models. Moreover, the tree-based versions of the SHAP method (designed to deal with random forest and gradient boosting) account for dependencies between variables (Lundberg et al., 2020). For these reasons, SHAP has recently become a popular method for visualizing variable and interaction importance. For a general discussion, see Molnar (2019) and for applications in finance Pedersen (2022).

that is, 0.0005 corresponds to a monthly return of five basis points. Figure 1 demonstrates that there is no single variable that drives the predictions of random forest, with 13 out of the 16 predicting variables having an average absolute SHAP value above 2.5 basis points.[17]

[Insert Figure 2 here]

Panel A of Figure 2 depicts the importance of each of the nine investor types estimated as the average absolute value of the sum of the SHAP values for the trading and holding variables corresponding to each investor. Panel A shows that although some investors—such as mutual funds, insurance companies, and wealth management firms—are more important, no single type of investor dominates the predictions, with the average absolute SHAP value of every investor above four basis points. Panel B of Figure 2 graphs the importance of trading versus lagged holding variables measured as the average across all observations of the absolute value of the sum of the SHAP values for all trading or lagged holding variables. The graph shows that lagged holdings are important for the predictions of random forests, with an average absolute SHAP value around 34 basis points. However, consistent with the results in Section 3.4, Panel B of Figure 2 shows that the trading variables contain information (beyond that contained in lagged holding variables) that is valuable for the predictions of random forest, with an average absolute SHAP value close to 25 basis points.

[Insert Figure 3 here]

Figure 3 illustrates the importance of the top 20 interactions among pairs of variables. We estimate interaction importance by averaging the absolute SHAP *interaction* value across all observations within the last estimation window. We find that random forest exploits the interactions between the lagged holdings of different types of investors as well as between the

[17] In contrast with the finding from Table 3 that the univariate portfolio based on short seller holdings generates a large and significant alpha, short seller holdings is not among the top ten most important variables for random forest, with a relatively modest average absolute SHAP value of around three basis points. To understand this, note first that, as shown in Table B.1 of Appendix B, the univariate portfolio based on short seller holdings generates an excess return of 0.71%, which is substantially smaller than its alpha of 1.51%, and the target variable for our machine-learning models is excess return rather than alpha. Second, although short seller holdings are not among the top 10 most important variables in terms of standalone variable importance, we show in this section that short seller holdings are instrumental for the interactions exploited by random forest.

trades of some investors and the lagged holdings of others. Short seller holdings and other institutional holdings, in particular, are instrumental for interaction importance, with one or the other appearing in 16 out of the 20 most important interactions.

[Insert Figure 4 here]

The top four interactions correspond to pairs of lagged institutional holdings (bank holdings and other institutional holdings, hedge fund holdings and other institutional holdings, mutual fund holdings and insurance firm holdings, and mutual fund holdings and wealth management holdings). To understand the nature of these interactions, the four panels of Figure 4 give SHAP plots illustrating the interactions. For each panel, the horizontal axis depicts the (standardized) first variable in the interaction and the vertical axis its SHAP value for each observation (black dots). To visualize the interaction, we split the observations into quintiles of the second variable and depict, for each quintile, the conditional mean SHAP value of the first variable (lines).

The pattern in the four SHAP plots in Figure 4 is quite similar. For instance, Panel C depicts the interaction between lagged mutual fund holdings and lagged insurance company holdings. To simplify the exposition, in the remainder of this section we drop the term "lagged" from the name of holding variables. Our first observation is that there is a strong interaction between mutual fund holdings and insurance company holdings. To see this, note that the lines depicting the conditional SHAP value of mutual fund holdings for each quintile of insurance company holdings differ substantially across quintiles, particularly for negative values of mutual fund holdings. For quintiles 3, 4, and 5 (high insurance company holdings), the lines indicate a positive relation between mutual fund holdings and future returns for negative values of standardized mutual fund holdings, but a mostly flat relation for positive values of standardized mutual fund holdings. In contrast, the lines for quintiles 1 and 2 (low insurance company holdings) have an inverted-U shape. For instance, the line for the first quintile is steeply increasing in mutual fund holdings in the range from -1.0 to around -0.85, decreasing from -0.85 to 0.0, and mildly increasing for positive values of mutual-fund holdings. That is, while there is a positive monotonic relation between mutual fund holdings

and SHAP values for stocks with high insurance company holdings, there is a nonmonotonic, inverted-U shape for stocks with low insurance company holdings. Another observation from Panel C of Figure 4 is that the relation between mutual-fund holdings and predicted returns is highly nonlinear.

[Insert Figure 5 here]

Figure 3 shows that there are also important interactions between trading variables and holding variables. For instance, among the top 20 interactions, there are several interactions between the trades of several types of investors and the holdings of other institutions, defined as institutions that we cannot classify as either mutual funds, hedge funds, banks, insurance companies, or wealth management firms. To investigate the nature of these interactions, the three panels in Figure 5 illustrate the interactions between the trades of three types of investors and the holdings of other institutions. For instance, Panel A depicts the interaction between mutual fund trades and lagged holdings of other institutions. The horizontal axis depicts the (standardized) mutual fund trades and the vertical axis their SHAP value for each observation (dots). To visualize the interaction, we use color to display the original value of the lagged holdings of other 13F institutions (red for high value and blue for low value).

Panel A of Figure 5 shows that, consistent with our findings from the univariate portfolios in Table 3, there is a negative relation between mutual fund trades and future returns, that is, the aggregate quarterly trades of mutual funds negatively predict next month stock returns. However, the dot coloring pattern demonstrates that this negative relation is particularly pronounced for stocks held by other institutions in the previous quarter (red dots). To see this, note that the red dots tend to have the highest SHAP values for negative values of standardized mutual fund trades, but the lowest for positive values of standardized mutual fund trades. Panel B illustrates the interaction between insurance company trades and holdings of other institutions, showing that there is (on average) a positive relation between the trades of insurance companies and future returns.[18] However, the relation is almost flat for stocks held

[18]This result contrasts with the result from the univariate portfolios in Table 3, where we find that insurance company trades negatively predict abnormal returns. The explanation is that the SHAP value plots depict the *marginal* contribution of insurance company trades to the return predictions of random forest, controlling for the other 15 predicting variables.

by other institutions in the previous quarter (red dots). Panel C illustrates the interaction between wealth management firm trades and lagged holdings of other institutions and it shows a very similar pattern to that in Panel B.[19]

The results in this section show that, while some investors (for instance, mutual funds, insurance companies, and wealth management firms) appear more important, there is no single type of investors that drives the predictions of random forest. Consistent with our findings from the Fama-MacBeth regressions, we find that both trades and lagged holdings are important for random forest. We also find that random forest exploits the interactions between the lagged holdings of different types of investors as well as between the trades of some investors and the lagged holdings of others. Short seller holdings and other institutional holdings, in particular, are instrumental for the interactions revealed by random forest, with one or the other appearing in 16 out of the 20 most important interactions.

4 Information environment and firm fundamentals

Prior studies show that expected stock returns are positively related to ex-ante information asymmetry (Diamond and Verrecchia, 1991; Verrecchia, 2001; O'Hara, 2003; Easley and O'Hara, 2004). If composite return predictors reflect valuable information that is not yet absorbed by the market, we would expect their ability to predict returns to be stronger for firms with a more uncertain information environment, where opportunities for informed investors are greater, but arbitrage costs are also relatively high. In addition, composite return predictors may also predict future firm fundamentals and earnings announcement returns. In this section, we test these hypotheses.

[19] Figure 5 shows that the holdings of other institutions play an important role in many of the interactions that are exploited by random forest. As mentioned above, we classify as "other" those institutions that we are not able to classify as mutual funds, hedge funds, banks, insurance companies, or wealth management firms. Thus, other institutions in our dataset are likely to consist of a mixed of different types of institutional investors, and thus, other institutional holdings may represent a proxy for general institutional holdings. Figure B.1 in Appendix B provides evidence for this hypothesis by depicting the interactions between the trades of mutual funds, insurance companies, and wealth management firms and the lagged holdings of the *remaining institutions*. For instance, Panel A depicts the interaction between mutual fund trading and the aggregate lagged holdings of institutional investors other than mutual funds. The three panels in Figure B.1 confirm that the insights from Figure 1 are robust to coloring the dots using the lagged holdings of all remaining institutions instead of the holdings of other institutions.

4.1 Portfolio performance and firm size

We first study how the LCP and NLCP portfolio returns vary across subsamples of stocks with different market capitalization. Given the evidence in the literature that there is greater information uncertainty about smaller firms (e.g., Atiase, 1985; Bamber, 1987; Llorente et al., 2002), we would expect that the predictive power of the composite predictors should be stronger for smaller firms. To test this hypothesis, we first sort all sample stocks into terciles by firm size at the end of each month. Then, within each tercile, we further sort stocks into decile portfolios based on the LCP or NLCP. We hold the stocks for one month. Table 9 reports the FF5+MOM alphas of the LCP and NLCP portfolios across terciles of stocks with different firm sizes.

[Insert Table 9 here]

Overall, both the LCP and NLCP show significant return predictability across different firm sizes. Consistent with the evidence that information asymmetry is greater for smaller firms, the long-short portfolio alphas are highest for small stocks, at 3.27% for the LCP and 3.68% for the NLCP, both significant at the 1% level. Alphas decrease substantially with firm size. For medium stocks, the long-short portfolio alphas for the LCP and NLCP are smaller at 1.03% and 1.29%, but they remain significant at the 1% level. For large stocks, the long-short portfolio alphas further decrease to 0.65% for LCP and 0.26% for NLCP, and while the LCP alpha remains significant at the 1% level, the NLCP alpha is significant only at the 10% level.[20]

To disentangle the effect of size on the performance of the LCP and NLCP, Table 10 reports the results from estimating Fama-MacBeth regressions with the same explanatory and dependent variables as in Table 7, but including interactions between firm size and the LCP and NLCP indicator variables. Table 10 shows that while the interaction between size and the LCP indicator variable is not significant, that between size and the NLCP is significantly

[20]Table B.5 in Appendix B reports the alphas for the value-weighted portfolios. Consistent with the findings in Table 9, both the LCP and NLCP value-weighted long-short portfolios earn significant alphas for small and medium stocks, although with smaller magnitudes. For large stocks, the alpha for LCP remains comparable in magnitude to the equal-weighted results and is significant at the 1% level, while the alpha for NLCP is no longer statistically significant.

negative. Consistent with Table 9, this result shows that the spread in expected returns between the top and bottom deciles generated by the NLCP is significantly larger for smaller stocks. Importantly though, the results in the fourth column of Table 10 show that, when we include the LCP and NLCP indicator variables and their interactions with size simultaneously in Fama-MacBeth regressions, only the NLCP indicator variable and its interaction with size remain significant. Thus, the predictability of the NLCP dominates that of the LCP in Fama-MacBeth regressions even when we control for the interaction terms of NLCP and firm size.

[Insert Table 10 here]

4.2 Effect of information environment

We next examine how the predictive power of the composite return predictors varies across sub-samples of stocks categorized by other commonly-used proxies for uncertainty and information asymmetry such as firm age, idiosyncratic volatility, illiquidity, and analyst coverage and disagreement (Llorente et al., 2002; Amihud, 2002; Yu, 2008; Diether et al., 2002). We measure firm age as the number of years since the firm was first covered by CRSP, idiosyncratic volatility as the standard deviation of the residuals from regressing daily excess stock returns on the three-factor Fama and French (1992) model over the previous six months, illiquidity following Amihud (2002), analyst coverage as the number of analysts with valid earnings per share forecast as of prior month-end, and analyst disagreement as analyst forecast dispersion (standard deviation of analyst forecasts divided by absolute value of average analyst forecast in the prior month). At the end of each month, we sort all sample stocks into halves based on an information-asymmetry proxy. We then run Fama and MacBeth (1973) regressions of stock returns on the I_LCP and I_NLCP indicator variables, and the firm characteristics defined in Section 3.3 for each sub-sample, and report the results in Table 11.

[Insert Table 11 here]

We find that the spread in the expected returns of firms with high versus low NLCP is higher for firms with higher idiosyncratic volatility, higher illiquidity, lower analyst coverage,

and higher analyst disagreement, although it is similar for young versus old firms. This is generally consistent with our conjecture that the NLCP predictability should be stronger for stocks with more uncertain information environment. We also find that the coefficient on the LCP indicator variable is not significant when we control for the NLCP indicator variable in the regressions, consistent with the findings from Table 7.

4.3 Predicting fundamentals and earnings announcement returns

We now test whether the LCP and NLCP contain information about future firm fundamentals related to operating performance and earning announcement returns. We consider four measures of operating performance: cash flows (CF), which is the difference between income before extraordinary items and total accruals, divided by total assets, gross margin (GM), which is is sales minus cost of goods sold scaled by current sales, return-on-equity (ROE), which is the sum of income before extraordinary items and interest expenses, divided by the lagged total equity, and return-on-asset (ROA), which is the sum of income before extraordinary items and interest expenses, divided by the lagged total assets. Earnings announcement returns (CAR) are cumulative abnormal returns in the $[-1, 1]$ three-day window around the earnings announcement day, where abnormal returns are the difference between the daily stock return and that of the corresponding portfolio among the six size and book-to-market Fama-French portfolios.

Table 12 reports the results for the Fama and MacBeth (1973) regressions of future operating performance and earning announcement returns on the indicator variables I_LCP and I_NLCP, controlling for the firm characteristics defined in Section 3.3. In addition, to account for auto-correlation in operating performance, we include the lagged values of the dependent variables as controls in the regressions.

[Insert Table 12 here]

Table 12 shows that the NLCP indicator variable significantly predicts future CF, GM, ROE, and ROA. Compared to a stock in the low NLCP decile, a stock in the high NLCP decile has an expected next-quarter CF ratio 0.914% higher, GM ratio 0.506% higher, ROE ratio

1.554% higher, and ROA ratio 1.252% higher, with all these differences being statistically significant at the 5% confidence level.[21] The results are qualitatively similar for the LCP indicator variable, except that the spread between the expected next-quarter ROE of stocks in the high and low LCP deciles is not statistically significant. Specifically, compared to a stock in the low LCP decile, a stock in the high LCP decile has an expected next-quarter CF ratio 0.95% higher, GM ratio 0.52% higher, and ROA ratio 0.98% higher.

Given that the NLCP can predict future firm fundamentals, it should also predict future earnings announcement returns, which capture the combined effect of reported fundamentals on stock returns. Indeed, Table 12 shows that the NLCP indicator variable significantly predicts future earnings announcement returns after controlling for firm characteristics. The spread between the expected earnings announcement return of stocks in the high and low NLCP deciles is 0.56%, significant at the 1% level. The LCP indicator variable, on the other hand, does not significantly predict earnings announcement returns.

Overall, these results suggest that the return predictability of the composite predictors originates at least partially from their ability to predict future firm fundamentals and earnings announcement returns.

5 Asset-pricing implications

In this section, we examine the asset-pricing implications of the composite return predictors. In particular, we study whether the composite return predictors assign stocks on the right side of anomalies, and whether a factor model constructed using the composite predictors spans the investment opportunity set.

5.1 Stock anomalies and composite predictors

The literature has studied whether the trades of different market participants are consistent with the predictability of well-known stock anomalies. For example, Edelen et al. (2016) show that institutional investors tend to trade on the wrong side of anomalies. McLean et al. (2025)

[21] As mentioned before, the spread in expected operating performance between the high and low deciles is double the slope coefficient of the indicator variables, as these variables range between minus one and one.

conduct a more systematic analysis of the trades of nine types of market participants, and find that while retail investors tend to trade on the wrong side of anomalies, firms and short sellers trade (on average) on the right side of anomalies. To gauge the information about anomalies contained in the composite return predictors, we compute the value of various anomalies for each decile portfolio of stocks sorted by LCP and NLCP.

We follow Green et al. (2017) to construct 102 anomalies based on firm characteristics. We then select the 12 anomalies whose returns are significant (t-stat >1.66) for our sample period from 2008 to 2020. The 12 significant anomalies are market capitalization (Banz, 1981), book-to-market (Rosenberg et al., 1985), gross margin (Novy-Marx, 2013), illiquidity (Amihud, 2002), idiosyncratic volatility (Ali et al., 2003), momentum_12m (Jegadeesh, 1990), momentum_1m (Jegadeesh and Titman, 1993), asset growth (Cooper et al., 2008), dividend yield (Litzenberger and Ramaswamy, 1981), analyst coverage (Elgers et al., 2001), price delay (Hou and Moskowitz, 2005), and combined fundamental (Mohanram, 2005). Table 13 reports the time-series average of the cross-sectional means of firm characteristics corresponding to the 12 stock anomalies for portfolios sorted by either LCP or NLCP. For easier interpretation, we change the sign of the four anomalies whose return is negative (market capitalization, idiosyncratic volatility, momentum_1m, and asset growth), so that the mean anomaly values for the LCP and NLCP long-short portfolios should be positive if they select stocks on the right side of the anomalies.

[Insert Table 13 here]

Panel B of Table 13 shows that the characteristic spreads for the NLCP long-short portfolio are significantly positive for 10 out of the 12 anomalies. Thus, the NLCP selects stocks on the right side of most anomalies. A notable exception is analyst coverage, for which the assignment of stocks based on the NLCP is significantly misaligned with the anomaly. The results for LCP in Panel A are very similar, except that the characteristic spreads are significantly positive for only eight characteristics. These findings indicate that, despite relying only on market participant trades and lagged holdings, the composite return predictors replicate some of the information available in well-known stock anomalies.

5.2 Factor models constructed from composite predictors

The results in the previous section show that both composite predictors are (on average) aligned with most anomalies. In this section, we further explore whether a factor model constructed using the LCP or NLCP can help to explain anomaly returns.

Table 14 reports the performance of several factor models in terms of their ability to explain anomaly returns and the Sharpe ratio generated by their factors. We consider four factor models: (i) the five Fama and French (2015) and momentum factors (FF5+MOM), (ii) the five Hou et al. (2021) and momentum factors (q5+MOM), (iii) a two-factor model including the market factor and the LCP factor defined as the return of the LCP equal-weighted long-short portfolio (LCP+MKT), and (iv) a two-factor model including the market factor and the NLCP factor defined as the return of the NLCP equal-weighted long-short portfolio (NLCP+MKT). We run time series regressions of the returns of the 12 significant anomalies on each of the four factor models. The first and second rows in the table report the average (across the 12 anomalies) of the absolute value of the alpha and alpha t-stat. The third row in the table reports the annualized Sharpe ratio of the mean-variance portfolio of the factors in each model, and the fourth and fifth rows report the p-value for the difference between the Sharpe ratio of each model and that of the FF5+MOM and q5+MOM models.

[Insert Table 14 here]

Table 14 shows that the average absolute value of the alpha of the 12 anomalies with respect to the LCP and NLCP factor models is 0.46 and 0.50, with average absolute t-statistics of 1.19 and 1.16. For comparison, the average absolute alpha with respect to the FF5+MOM and q5+MOM models is 0.59 and 0.44, with average absolute t-statistics of 2.03 and 1.41. That is, the two-factor model with only the market and our LCP or NLCP factor performs similar to the prominent FF5+MOM and q5+MOM models in terms of its ability to explain anomaly returns. This is a notable result because the LCP and NLCP factors are constructed using only trades and holdings of market participants, ignoring any anomaly information.

Barillas and Shanken (2018) show that (under the assumption that a good factor model should span not only the test assets but also the factors in other models) test assets are

irrelevant and it suffices to compare models in terms of the Sharpe ratio generated by their factors. We find that the NLCP factor model produces an annualized Sharpe ratio of 3.09, which is substantially higher than those generated by the FF5+MOM, q5+MOM, and LCP models (1.13, 1.71, and 2.32, respectively). Moreover, the difference between the Sharpe ratio of the NLCP model and those of the FF5+MOM and q5+MOM models is significant.

In summary, the Sharpe ratio results confirm that the nonlinear interaction effects exploited by the NLCP contain information that helps to span the overall investment opportunity set better than the LCP and existing factor models. In other words, the NLCP identifies information that while being important for asset pricing, is not captured by either existing factor models or the set of anomalies we consider.

6 Conclusion

We use machine learning to study whether trades and holdings of multiple market participants contain information about stocks. Our machine learning approach captures nonlinearities and interactions in the relation between trades and holdings and future stock returns. A long-short portfolio based on the nonlinear composite predictor (NLCP) yields monthly alphas with respect to prominent factor models exceeding 2.5%. The predictive power of the NLCP is not driven by just a few types of market participants, and both trades and holdings as well as their interactions are important to predict returns.

The predictability of the NLCP is stronger for stocks with lower market capitalization, higher idiosyncratic volatility, lower liquidity, lower analyst coverage, and higher analyst disagreement, suggesting that the information in the NLCP is more useful for firms with higher information uncertainty. The NLCP contains information about firm fundamentals (predicting future cash flows, gross margin, and return on assets) and predicts also earnings announcement returns. Furthermore, the NLCP assigns stocks on the right side of most anomalies and a two-factor model with the market factor and an NLCP-based factor generates higher Sharpe ratio than prominent factor models.

Our findings suggest that combining the trades and holdings of multiple participants and accounting for nonlinearities and interactions provides valuable information for price discovery.

References

Aggarwal, R. K. and Jorion, P. (2010). "The performance of emerging hedge funds and managers". Journal of Financial Economics 96, 238–256.

Ali, A., Hwang, L.-S., and Trombley, M. A. (2003). "Arbitrage risk and the book-to-market anomaly". Journal of Financial Economics 69, 355–373.

Amihud, Y. (2002). "Illiquidity and stock returns: Cross-section and time-series effects". Journal of Financial Markets 5, 31–56.

Asquith, P., Pathak, P. A., and Ritter, J. R. (2005). "Short interest, institutional ownership, and stock returns". Journal of Financial Economics 78, 243–276.

Atiase, R. (1985). "Predisclosure information, firm capitalization, and security price behavior around earnings announcements". Journal of Accounting Research 23, 21–36.

Baker, M., Litov, L., Wachter, J. A., and Wurgler, J. (2010). "Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements". Journal of Financial and Quantitative Analysis 45, 1111–1131.

Bali, T. G., Beckmeyer, H., Moerke, M., and Weigert, F. (2023). "Option return predictability with machine learning and big data". The Review of Financial Studies 36, 3548–3602.

Bamber, L. S. (1987). "Unexpected earnings, firm size, and trading volume around quarterly earnings announcements". The Accounting Review 62, 510–532.

Banz, R. W. (1981). "The relationship between return and market value of common stocks". Journal of Financial Economics 9, 3–18.

Barber, B. M. and Odean, T. (2000). "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors". The Journal of Finance 55, 773–806.

Barber, B. M., Odean, T., and Zhu, N. (2009). "Do Retail Trades Move Markets?" The Review of Financial Studies 22, 151–186.

Barber, B. M., Huang, X., Jorion, P., Odean, T., and Schwarz, C. (2024). "A (sub) penny for your thoughts: Tracking retail investor activity in TAQ". The Journal of Finance 79, 2403–2427.

Barillas, F. and Shanken, J. (2018). "Comparing asset pricing models". The Journal of Finance 73, 715–754.

Bianchi, D., Büchner, M., and Tamoni, A. (2021). "Bond risk premiums with machine learning". The Review of Financial Studies 34, 1046–1089.

Binsbergen, J. H. van, Han, X., and Lopez-Lira, A. (2022). "Man versus machine learning: The term structure of earnings expectations and conditional biases". The Review of Financial Studies.

Boehmer, E., Huszar, Z. R., and Jordan, B. D. (2010). "The good news in short interest". Journal of Financial Economics 96, 80–97.

Boehmer, E., Jones, C. M., Zhang, X., and Zhang, X. (2021). "Tracking retail investor activity". The Journal of Finance 76, 2249–2305.

Boehmer, E., Jones, C. M., Wu, J., and Zhang, X. (2020). "What do short sellers know?" Review of Finance 24, 1203–1235.

Breiman, L. (2001). "Random forests". Machine Learning 45, 5–32.

Bryzgalova, S., Pelger, M., and Zhu, J. (2021). "Forest through the trees: Building cross-sections of stock returns". Available at SSRN 3493458.

Cao, C., Liang, B., Lo, A. W., and Petrasek, L. (2018). "Hedge fund holdings and stock market efficiency". The Review of Asset Pricing Studies 8, 77–116.

Chatigny, P., Goyenko, R., and Zhang, C. (2022). "Asset pricing with attention guided deep learning". Available at SSRN 3971876.

Chinco, A., Clark-Joseph, A. D., and Ye, M. (2019). "Sparse signals in the cross-section of returns". The Journal of Finance 74, 449–492.

Cooper, M. J., Gulen, H., and Schill, M. J. (2008). "Asset growth and the cross-section of stock returns". The Journal of Finance 63, 1609–1651.

Da, R. (2022). "Market efficiency with many investors". Chicago Booth working paper.

Da, Z., Dong, X., Wu, K., and Zhou, D. (2024). "Inside and outside informed trading". Mendoza College of Business working paper.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks". The Journal of Finance 52, 1035–1058.

Davis, C., Kargar, M., and Li, J. (2024). "Why is asset demand inelastic?" Available at SSRN: https://ssrn.com/abstract=4195089.

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

Desai, H., Ramesh, S., Thiagarajan, R., and Balachandran, B. V. (2002). "An investigation of the informational role of short interest in the Nasdaq market". The Journal of Finance 57, 2263–2287.

Diamond, D. W. and Verrecchia, R. E. (1991). "Disclosure, liquidity, and the cost of capital". The Journal of Finance 46, 1325–1359.

Diether, K. B., Lee, K.-H., and Werner, I. M. (2009). "Short-sale strategies and return predictability". The Review of Financial Studies 22, 575–607.

Diether, K. B., Malloy, C. J., and Scherbina, A. (2002). "Differences of opinion and the cross section of stock returns". The Journal of Finance 57, 2113–2141.

Dong, X., Li, Y., Rapach, D., and Zhou, G. (2022). "Anomalies and the expected market return". The Journal of Finance 77, 639–681.

Easley, D. and O'Hara, M. (2004). "Information and the cost of capital". The Journal of Finance 59, 1553–1583.

Edelen, R. M., Ince, O. S., and Kadlec, G. B. (2016). "Institutional investors and stock return anomalies". Journal of Financial Economics 119, 472–488.

Egan, M. L., MacKay, A., and Yang, H. (2024). "What drives variation in investor portfolios? Estimating the roles of beliefs and risk preferences". NBER Working Paper.

Elgers, P. T., Lo, M. H., and Jr., R. J. P. (2001). "Delayed security price adjustments to financial analysts' forecasts of annual earnings". The Accounting Review 76, 613–632.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996). "The persistence of risk-adjusted mutual fund performance". Journal of Business 69, 133–157.

Fama, E. F. and French, K. R. (1992). "The cross-section of expected stock returns". The Journal of Finance 47, 427–465.

— (2010). "Luck versus skill in the cross-section of mutual fund returns". The Journal of Finance 65, 1915–1947.

Fama, E. F. and French, K. R. (2015). "A five-factor asset pricing model". Journal of Financial Economics 116, 1–22.

Fama, E. F. and MacBeth, J. D. (1973). "Risk, return, and equilibrium: Empirical tests". Journal of Political Economy 81, 607–636.

Feng, G., Giglio, S., and Xiu, D. (2020). "Taming the factor zoo: A test of new factors". The Journal of Finance 75, 1327–1370.

Ferreira, M. A., Keswani, A., Miguel, A. F., and Ramos, S. B. (2013). "The determinants of mutual fund performance: A cross-country study". Review of Finance 17, 483–525.

Gao, M. and Zhang, C. (2023). "Optimizing Return Forecasts: A Bayesian Intermediary Asset Pricing Approach". Available at SSRN 4545015.

Gompers, P. A. and Metrick, A. (2001). "Institutional investors and equity prices". The Quarterly Journal of Economics 116, 229–259.

Green, J., Hand, J. R. M., and Zhang, X. F. (2017). "The characteristics that provide independent information about average U.S. monthly stock returns". The Review of Financial Studies 30, 4389–4436.

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

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer Science & Business Media.

He, Z. and Krishnamurthy, A. (2013). "Intermediary asset pricing". American Economic Review 103, 732–770.

Hirshleifer, D., Teoh, S. H., and Yu, J. J. (2011). "Short arbitrage, return asymmetry, and the accrual anomaly". The Review of Financial Studies 24, 2429–2461.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2021). "An augmented q-factor model with expected growth". Review of Finance 25, 1–41.

Hou, K. and Moskowitz, T. J. (2005). "Market frictions, price delay, and the cross-section of expected returns". The Review of Financial Studies 18, 981–1020.

Huebner, P. (2024). "The making of momentum: A demand-system perspective". Available at SSRN: https://ssrn.com/abstract=4395945.

Jegadeesh, N. (1990). "Evidence of predictable behavior of security returns". The Journal of Finance 45, 881–898.

Jegadeesh, N. and Titman, S. (1993). "Returns to buying winners and selling losers: Implications for stock market efficiency". The Journal of Finance 48, 65–91.

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

Jiao, Y., Massa, M., and Zhang, H. (2016). "Short selling meets hedge fund 13F: An anatomy of informed demand". Journal of Financial Economics 122, 544–567.

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

Kaniel, R., Liu, S., Saar, G., and Titman, S. (2012). "Individual investor trading and return patterns around earnings announcements". The Journal of Finance 67, 639–680.

Karpoff, J. M. and Lou, X. (2010). "Short sellers and financial misconduct". The Journal of Finance 65, 1879–1913.

Kelley, E. K. and Tetlock, P. C. (2013). "How wise are crowds? Insights from retail orders and stock returns". The Journal of Finance 68, 1229–1265.

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

Koijen, R. S. J. and Yogo, M. (2019). "A demand system approach to asset pricing". Journal of Political Economy 127, 1475–1515.

Kosowski, R., Naik, N. Y., and Teo, M. (2007). "Do hedge funds deliver alpha? A Bayesian and bootstrap analysis". Journal of Financial Economics 84, 229–264.

Kozak, S., Nagel, S., and Santosh, S. (2020). "Shrinking the cross-section". Journal of Financial Economics 135, 271–292.

Leippold, M., Wang, Q., and Zhou, W. (2022). "Machine learning in the Chinese stock market". Journal of Financial Economics 145, 64–82.

Li, B. and Rossi, A. G. (2021). "Selecting mutual funds from the stocks they hold: A machine learning approach". Available at SSRN 3737667.

Litzenberger, R. H. and Ramaswamy, K. (1981). "The effects of dividends on common stock prices tax effects or information effects?" The Journal of Finance 37, 429–443.

Llorente, G., Michaely, R., Saar, G., and Wang, J. (2002). "Dynamic volume-return relation of individual stocks". The Review of Financial Studies 15, 1005–1047.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I. (2020). "From local explanations to global understanding with explainable AI for trees". Nature Machine Intelligence 2, 56–67.

Lundberg, S. M. and Lee, S.-I. (2017). "A unified approach to interpreting model predictions". Advances in Neural Information Processing Systems 30, 1–10.

Massa, M., Qian, W., Xu, W., and Zhang, H. (2015). "Competition of the informed: Does the presence of short sellers affect insider selling?" Journal of Financial Economics 118, 268–288.

McLean, R. D., Pontiff, J., and Reilly, C. (2025). "Taking sides on return predictability". Available at SSRN 3637649. Forthcoming in Journal of Financial Economics.

Mohanram, P. S. (2005). "Separating winners from losers among low book-to-market stocks using financial statement analysis". Review of Accounting Studies 10, 133–170.

Molnar, C. (2019). *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. https://christophm.github.io/interpretable-ml-book/. Lulu.com.

Newey, W. K. and West, K. D. (1994). "Automatic lag selection in covariance matrix estimation". The Review of Economic Studies 61, 631–653.

Noh, D., Oh, S., and Song, J. (2024). "Unpacking the demand for sustainable equity investing". Available at SSRN: https://ssrn.com/abstract=3639693.

Novy-Marx, R. (2013). "The other side of value: The gross profitability premium". Journal of Financial Economics 108, 1–28.

O'Hara, M. (2003). "Presidential address: Liquidity and price discovery". The Journal of Finance 58, 1335–1354.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2017). "Do funds make more when they trade more?" The Journal of Finance 72, 1483–1528.

Patton, A. J. and Weller, B. M. (2020). "What you see is not what you get: The costs of trading market anomalies". Journal of Financial Economics 137, 515–549.

Pedersen, L. H. (2022). "Big data asset pricing 5: Machine learning in asset pricing". Available at SSRN 4068797.

Rapach, D., Strauss, J., Tu, J., and Zhou, G. (2019). "Industry return predictability: A machine learning approach". Journal of Financial Data Science 1, 9–28.

Rosenberg, B., Reid, K., and Lanstein, R. (1985). "Persuasive evidence of market inefficiency". Journal of Portfolio Management 11, 9–16.

Schapire, R. E. and Freund, Y. (2012). *Boosting: Foundations and algorithms*. MIT Press.

Sias, R. W. and Whidbee, D. A. (2010). "Insider trades and demand by institutional and individual investors". The Review of Financial Studies 23, 1544–1595.

Tibshirani, R. (1996). "Regression shrinkage and selection via the Lasso". Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Verrecchia, R. E. (2001). "Essays on disclosure". Journal of Accounting and Economics 32, 97–180.

Wermers, R. (2000). "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses". The Journal of Finance 55, 1655–1695.

Yan, X. and Zhang, Z. (2009). "Institutional investors and equity returns: Are short-term institutions better informed?" The Review of Financial Studies 22, 893–924.

Yu, F. F. (2008). "Analyst coverage and earnings management". Journal of Financial Economics 88, 245–271.

Zou, H. (2006). "The adaptive Lasso and its oracle properties". Journal of the Royal Statistical Society. Series B (Methodological) 101, 1418–1429.

Figure 1. Variable importance

This figure illustrates the importance of each of the 16 predicting variables for random forest. We estimate variable importance by averaging the absolute SHAP values of each variable across all observations within the last estimation window, which spans December 2015 to November 2020. The SHAP values are expressed as monthly returns in decimal form, that is, 0.001 corresponds to a monthly return of 10 basis points.

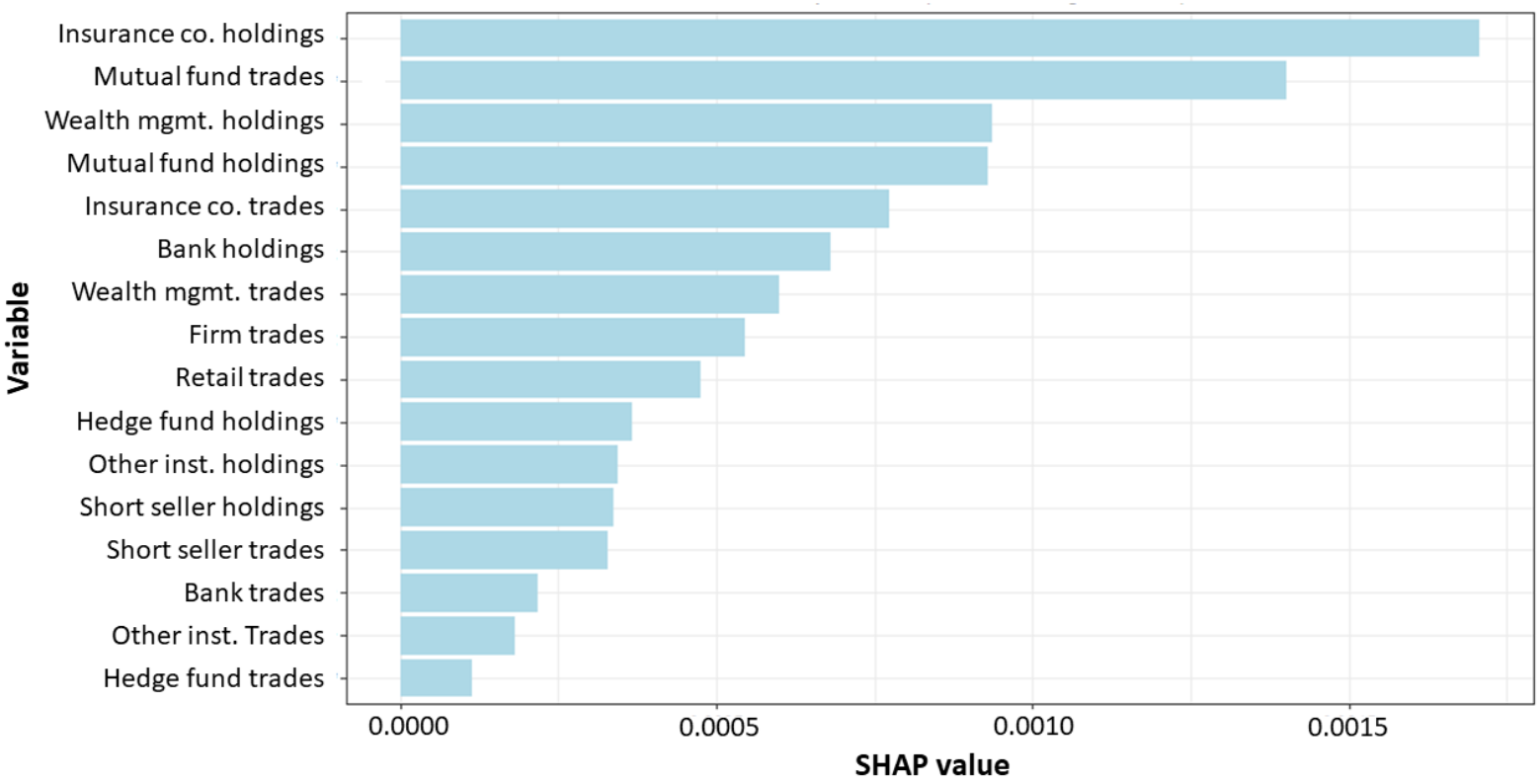
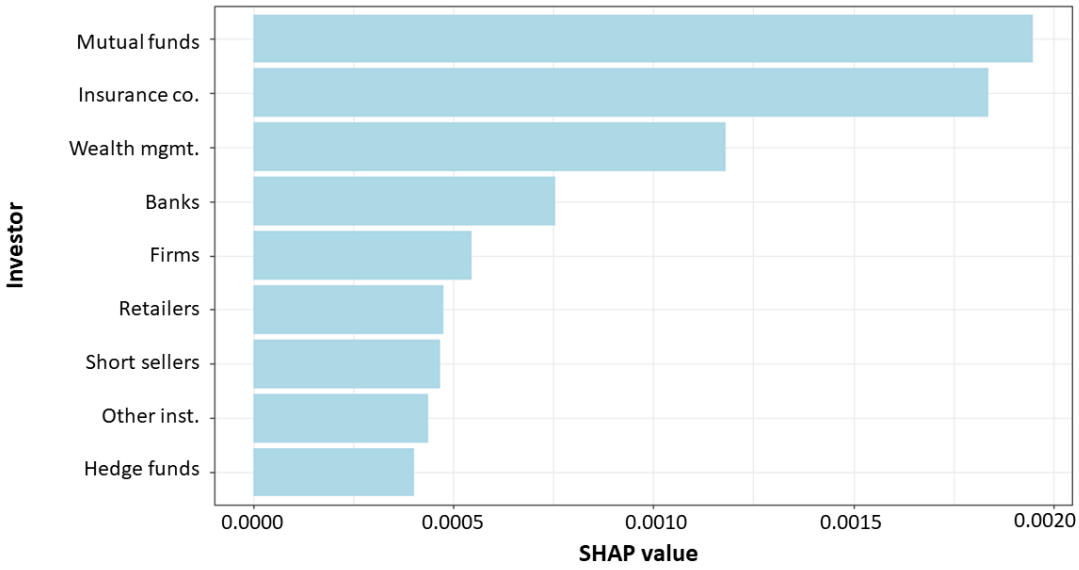


Figure 2. Investor and trades-versus-holdings importance

This figure illustrates the importance of each investor type and that of trades versus holdings for random forest. Panel A graphs the importance of each of the nine investor types estimated as the average absolute value of the sum of the SHAP values for the trading and holding variables corresponding to each investor. Panel B graphs the importance of trading versus holding variables measured as the average of the absolute value of the sum of the SHAP values for all trading or holding variables. We estimate importance for the last estimation window, which spans December 2015 to November 2020.

Panel A. Investor importance



Panel B. Trade versus holdings importance

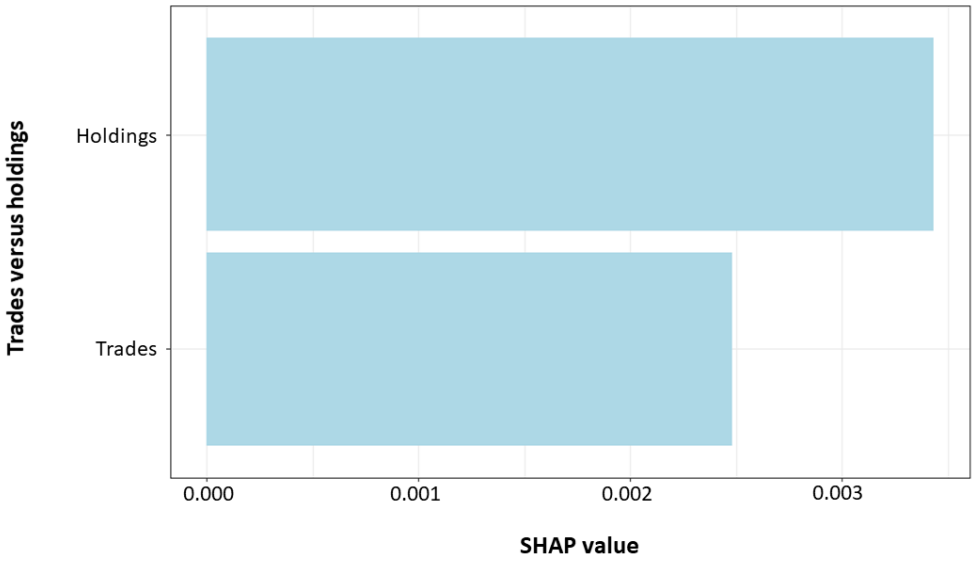


Figure 3. Interaction importance

This figure illustrates the importance of the top 20 interactions among pairs of variables. We estimate interaction importance by averaging the absolute SHAP interaction value across all observations within the last estimation window, which spans December 2015 to November 2020.

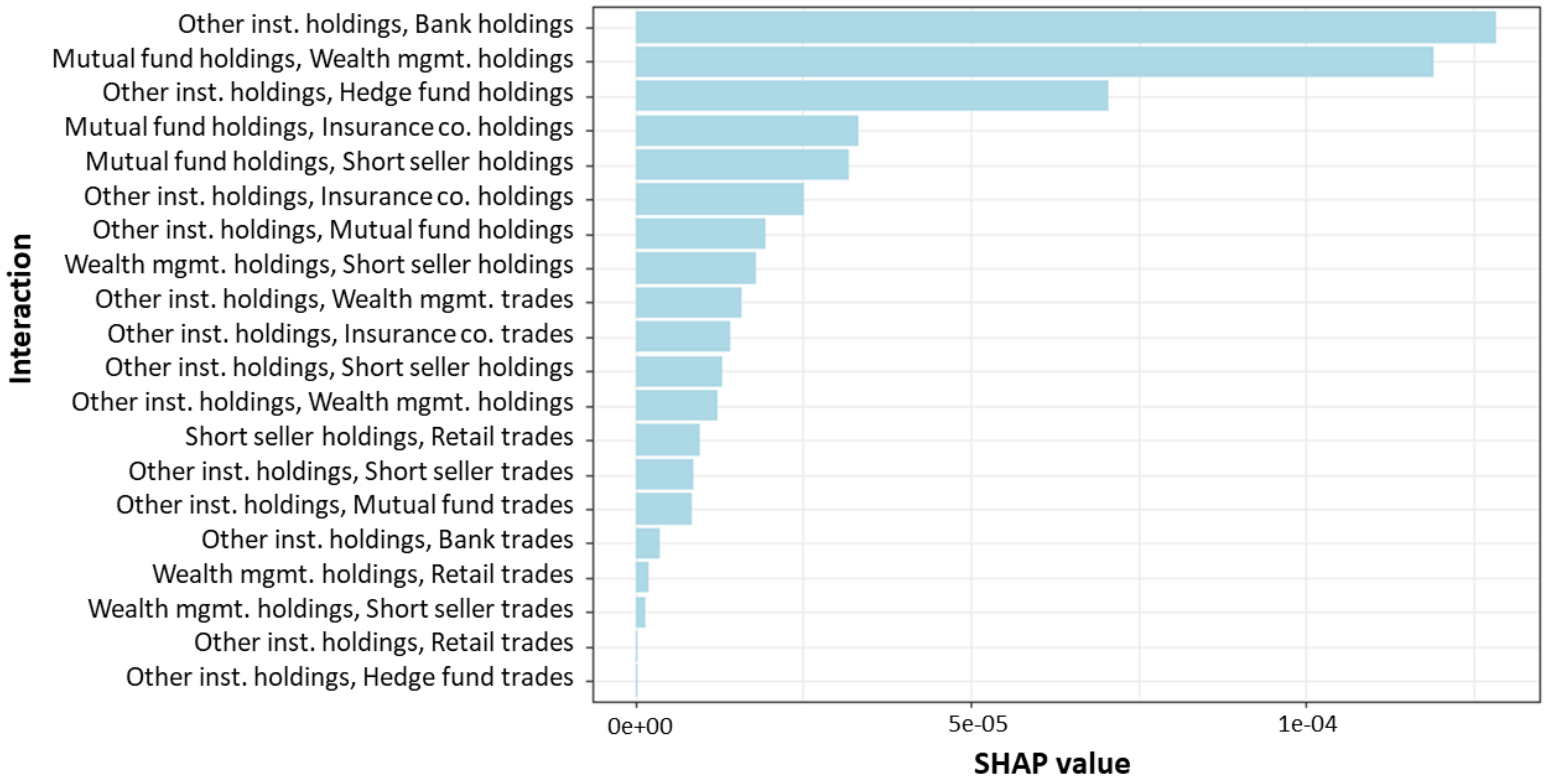
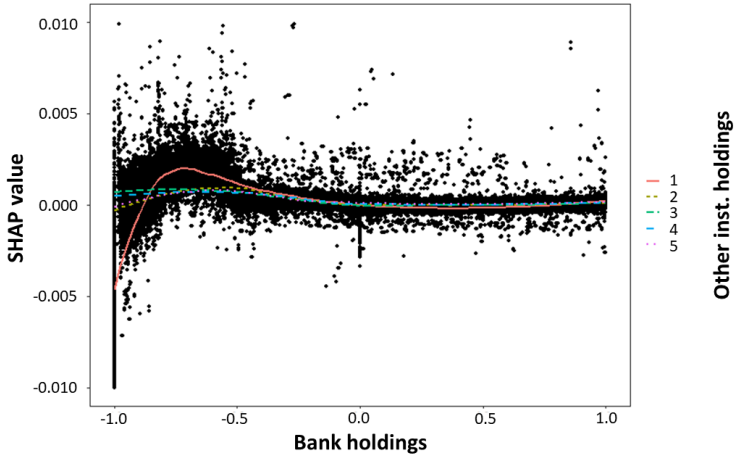


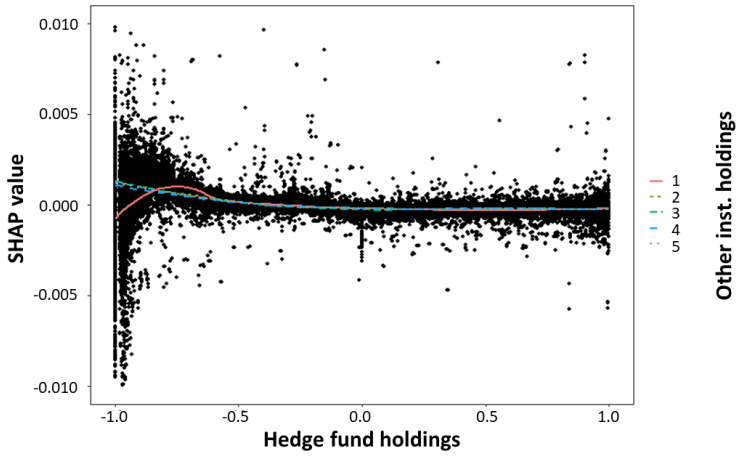
Figure 4. Top four interactions

This figures illustrates the four most important interactions, ranked by average absolute SHAP interaction value. Panel A depicts the interaction between bank holdings and other holdings, Panel B between hedge fund holdings and other holdings, Panel C between mutual fund holdings and insurance company holdings, and Panel D between mutual fund holdings and wealth management firm holdings. For each panel, the horizontal axis depicts the (standardized) first variable and the vertical axis its SHAP value for each observation (black dots). To visualize the interaction, we split all observations into quintiles of the second variable and depict, for each quintile, the conditional mean SHAP value of the first variable (solid lines). We compute SHAP values for the last estimation window, which spans the period from December 2015 to November 2020.

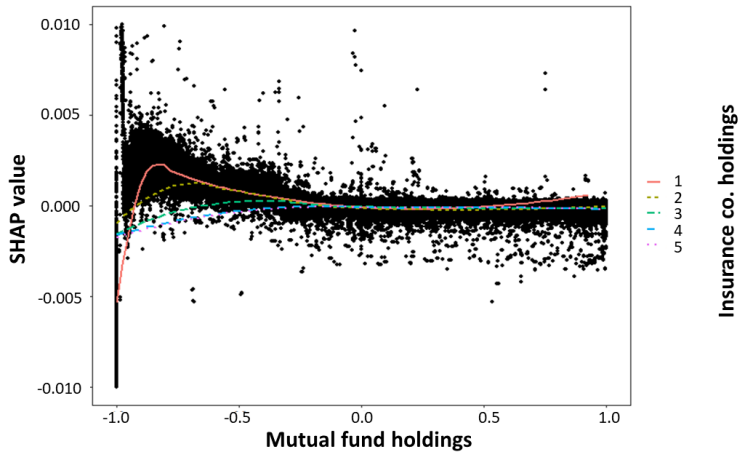
Panel A. Bank vs. other



Panel B. Hedge fund vs. other



Panel C. Mutual fund vs. insurance



Panel D. Mutual fund vs. wealth manag.

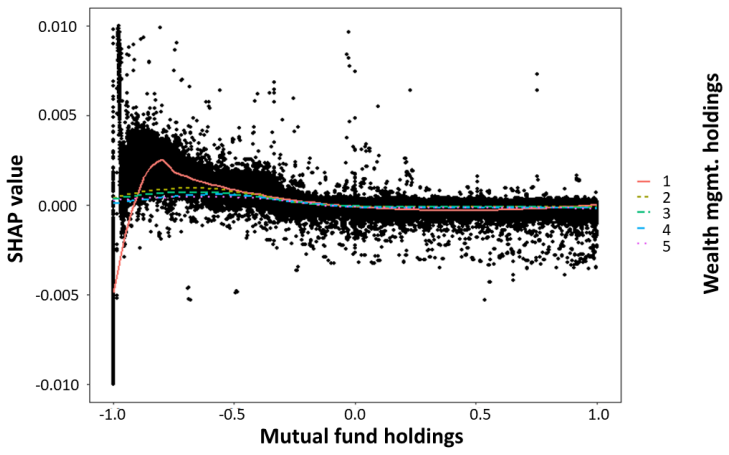
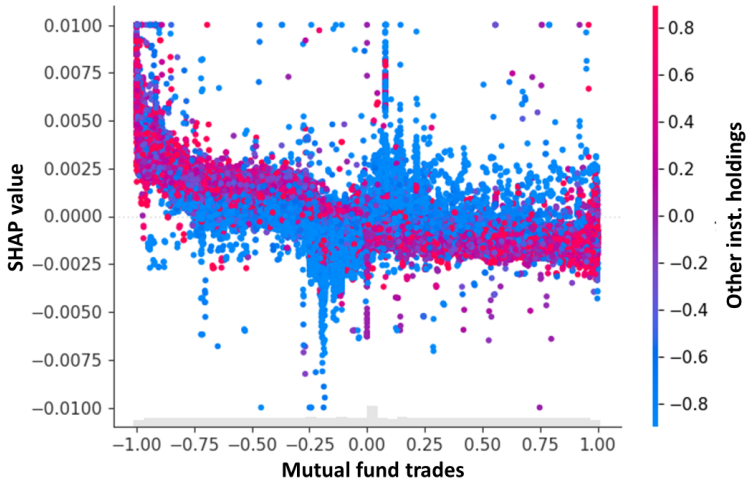


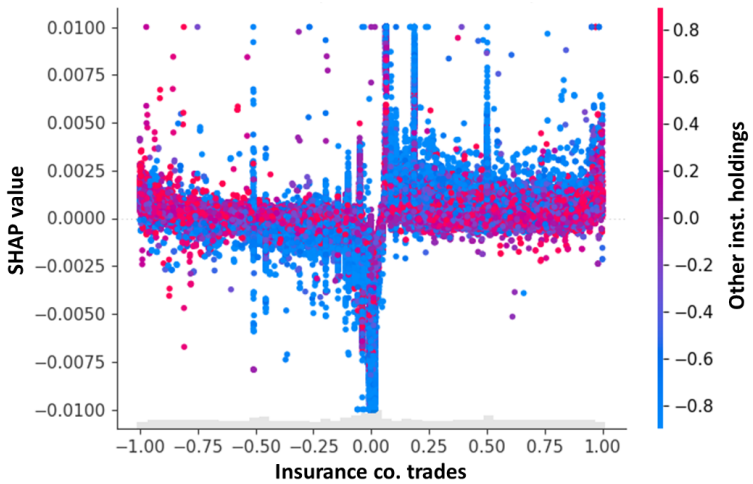
Figure 5. Interactions between trading variables and other institutional holdings

This figure illustrates the interactions between the trades of three types of institutions and the holdings of "other" institutions. Panel A depicts the interaction between mutual fund trades and other institutions holdings, Panel B between insurance company trades and other institutions holdings, and Panel C between wealth management trades and other institutions holdings. For each panel, the horizontal axis depicts the (standardized) trading variable and the vertical axis its SHAP value for each observation (dots). To visualize the interaction, we use color to display the original value of the other institutions holdings variable (red for high value and blue for low value). We compute SHAP values for the last estimation window, which spans December 2015 to November 2020.

Panel A. Mutual fund trades vs. other inst. holdings



Panel B. Insurance co. trades vs. other inst. holdings



Panel C. Wealth mgmt. trades vs. other inst. holdings



Table 1. Composite return predictors, predicting variables, and control variables

Panels A, B, and C of this table provide the definition for the 16 predicting variables, the linear and nonlinear composite return predictors, and the six control variables considered in the Fama and MacBeth (1973) regressions, respectively.

Variable	Definition
Panel A: Predicting variables	
Bank holdings	Quarterly number of shares held by banks, scaled by the total number of shares outstanding
Bank trades	Changes in bank holdings from the previous quarter
Hedge fund holdings	Quarterly number of shares held by hedge funds, scaled by the total number of shares outstanding
Hedge fund trades	Changes in hedge fund holdings from the previous quarter
Insurance co. holdings	Quarterly number of shares held by insurance companies, scaled by the total number of shares outstanding
Insurance co. trades	Changes in insurance company holdings from the previous quarter
Mutual fund holdings	Quarterly number of shares held by mutual funds, scaled by the total number of shares outstanding
Mutual fund trades	Changes in mutual fund holdings from the previous quarter
Wealth mgmt. holdings	Quarterly number of shares held by wealth management firms, scaled by the total number of shares outstanding
Wealth mgmt. trades	Changes in wealth management holdings from the previous quarter
Other inst. holdings	Quarterly number of shares held by remaining 13F institutions, after excluding mutual funds, hedge funds, banks, insurance companies, and wealth management firms, scaled by the total number of shares outstanding
Other inst. trades	Changes in other inst. holdings from the previous quarter
Firm trades	Quarter-end share issues minus share repurchases, scaled by the number of shares outstanding. The variable is multiplied by –1, so more negative (positive) values indicate higher (lower) levels of firm trading
Short seller holdings	Quarter-end short interest, scaled by the total number of shares outstanding. The variable is multiplied by –1, so higher values indicate lower levels of short selling
Short seller trades	Changes in signed short interest from the previous quarter
Retail trades	Average daily order imbalance each month, calculated following Barber et al., 2024 as the difference between retail buy and sell volumes normalized by their sum, based on off-exchange marketable orders
Panel B: Composite return predictors	
Linear composite predictor (LCP)	Monthly average of the out-of-sample excess returns predicted by the six linear models defined in Section 2.2
Nonlinear composite predictor (NLCP)	Monthly average of out-of-sample excess returns predicted by the six nonlinear models defined in Section 2.2

Table 1. (Cont.) Composite return predictors, predicting variables, and control variables

Variable	Definition
Panel C: Control variables	
Firm size (SIZE)	Natural logarithm of firm market capitalization
Book-to-market ratio (BM)	Most recent fiscal year-end book value divided by market capitalization
Momentum (MOM)	Cumulative returns from month -12 to month -2
Short-term reversal (STR)	Prior month's return
Asset growth (AG)	Annual asset growth from the previous fiscal year
Gross profitability (GP)	Gross profit over total assets from last fiscal year
Turnover ratio (TO)	Trading volume over number of shares outstanding last month
Idiosyncratic volatility (IVOL)	Standard deviation of residuals from Fama and French (1992) three-factor model of daily stock excess returns over previous six months

Table 2. Summary statistics

This table provides summary statistics (mean, median, standard deviation, and 10th and 90th percentiles) estimated at the stock-month panel-data level for holdings, trades, and the linear (LCP) and nonlinear (NLCP) composite predictors, for our out-of-sample period from January 2013 to December 2020. We consider nine market participants: banks, firms, hedge funds, insurance companies, mutual funds, other institutions, short sellers, wealth management firms, and retail investors. For firms and retail investors, we consider only trades because we lack information about holdings. We follow McLean et al. (2025) to construct all quarterly trades and quarterly holdings except retail trades, which we construct like Barber et al. (2024). We use a five-year rolling-window approach to train the six linear and six nonlinear prediction models described in Section 2.2 to predict next month's stock excess returns, using quarterly trades and lagged quarterly holdings of nine market participants as predicting variables. The LCP and NLCP are the average predictors across the six linear and six nonlinear models.

Variable	Mean	Median	St.Dev	p10	p90
Panel A: Holdings					
Bank	6.60%	5.54%	5.97%	0.01%	14.66%
Hedge fund	9.52%	8.12%	8.14%	0.40%	20.00%
Insurance co.	1.97%	1.36%	2.33%	0.00%	4.81%
Mutual fund	19.99%	19.94%	14.78%	0.28%	39.96%
Other inst.	20.06%	19.39%	14.78%	1.47%	37.62%
Short seller	-15.46%	-9.20%	20.23%	-37.97%	-0.49%
Wealth mgmt.	2.93%	1.79%	3.33%	0.00%	7.88%
Panel B: Trades					
Bank	-0.05%	0.00%	1.90%	-1.39%	1.38%
Hedge fund	0.02%	0.00%	3.14%	-2.21%	2.29%
Insurance co.	-0.01%	0.00%	0.85%	-0.50%	0.48%
Mutual fund	0.22%	0.02%	3.60%	-2.86%	3.52%
Other inst.	-0.14%	-0.05%	10.76%	-5.10%	4.78%
Short seller	-0.12%	0.01%	14.11%	-5.24%	4.85%
Wealth mgmt.	0.03%	0.00%	1.01%	-0.35%	0.46%
Firm	-1.60%	-0.08%	11.33%	-2.32%	0.98%
Retail	-2.61%	-1.49%	13.41%	-16.69%	9.87%
Panel C: Composite predictors					
LCP	0.63%	0.62%	0.58%	-0.10%	1.36%
NLCP	1.22%	1.17%	0.83%	0.35%	2.23%

Table 3. Alpha of univariate portfolios of stocks sorted by trades or lagged holdings

This table reports the monthly alpha of univariate portfolios obtained by sorting stocks based on the trades (or lagged holdings) of a single type of market participant. Panels A and B report the results for the univariate portfolios constructed using trades and lagged holdings, respectively. For each panel, the first column reports the name of the sorting variable, the second to seventh columns report the alpha of the low, second, third, eighth, ninth, and high decile portfolios, respectively, and the eighth column reports the alpha for the long-short portfolio, which goes long stocks in the high decile and short stocks in the low decile. At the end of each month, we rank all sample stocks based on either the trades or the lagged holdings of each market participant and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile portfolio, as well as on the long-short portfolio, and report their alphas with respect to the five Fama and French (2015) and momentum factors. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. To facilitate the comparison, we evaluate performance in the period from 2013 to 2020, which coincides with the out-of-sample period for the evaluation of the composite return predictors.

	L	2	3	8	9	H	**H - L**
Panel A: Trade sorts							
Banks	0.17%	0.11%	0.15%	0.02%	-0.07%	-0.34%**	-0.5%***
	(1.13)	(0.88)	(1.18)	(0.17)	(-0.70)	(-2.33)	(-2.74)
Firms	-0.39%	-0.14%	0.31%**	0.13%	0.28%***	0.00%	0.39%
	(-1.20)	(-0.82)	(2.36)	(0.65)	(3.23)	(0.00)	(1.25)
Hedge funds	-0.21%	-0.08%	0.10%	0.08%	0.16%	0.16%	0.37%**
	(-1.51)	(-0.62)	(0.76)	(0.69)	(1.57)	(1.04)	(2.49)
Insurance co.	0.14%	0.10%	0.13%	-0.10%	-0.10%	-0.09%	-0.23%*
	(1.30)	(0.82)	(0.94)	(-0.77)	(-1.37)	(-0.91)	(-1.67)
Mutual funds	0.26%**	0.16%	0.24%**	-0.12%	-0.20%*	-0.13%	-0.39%**
	(2.07)	(0.90)	(2.23)	(-1.03)	(-1.87)	(-1.00)	(-2.07)
Other inst.	-0.09%	0.08%	-0.07%	0.16%	0.07%	0.29%*	0.37%***
	(-0.51)	(0.76)	(-0.56)	(1.04)	(0.64)	(1.81)	(2.69)
Short sellers	-0.66%***	-0.16%	0.01%	0.29%**	-0.17%	-0.45%**	0.21%
	(-2.91)	(-0.92)	(0.07)	(2.18)	(-1.14)	(-2.56)	(0.87)
Wealth mgmt.	0.18%	0.17%	0.33%**	0.14%	-0.05%	-0.12%	-0.30%*
	(1.14)	(1.64)	(2.49)	(1.08)	(-0.41)	(-0.81)	(-1.80)
Retail	0.14%	0.23%**	0.01%	-0.26%	-0.04%	0.62%***	0.48%**
	(1.32)	(2.00)	(0.06)	(-1.35)	(-0.21)	(3.02)	(2.17)
Panel B: Lagged-holdings sorts							
Banks	-0.41%	0.31%	0.60%**	-0.10%	0.08%	-0.13%	0.27%
	(-0.99)	(0.93)	(2.21)	(-1.34)	(1.11)	(-1.44)	(0.67)
Hedge funds	-0.54%	0.40%*	0.16%	0.13%	0.09%	-0.02%	0.52%
	(-1.43)	(1.76)	(1.04)	(1.06)	(0.95)	(-0.15)	(1.44)
Insurance co.	0.25%	-0.09%	0.45%	-0.08%	0.00%	-0.07%	-0.32%
	(0.69)	(-0.25)	(1.53)	(-1.18)	(0.01)	(-0.85)	(-0.93)
Mutual funds	-0.78%*	0.74%**	0.48%**	-0.01%	-0.09%	-0.10%	0.69%
	(-1.73)	(2.40)	(2.30)	(-0.06)	(-1.14)	(-1.01)	(1.42)
Other inst.	-0.47%	0.37%*	0.15%	0.01%	0.15%	-0.05%	0.42%
	(-1.11)	(1.84)	(1.01)	(0.14)	(1.62)	(-0.35)	(1.14)
Short sellers	-0.47%**	-0.37%**	-0.42%***	0.24%**	0.33%	1.04%***	1.51%***
	(-2.16)	(-2.15)	(-2.89)	(2.29)	(1.46)	(4.03)	(5.17)
Wealth mgmt.	-0.47%	0.16%	0.26%**	0.01%	0.11%	0.11%	0.58%
	(-1.06)	(0.52)	(2.09)	(0.14)	(0.94)	(0.93)	(1.35)

Table 4. Alpha of portfolios based on composite return predictors

This table reports the monthly alpha of the portfolios of stocks sorted by the linear (LCP) and nonlinear (NLCP) composite return predictors. The first column reports the label for each decile portfolio and the long-short portfolio, the second and third columns report the alpha of each portfolio based on the LCP with respect to the five Fama and French (2015) and momentum factors (FF5+MOM) and the five Hou et al. (2021) and momentum factors (q5+MOM), the fourth and fifth columns for the portfolios based on the NLCP, and the sixth and seventh columns for the difference between the returns of the portfolios based on NLCP and LCP (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile portfolio, as well as on the long-short portfolio, and report their alphas. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	LCP		NLCP		NLCP minus LCP	
	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM
L	-0.98%***	-0.48%*	-1.49%***	-1.00%**	-0.51%***	-0.51%**
	(-3.72)	(-1.93)	(-4.11)	(-2.49)	(-2.69)	(-2.22)
2	-0.36%**	-0.01%	-0.06%	0.29%	0.30%*	0.30%*
	(-2.06)	(-0.03)	(-0.35)	(1.56)	(1.71)	(1.76)
3	-0.13%	0.17%	0.03%	0.31%*	0.16%	0.14%
	(-0.73)	(0.95)	(0.22)	(1.97)	(1.11)	(0.99)
4	0.23%	0.53%**	-0.02%	0.28%**	-0.24%	-0.25%
	(1.28)	(2.51)	(-0.16)	(2.15)	(-1.54)	(-1.37)
5	0.12%	0.35%**	0.09%	0.31%**	-0.03%	-0.03%
	(0.95)	(2.16)	(0.56)	(2.24)	(-0.16)	(-0.17)
6	0.29%**	0.55%***	0.12%	0.42%**	-0.17%	-0.13%
	(2.14)	(3.65)	(0.90)	(2.33)	(-1.15)	(-0.76)
7	0.34%***	0.53%***	0.03%	0.31%*	-0.31%**	-0.22%**
	(2.76)	(4.08)	(0.22)	(1.78)	(-2.19)	(-2.01)
8	0.32%**	0.55%***	0.43%***	0.68%***	0.12%	0.13%
	(2.06)	(4.54)	(3.12)	(4.07)	(1.01)	(1.00)
9	0.39%***	0.73%***	0.63%***	0.91%***	0.24%	0.18%
	(3.32)	(5.77)	(3.57)	(5.31)	(1.57)	(1.36)
H	0.80%***	1.20%***	1.25%***	1.61%***	0.45%*	0.41%
	(3.77)	(5.52)	(4.87)	(5.71)	(1.95)	(1.62)
H-L	1.78%***	1.68%***	2.74%***	2.60%***	0.96%***	0.92%***
	(10.03)	(6.01)	(9.80)	(7.60)	(3.46)	(2.65)

Table 5. Alpha of returns net of transaction costs of LCP and NLCP long-short portfolios

This table reports the monthly alpha of the returns of the long-short portfolio of stocks sorted by the linear and nonlinear composite predictors, as well as the alpha of the difference between the returns of the NLCP and LCP long-short portfolios in the absence of transaction costs (0bps) and in the presence of proportional transaction costs of 30 and 60 basis points (30bps and 60bps). The first column reports the level of transaction costs, the second and third columns report the alpha of the LCP long-short portfolio with respect to the five Fama and French (2015) and momentum factors (FF5+MOM) and the five Hou et al. (2021) and momentum factors (q5+MOM), the fourth and fifth columns for the NLCP long-short portfolio, and the sixth and seventh columns for the difference between the returns of the NLCP and LCP long-short portfolios (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted returns net or proportional transaction costs on each decile as well as on the long-short portfolio and report the alphas. Each month, we compute transaction costs by comparing portfolio weights before and after rebalancing. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	LCP		NLCP		NLCP minus LCP	
	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM
0bps	1.78%***	1.68%***	2.74%***	2.60%***	0.96%***	0.92%***
30bps	1.46%***	1.35%***	2.26%***	2.11%***	0.79%***	0.76%**
60bps	1.15%***	1.02%***	1.78%***	1.61%***	0.63%**	0.60%*

Table 6. Persistence of alpha of long-short portfolios based on composite return predictors

This table reports the monthly alpha of the long-short portfolios of stocks sorted by LCP and NLCP, for specific months after portfolio formation, ranging from the first month to the 24th month after portfolio formation. The first column reports the number of the month following portfolio formation. The second and third columns report the FF5+MOM and q5+MOM alphas of the LCP long-short portfolio, the fourth and fifth columns for the NLCP long-short portfolio, and the sixth and seventh columns for the difference between the returns of the NLCP and LCP long-short portfolios (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted excess returns on the long-short portfolio, for up to 24 months after portfolio formation, and report the alphas for each month separately. We compute Newey-West adjusted t-statistics and use ***, **, and * to indicate statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	LCP		NLCP		NLCP minus LCP	
	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM
m + 1	1.76%***	1.67%***	2.74%***	2.60%***	0.98%***	0.93%***
m + 2	1.59%***	1.43%***	2.28%***	2.11%***	0.69%***	0.68%**
m + 3	1.46%***	1.25%***	1.85%***	1.69%***	0.39%	0.44%
m + 4	1.43%***	1.23%***	1.85%***	1.65%***	0.42%	0.42%
m + 5	1.55%***	1.36%***	1.60%***	1.52%***	0.04%	0.16%
m + 6	1.31%***	1.10%***	1.56%***	1.60%***	0.25%	0.50%*
m + 7	1.55%***	1.39%***	1.86%***	1.88%***	0.31%	0.50%*
m + 8	1.35%***	1.11%**	1.81%***	1.78%***	0.46%*	0.67%***
m + 9	1.29%***	1.06%**	1.78%***	1.69%***	0.49%**	0.63%**
m + 10	0.97%***	0.73%*	1.88%***	1.92%***	0.91%***	1.18%***
m + 11	0.97%***	0.86%**	1.51%***	1.52%***	0.55%*	0.66%**
m + 12	0.84%**	0.62%	1.28%***	1.37%***	0.44%	0.75%**
m + 13	1.06%***	0.85%**	1.19%***	1.23%***	0.13%	0.38%
m + 14	0.95%***	0.62%	1.01%**	0.96%**	0.06%	0.33%
m + 15	0.79%**	0.46%	1.42%***	1.20%**	0.63%***	0.73%***
m + 16	0.80%**	0.45%	1.22%***	0.98%**	0.42%*	0.53%*
m + 17	0.75%**	0.42%	0.78%*	0.64%	0.03%	0.22%
m + 18	0.81%**	0.52%	0.97%*	0.87%	0.15%	0.34%
m + 19	0.77%*	0.49%	1.36%***	1.21%**	0.58%**	0.72%**
m + 20	0.64%*	0.33%	1.18%**	1.05%*	0.54%*	0.71%**
m + 21	0.59%	0.39%	1.04%**	0.97%*	0.45%	0.58%**
m + 22	0.63%*	0.42%	0.63%	0.54%	0.00%	0.12%
m + 23	0.66%*	0.49%	0.90%*	0.89%	0.24%	0.40%
m + 24	0.72%*	0.55%	1.09%***	1.07%**	0.37%	0.52%

Table 7. Fama-MacBeth regressions of monthly returns

This table reports the results for the Fama and MacBeth (1973) regressions of monthly stock returns on the LCP and NLCP indicator variables controlling for common firm characteristics. Each month, we sort sample stocks into decile portfolios based on their LCP and NLCP values. We then construct the indicator variables, I_LCP or I_NLCP, by setting them to one if the corresponding LCP or NLCP value falls within the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of stock returns on the indicator variables controlling for the eight firm characteristics defined in Panel C of Table 1: firm size (SIZE), book-to-market ratio (BM), momentum (MOM), short-term reversal (STR), asset growth (AG), gross profitability (GP), turnover (TO), and idiosyncratic volatility (IVOL). We winsorize characteristics at the 1st and 99th percentiles and standardize them to have zero mean and unit standard deviation. The first column reports the symbol of each explanatory variable, the second, third, and fourth columns report the results for the regressions that include only the I_LCP indicator variable, only the I_NLCP indicator variable, and both the I_LCP and I_NLCP indicator variables, respectively. We run the regressions for the period from 2013 to 2020, which coincides with the out-of-sample period we use to evaluate portfolio performance. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The coefficients are reported in percentage.

Variable	Ret	Ret	Ret
I_LCP	0.3233***		0.0006
	(3.02)		(0.01)
I_NLCP		0.7701***	0.7798***
		(5.03)	(4.77)
SIZE	-0.2483**	-0.2376*	-0.2312*
	(-2.05)	(-1.96)	(-1.93)
BM	-0.1462	-0.1551	-0.1574
	(-1.11)	(-1.17)	(-1.18)
MOM	-0.2008	-0.2124*	-0.2120*
	(-1.60)	(-1.68)	(-1.69)
STR	-0.2794***	-0.3019***	-0.2992***
	(-3.15)	(-3.41)	(-3.38)
AG	-0.0685	-0.0635	-0.0617
	(-1.11)	(-1.04)	(-1.01)
GP	0.1145	0.0989	0.0978
	(1.10)	(0.95)	(0.94)
TO	-1.1789***	-1.1399***	-1.1409***
	(-7.01)	(-6.79)	(-6.77)
IVOL	0.4563**	0.4592**	0.4604**
	(2.54)	(2.55)	(2.58)
Obs.	230,862	230,862	230,862
R-squared	0.04	0.04	0.04

Table 8. Trading- and holdings-based predictors

This table reports the results for the Fama and MacBeth (1973) regressions of monthly stock returns on indicator variables obtained using the LCP and NLCP based on only the combined trades (LCP_Trading and NLCP_Trading) or only the lagged combined holdings (LCP_Holding and NLCP_Holding) of all market participants. The indicator variables are equal to one if the composite predictor is in the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of next-month stock returns on the indicator variables controlling for the eight firm characteristics defined in Panel C of Table 1. We winsorize characteristics at the 1st and 99th percentiles and standardize them to have zero mean and unit standard deviation. The first column reports the symbol of each explanatory variable, the second third, fourth, and fifth columns report the results for the regressions that include (individually) the LCP_Trading, LCP_Holding, NLCP_Trading, or NLCP_Holding indicator variables, respectively, and the sixth column for the regression that includes all four indicator variables. We run the regressions for the period from 2013 to 2020, which coincides with the out-of-sample period we use to evaluate portfolio performance. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The coefficients are reported in percentage.

Variable	Ret	Ret	Ret	Ret	Ret
I_LCP_Trading	0.2125**				0.1254
	(2.12)				(1.07)
I_LCP_Holding		0.3261**			0.0962
		(2.46)			(0.96)
I_NLCP_Trading			0.4262***		0.3187**
			(3.17)		(2.21)
I_NLCP_Holding				0.5813***	0.4617***
				(3.66)	(3.20)
SIZE	-0.2823**	-0.2484**	-0.2898**	-0.2223*	-0.2547**
	(-2.32)	(-2.01)	(-2.45)	(-1.82)	(-2.14)
BM	-0.1411	-0.1452	-0.1446	-0.1475	-0.1536
	(-1.08)	(-1.11)	(-1.10)	(-1.11)	(-1.15)
MOM	-0.1822	-0.2158*	-0.1855	-0.2287*	-0.2122
	(-1.44)	(-1.68)	(-1.46)	(-1.79)	(-1.64)
STR	-0.2810***	-0.2773***	-0.2872***	-0.2830***	-0.2880***
	(-3.15)	(-3.12)	(-3.21)	(-3.22)	(-3.29)
AG	-0.0767	-0.0708	-0.0729	-0.0705	-0.0625
	(-1.23)	(-1.16)	(-1.17)	(-1.15)	(-1.04)
GP	0.1219	0.1092	0.1154	0.0953	0.0871
	(1.17)	(1.05)	(1.09)	(0.91)	(0.83)
TO	-1.1910***	-1.1694***	-1.1689***	-1.1533***	-1.1365***
	(-7.11)	(-6.95)	(-7.06)	(-6.81)	(-6.74)
IVOL	0.4347**	0.4636**	0.4542**	0.4646***	0.4796***
	(2.40)	(2.58)	(2.48)	(2.64)	(2.75)
Obs.	230,862	230,862	230,862	230,862	230,862
R-squared	0.04	0.04	0.04	0.04	0.05

Table 9. Portfolio performance and firm size

This table reports the alpha of the portfolios of stocks sorted by the LCP (or NLCP) across sub-samples of stocks categorized by firm size. At the end of each month, we rank all sample stocks based on firm size and sort them into tercile portfolios. We then rank stocks within each tercile based on either the LCP or NLCP, and sort them into decile portfolios. We then compute the equal-weighted returns on each decile as well as on the long-short portfolio and report the alphas with respect to the five Fama and French (2015) and momentum factors (FF5+MOM). The first column reports the label for each decile portfolio and the long-short portfolio, the second and third columns report the alphas for small firms, the fourth and fifth columns for medium firms, and the last two columns for large firms. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	Small		Medium		Large	
	LCP	NLCP	LCP	NLCP	LCP	NLCP
L	-2.02%***	-2.39%***	-0.61%**	-1.00%***	-0.39%**	-0.29%**
	(-3.31)	(-3.79)	(-2.54)	(-4.03)	(-2.55)	(-2.05)
2	-0.61%	-0.87%*	-0.18%	-0.13%	-0.16%	-0.32%*
	(-1.32)	(-1.66)	(-1.07)	(-0.92)	(-1.22)	(-1.95)
3	0.17%	0.28%	-0.14%	0.34%**	0.14%	0.01%
	(0.38)	(0.88)	(-0.75)	(2.32)	(1.08)	(0.06)
4	0.68%*	0.43%	-0.04%	0.12%	-0.25%**	-0.11%
	(1.77)	(1.47)	(-0.34)	(0.88)	(-2.06)	(-0.94)
5	0.93%***	0.47%	0.08%	-0.18%	0.12%	0.02%
	(3.13)	(1.62)	(0.62)	(-1.16)	(0.91)	(0.14)
6	0.90%***	0.59%**	0.17%	0.36%	-0.12%	0.00%
	(3.50)	(2.11)	(0.96)	(1.56)	(-1.07)	(0.03)
7	0.75%**	0.70%**	0.04%	-0.22%	0.04%	-0.01%
	(2.01)	(2.39)	(0.21)	(-1.50)	(0.50)	(-0.22)
8	0.92%**	1.20%***	0.28%*	0.02%	-0.24%***	0.02%
	(2.39)	(2.92)	(1.86)	(0.12)	(-4.95)	(0.50)
9	0.67%**	1.95%***	0.05%	0.45%***	-0.07%	0.04%
	(2.39)	(5.07)	(0.41)	(3.52)	(-0.69)	(0.42)
H	1.25%***	1.28%***	0.42%*	0.29%	0.26%***	-0.02%
	(3.04)	(3.14)	(1.86)	(1.53)	(3.13)	(-0.22)
H-L	3.27%***	3.68%***	1.03%***	1.29%***	0.65%***	0.26%*
	(8.55)	(7.03)	(3.44)	(5.28)	(3.39)	(1.82)

Table 10. Fama-MacBeth regressions: interactions with firm size

This table reports the results for the Fama and MacBeth (1973) regressions of monthly stock returns on the LCP and NLCP indicator variables after including interaction between firm size and the LCP and NLCP indicator variables. Each month, we sort sample stocks into decile portfolios based on their LCP and NLCP values. We then construct the indicator variables, I_LCP or I_NLCP, by setting them to one if the corresponding LCP or NLCP value falls within the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of stock returns on the indicator variables and their interaction terms with firm size, while also controlling for the eight firm characteristics defined in Panel C of Table 1. We winsorize characteristics at the 1st and 99th percentiles and standardize them to have zero mean and unit standard deviation. The first column reports the symbol of each explanatory variable, the second, third, and fourth columns report the results for the regressions that include the I_LCP indicator variable and its interaction term with size, the I_NLCP indicator variable and its interaction term with size, and both the I_LCP and I_NLCP indicator variables and their interaction terms with size, respectively. We run the regressions for the period from 2013 to 2020, which coincides with the out-of-sample period we use to evaluate portfolio performance. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The coefficients are reported in percentage.

Variable	Ret	Ret	Ret
I_LCP	0.3220***		0.0278
	(3.21)		(0.26)
I_LCP * SIZE	-0.0765		0.1881
	(-0.58)		(1.11)
I_NLCP		0.6119***	0.5963***
		(4.77)	(4.97)
I_NLCP * SIZE		-0.3242**	-0.4166**
		(-2.60)	(-2.52)
SIZE	-0.2471**	-0.2291*	-0.2299*
	(-2.00)	(-1.89)	(-1.88)
BM	-0.1472	-0.1621	-0.1612
	(-1.13)	(-1.21)	(-1.20)
MOM	-0.1985	-0.2162*	-0.2109*
	(-1.59)	(-1.69)	(-1.67)
STR	-0.2788***	-0.3068***	-0.3012***
	(-3.17)	(-3.47)	(-3.42)
AG	-0.0685	-0.0614	-0.0588
	(-1.10)	(-1.01)	(-0.96)
GP	0.1122	0.0896	0.0908
	(1.08)	(0.86)	(0.87)
TO	-1.1801***	-1.1261***	-1.1353***
	(-6.95)	(-6.77)	(-6.74)
IVOL	0.4563**	0.4587**	0.4570**
	(2.53)	(2.53)	(2.55)
Obs.	230,862	230,862	230,862
R-squared	0.04	0.05	0.05

Table 11. Effect of information environment

This table reports the results for Fama and MacBeth (1973) regressions for sub-samples of stocks categorized based on five firm characteristics that proxy information asymmetry: firm age, idiosyncratic volatility, illiquidity, analyst coverage, and analyst disagreement. Firm age is the number of years since the firm was first covered by CRSP, idiosyncratic volatility as the standard deviation of the residuals from regressing daily excess stock returns on the three-factor Fama and French (1992) model over the previous six months, illiquidity is the measure of Amihud (2002), analyst coverage as the number of analysts with valid earnings per share forecast as of prior month-end, and analyst disagreement as analyst forecast dispersion (standard deviation of analyst forecasts divided by absolute value of average analyst forecast in the prior month). Each month and within each stock sub-sample, we sort stocks into decile portfolios based on their LCP or NLCP values. We then construct the indicator variables, I_LCP or I_NLCP, by setting them to one if the corresponding LCP or NLCP value falls within the high decile, minus one if it is in the low decile, and zero otherwise. We then run monthly regressions of stock returns on I_LCP and I_NLCP, and the eight firm characteristics defined in Panel C of Table 1. We winsorize characteristics at the 1st and 99th percentiles and standardize them to have zero mean and unit standard deviation. The first column reports the symbol of each explanatory variable, the rest of the columns report the regression results for each stock sub-sample. We run the regressions for the period from 2013 to 2020, which coincides with the out-of-sample period we use to evaluate portfolio performance. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The coefficients are reported in percentage.

	Firm age		Idio. vol.		Illiquidity		Analyst coverage		Analyst disagreement	
	Young	Old	Low	High	Low	High	Low	High	Low	High
	Ret	Ret	Ret	Ret	Ret	Ret	Ret	Ret	Ret	Ret
I_LCP	0.066	-0.056	-0.027	-0.199	-0.007	-0.025	-0.213	0.145	0.024	-0.025
	(0.37)	(-0.42)	(-0.26)	(-0.91)	(-0.08)	(-0.11)	(-0.99)	(1.39)	(0.17)	(-0.19)
I_NLCP	0.801***	0.807***	0.317***	1.005***	0.239*	0.879***	0.996***	0.181	0.430**	0.687***
	(4.02)	(3.67)	(2.71)	(4.09)	(1.84)	(4.45)	(4.41)	(1.14)	(2.62)	(2.82)
Character.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	103,264	127,598	122,333	108,529	123,587	107,247	102,829	94,004	83,219	78,427
R-squared	0.05	0.05	0.05	0.05	0.08	0.05	0.05	0.10	0.06	0.08

Table 12. Predicting operating performance and earnings announcement returns

This table reports the results for the Fama and MacBeth (1973) regressions of future operating performance and earning announcement returns on the indicator variables I_LCP and I_NLCP, controlling for firm characteristics and current measures of operating performance. We use four measures of operating performance: cash flows (CF), which is the difference between income before extraordinary items and total accruals, divided by total assets, gross margin (GM), which is is sales minus cost of goods sold scaled by current sales, return-on-equity (ROE), which is the sum of income before extraordinary items and interest expenses, divided by the lagged total equity, and return-on-asset (ROA), which is the sum of income before extraordinary items and interest expenses, divided by the lagged total assets. Earnings announcement returns (CAR) are cumulative abnormal returns in the $[-1, 1]$ three-day window around the earnings announcement day, where abnormal returns are the difference between the daily stock return and that of the corresponding portfolio among the six size and book-to-market Fama-French portfolios. We then run monthly regressions of the four measures of future operating performance and earnings announcement returns on I_LCP or I_NLCP, the eight firm characteristics defined in Panel C of Table 1, and current measures of operating performance (CF0, GM0, ROE0, ROA0). We run the regressions for the period from 2013 to 2020, which coincides with the out-of-sample period we use to evaluate portfolio performance. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The coefficients for the four measures of operating performance and earnings announcement returns reported in percentage.

Variable	CF	CF	GM	GM	ROE	ROE	ROA	ROA	CAR	CAR
I_LCP	0.476***		0.263**		0.487		0.493**		0.06	
	(3.70)		(2.35)		(1.52)		(2.53)		(0.72)	
I_NLCP		0.457***		0.253**		0.777***		0.626**		0.28***
		(6.09)		(2.78)		(3.30)		(2.10)		(3.39)
CF0	33.186***	33.181***								
	(90.62)	(91.98)								
GM0			45.488***	45.484***						
			(372.73)	(374.20)						
ROE0					67.562***	67.537***				
					(49.59)	(49.65)				
ROA0							44.324***	44.307***		
							(17.57)	(17.55)		
Character.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	70,648	70,648	75,254	75,254	73,419	73,419	68,130	68,130	203,949	203,949
R-squared	0.91	0.91	0.96	0.96	0.91	0.91	0.89	0.89	0.01	0.01

Table 13. Average anomalies for portfolios of stocks sorted by composite return predictors

Panels A and B of this table report the average of each stock anomaly for the portfolios of stocks sorted by the LCP and NLCP. For each panel, the first column reports the name of the anomalies, the second to eighth columns report the averaged anomaly value of the low, second, third, eighth, ninth, high decile portfolios, and the long-short portfolio, respectively. The last column reports the t-statistics for the long-short portfolio anomalies. At the end of each month, we rank all sample stocks based on either the LCP or NLCP and then sort them into decile portfolios. We then compute the time-series average of the cross-sectional means of each stock anomaly on each decile as well as on the long-short portfolio. Out of the 102 anomalies documented in Green et al. (2017), we select 12 anomalies whose return is significant (t-stat >1.66) for our sample period from 2008 to 2020. For easier interpretation, we change the sign of anomalies whose return is negative, including market capitalization, idiosyncratic volatility, momentum_1m, and asset growth, so that the mean anomaly values for the LCP and NLCP long-short portfolios should be positive if they select stocks on the right side of the anomaly. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	L	2	3	8	9	H	H - L	t-stat
Panel A: Anomalies for the portfolios of stocks sorted by LCP								
(-) Market capitalization	12.858	13.420	13.613	13.778	13.536	13.107	0.25**	2.23
Book-to-market	0.423	0.441	0.457	0.534	0.551	0.609	0.19***	14.97
Gross margin	0.232	0.298	0.321	0.350	0.353	0.349	0.12***	10.80
Illiquidity	0.144	0.207	0.295	0.915	1.117	1.246	1.10***	8.60
(-) Idiosyncratic volatility	0.052	0.050	0.051	0.060	0.065	0.078	0.03***	13.33
Momentum_12m	0.019	0.061	0.092	0.100	0.098	0.068	0.05**	2.29
(-) Momentum_1m	0.017	0.016	0.014	0.008	0.005	-0.002	-0.02*	-1.98
(-) Asset growth	0.051	0.079	0.092	0.143	0.161	0.182	0.13***	16.84
Dividend yield	0.010	0.011	0.012	0.011	0.011	0.010	0.00	0.37
Analyst coverage	6.991	7.968	8.533	7.596	6.819	5.106	-1.89***	-6.82
Price delay	0.057	0.042	0.038	0.053	0.053	0.066	0.01	1.03
Combined fundamental	3.871	4.174	4.285	4.286	4.203	3.974	0.10*	1.90
Panel B: Anomalies for the portfolios of stocks sorted by NLCP								
(-) Market capitalization	12.078	13.086	13.733	14.003	13.641	12.330	0.25**	2.27
Book-to-market	0.492	0.474	0.458	0.499	0.565	0.676	0.18***	11.48
Gross margin	0.177	0.336	0.351	0.343	0.342	0.319	0.14***	16.27
Illiquidity	0.385	0.358	0.276	0.739	1.313	2.365	1.98***	8.10
(-) Idiosyncratic volatility	0.066	0.057	0.052	0.052	0.059	0.088	0.02***	12.13
Momentum_12m	-0.037	0.082	0.109	0.106	0.080	0.032	0.07***	2.65
(-) Momentum_1m	0.027	0.017	0.015	0.008	0.003	-0.013	-0.04***	-4.11
(-) Asset growth	0.056	0.091	0.103	0.138	0.157	0.147	0.09***	12.35
Dividend yield	0.009	0.011	0.011	0.011	0.011	0.012	0.00***	2.77
Analyst coverage	5.075	7.903	8.691	7.948	6.352	3.801	-1.27***	-4.81
Price delay	0.081	0.038	0.029	0.049	0.078	0.134	0.05***	5.05
Combined fundamental	3.514	4.260	4.439	4.297	4.072	3.632	0.12**	2.42

Table 14. Performance of factor models based on composite return predictors

This table reports the performance of several factor models in terms of their ability to explain anomaly returns and the Sharpe ratio generated by their factors. We consider four factor models: (i) the five Fama and French (2015) and momentum factors (FF5+MOM), (ii) the five Hou et al. (2021) and momentum factors (q5+MOM), (iii) a two-factor model including the market factor and the LCP factor defined as the return of the LCP equal-weighted long-short portfolio (LCP+MKT), and (iv) a two-factor model including the market factor and the NLCP factor defined as the return of the NLCP equal-weighted long-short portfolio (NLCP+MKT). Out of the 102 anomalies documented in Green et al. (2017), we select the 12 anomalies whose return is significant (t-stat >1.66) for our sample period from 2008 to 2020. We run time series regressions of the 12 anomaly returns on each of the four factor models. We run time series regressions of the returns of the 12 anomalies on each of the four factor models. The first and second rows in the table report the average (across the 12 anomalies) of the absolute value of the alpha and alpha t-stat. The third row in the table reports the annualized Sharpe ratio of the mean-variance portfolio of the factors in each model, and the fourth and fifth rows report the p-value for the difference between the Sharpe ratio of each model and that of the FF5+MOM and q5+MOM models.

	FF5+MOM	q5+MOM	LCP+MKT	NLCP+MKT
Average absolute alpha	0.59	0.44	0.46	0.50
Average absolute t-stat	2.03	1.41	1.19	1.16
Annualized Sharpe ratio (SR)	1.13	1.71	2.32	3.09
P-value of SR difference compared to FF5+MOM	NA	0.18	0.30	0.02
P-value of SR difference compared to q5+MOM	0.17	NA	0.55	0.05

Appendix A Predicting variables and models

This appendix provides additional details about the predicting variables and models we consider.

A.1 Predicting variables

As described in Section 2.1, the predicting variables consist of quarterly trades and lagged quarterly holdings of nine types of market participants, including six types of institutions that report their holdings on 13F form filings as well as firms, short sellers, and retail investors.

We gather institutional holdings data from *Thomson/Refinitiv S12* and *13F*, following the approach of McLean et al. (2025) to categorize institutions into six types. First, we combine mutual fund holdings reported in S12 form filings with their corresponding 13F form filings, treating the reported number of shares by mutual funds as their raw holdings. Second, we employ the type codes developed by Brian Bushee (available at https://accounting-faculty.wharton.upenn.edu/bushee/) to identify institutions classified as banks or insurance companies, extracting their corresponding holdings. Third, we use the search scheme in McLean et al. (2025) to identify wealth management firms and hedge funds. For wealth management firms, we conduct case-insensitive searches for terms such as "wealth manag", "wealth MGNT", "private", "PRVT" and "advisor," and for hedge funds for terms like "LLC", "L.L.C.", "L L C", "L. L. C.", "LP", "L.P", "L P", "L. P", or "partner". Fourth, we designate the remaining institutions as "other" institutions and acquire their reported holdings. For all six types of institutions, we scale their quarterly lagged holdings by the total number of shares outstanding to obtain: *bank holdings*, *hedge fund holdings*, *insurance co. holdings*, *mutual fund holdings*, *other inst. holdings*, and *wealth mgmt. holdings*. We calculate quarterly trades as the changes in quarterly holdings to obtain: *bank trades*, *hedge fund trades*, *insurance co. trades*, *mutual fund trades*, *other inst. trades*, and *wealth mgmt. trades*.

We use the approach proposed by Barber et al. (2024) to measure retail trades.[22] First,

[22] Barber et al. (2024) modify the approach of Boehmer et al. (2021) for identifying and signing retail trades. They argue that retail orders that are internalized or executed by wholesalers often receive a small amount of price improvement relative to the quote midpoint, allowing for separating retail orders from institutional ones based on the sub-penny pricing of execution.

we acquire daily off-exchange marketable orders from the *Trade and Quote (TAQ)* dataset and exclude trades executed between 40% and 60% of the National Best Bid or Offer (NBBO). We then label a transaction as a retail buy if its execution price surpasses the quote midpoint, and classify it as a retail sell if the price falls below the quote midpoint. Next, we aggregate intraday retail buy and sell volumes for each stock on each trading day and calculate the order imbalance for the ith stock on day t as:

$$\text{oib}_{i,t} = \frac{\text{indbvol}_{i,t} - \text{indsvol}_{i,t}}{\text{indbvol}_{i,t} + \text{indsvol}_{i,t}}, \tag{A.1}$$

where $\text{indbvol}_{i,t}$ and $\text{indbvol}_{i,t}$ are the intraday retail buy and sell order volumes. We then compute monthly retail trades as the average daily order imbalance for each month.

To compute short seller trades and holdings, we acquire quarter-end short-interest data from *Compustat* and normalize it by the number of shares outstanding to obtain the short-interest ratio. Following McLean et al. (2025), we adjust the sign of the short-interest ratio so that larger values imply smaller short positions. Next, we compute short seller trades as the change in signed short interest from the previous quarter. Thus, increases in the ratio indicate negative short seller trades, while decreases (reflecting a net closing of short positions) translate into positive values of short seller trades. To measure firm trading, we subtract quarter-end share repurchases from share issues, scaled by the number of shares outstanding. We also sign this variable so that decreases (or increases) in shares outstanding denote positive (or negative) values of firm trading. Panel A of Table 1 inf the main body of the manuscript provides the definition for the 16 predicting variables.

A.2 Predicting models

We consider 16 models used in the recent asset-pricing literature (Rapach et al., 2019; Gu et al., 2020; DeMiguel et al., 2023) to predict out-of-sample excess returns using the trades and lagged holdings of the nine market participants as predicting variables. The models can be categorized as linear or nonlinear.

The set of hyper-parameters we consider for each model is described in the remainder of this section and it is similar to that considered by Gu et al. (2020), with minor adjustments

to account for our using a smaller number of predicting variables. Following DeMiguel et al. (2023) and the references therein, we select hyper-parameters using five-fold cross validation (Hastie et al., 2009), that is, we divide the sample into five *folds* and then remove each *fold* in turn and use it to evaluate the estimation errors associated with estimating the models in the other four folds using different sets of hyper-parameters. We choose the value of hyper-parameters that minimizes the average estimation error. We use the *sklearn* Python package to train the models.

A.2.1 Linear models

Ordinary least squares (OLS)

We consider a panel with N stocks and T time periods, as well as p predicting variables. The first linear model we consider is OLS, which is a good benchmark because of its widespread use and interpretability. The goal is to select the parameter vector that minimizes the quadratic loss function:

$$\min_{\beta,\zeta} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1}-\zeta-\sum_{j=1}^{p}\beta_j x_{ij,t}\right)^2, \tag{A.2}$$

where $r_{i,t+1}$ denotes the excess return of the ith stock in month $t+1$, $x_{ij,t}$ is the jth predicting variable for the ith stock in month t, ζ is the intercept, and β_j is the slope for the jth predicting variable. A drawback of OLS is that it may overfit the data when there are many predicting variables.

Ridge regression (Ridge)

A popular approach to combat overfitting is to use shrinkage to obtain smaller slope estimates and thus avoid spurious large estimators (Gu et al., 2020). Ridge regression, in particular, improves forecast accuracy by trading off a small increase in estimation bias for a large reduction in estimation variance. Ridge regression shrinks the regression coefficients via parameter penalization (Hastie et al., 2009), where the goal is to minimize a penalized sum of squared residuals:

$$\min_{\beta,\zeta} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1}-\zeta-\sum_{j=1}^{p}x_{ij,t}\beta_j\right)^2+\lambda\sum_{j=1}^{p}\beta_j^2, \tag{A.3}$$

in which λ is the shrinkage intensity. We consider a single hyper-parameter, the shrinkage intensity, which we select using five-fold cross validation from a grid of 50 different values in the interval $[0.0001, 0.1]$.

Adaptive least absolute shrinkage and selection operator (ALasso)

Another common approach to alleviate overfitting is least absolute shrinkage and selection operator (lasso); see for instance Tibshirani (1996) and Rapach et al. (2019). In short, lasso performs both shrinkage and variable selection, by minimizing the following penalized loss function:

$$\min_{\beta,\zeta} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1} - \zeta - \sum_{j=1}^{p} x_{ij,t}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j|, \tag{A.4}$$

where λ is the shrinkage parameter and the ridge penalty $\sum_{j=1}^{p}\beta_j^2$ in Problem (A.3) is replaced by the lasso penalty $\sum_{j=1}^{p}|\beta_j|$ in Problem (A.4).

Following Zou (2006), we use the *adaptive* lasso to further improve estimation. Specifically, we assign a weight $w_j = 1/|\hat{\beta}_j|$, where $\hat{\beta}_j$ is the OLS slope for the jth predicting variable, to the jth slope in the regularization term:

$$\min_{\beta,\zeta} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1} - \zeta - \sum_{j=1}^{p} x_{ij,t}\beta_j\right)^2 + \lambda\sum_{j=1}^{p} w_j|\beta_j|. \tag{A.5}$$

We select the shrinkage intensity λ using five-fold cross validation from a grid of 50 different values in the interval $[0.0001, 0.1]$.

Elastic net (ENet)

We also use the elastic net model (Gu et al., 2020; DeMiguel et al., 2023), which generalizes ridge and lasso regression:

$$\min_{\beta,\zeta} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1} - \zeta - \sum_{j=1}^{p} x_{ij,t}\beta_j\right)^2 + \lambda\rho\sum_{j=1}^{p}|\beta_j| + \lambda(1-\rho)\sum_{j=1}^{p}\beta_j^2, \tag{A.6}$$

where $\sum_{j=1}^{p}|\beta_j|$ is the 1-norm and $\sum_{j=1}^{p}\beta_j^2$ is the 2-norm of the parameter vector. The 1-norm term controls the sparsity in the estimated parameter vector β, while the 2-norm

term enhances its stability. The ridge regression in (A.3) and the lasso regression in (A.4) are obtained as particular cases when $\rho = 0$ and when $\rho = 1$. Other values of ρ lead to a combination of ridge and lasso penalties. We set the hyper-parameter ρ to be 0.5 and λ from a grid of 50 different values in the interval $[0.0001, 0.1]$.

Principal components regression (PCR)

When predictors are highly correlated, regularization models such as lasso and elastic net can produce suboptimal forecasts. To address this challenge, we consider the dimension-reduction approach known as principal components regression (PCR), which is carried out in two steps. First, principal components analysis (PCA) is used to select a small number K of linear combinations of the p predictors that capture most of the covariation among the predictors. Second, OLS regression is performed using the K principal components as explanatory variables. We set the number of components $K = 4$.

Partial least squares (PLS)

A disadvantage of PCR is that the target variable (excess returns in our case) is not considered when selecting the principal components. To address this limitation, we also consider partial least squares (PLS), which accounts for the correlation between the target and predicting variables when performing the dimension reduction step (Gu et al., 2020). For each predicting variable, PLS first estimates a univariate OLS regression of the target variable on the jth predictor, obtaining an individual slope coefficient ϕ_j. Then, the first component is produced by weighting each predictor proportionally to its slope coefficient ϕ_j, that is, giving more weight to stronger predictors and less to weaker ones. To construct additional components, both the target variable and predictors are orthogonalized with respect to the components already constructed. The above steps are then repeated on the orthogonalized data until the desired number of components is reached. We set the number of components $K = 4$.

A.2.2 Nonlinear models

We consider three types of predicting models that capture nonlinearities and interactions in the relation between the trades and holdings of market participants and future excess returns.

Gradient boosted regression trees (GBRT)

Regression trees are non-parametric models that partition the data into groups of observations with similar characteristics. Within each group, the forecast is simply the average excess return. While regression trees are capable of capturing nonlinearities and interactions, they often suffer from overfitting.

To address overfitting, we use gradient boosted regression trees (Gu et al., 2020; DeMiguel et al., 2023), which build an ensemble of shallow trees in a sequence. The model starts with a shallow tree (for instance with depth $L = 1$), which typically provides a weak prediction with high bias. A second shallow tree (with the same depth L) is then fitted to the residuals of the first tree and added to the prediction, scaled by a learning rate ν to mitigate overfitting. Forecasts from these two trees are added together to form an ensemble prediction of the outcome. This procedure is iterated until there are a total of N trees in the ensemble.

The final output is an additive combination of shallow trees, controlled by three main hyper-parameters: the depth of each tree (L), the learning rate (ν), and the number of trees (N). In our implementation, we consider learning rates $\nu \in \{0.01, 0.1\}$, tree depths $L \in \{1, 2\}$, and number of trees $N = 500$. By focusing sequentially on correcting errors and employing regularization techniques, GBRT improves the forecasting performance by reducing the prediction variance and bias (Gu et al., 2020; DeMiguel et al., 2023; Schapire and Freund, 2012).

Random forest (RF)

Similar to gradient boosting, random forest combines predictions from multiple decision trees. One key distinction between the two methods is that RF aggregates independent decision trees through bootstrap aggregation, or "bagging" (Breiman, 2001), while GBRT aggregates the trees sequentially. Specifically, random forests generate multiple bootstrap samples from the original dataset and fit a separate regression tree to each. The final prediction is obtained by averaging the forecasts across all trees, which helps to reduce variance.

To further lower the correlation among trees—a common limitation of bagging—random forests introduce an additional layer of randomness. At each node split, only a random subset of predictor variables is considered, preventing dominant predictors from driving early splits

across all trees and reducing the average correlation among trees. In our setting, we follow Gu et al. (2020) and set the number of trees to 300. The maximum tree depth is allowed to vary between one and six. For each node split, the number of predictors considered is selected from the set {1, 2, 3, 5, 10}.

Artificial neural networks (ANN)

Lastly, we employ artificial neural networks, which are highly flexible tools for modeling complex nonlinear relationships. The ANNs consist of multiple layers of interconnected units that transform predictor inputs into forecasts through layered nonlinear interactions.

We focus on commonly-used "feed-forward" neural networks, which include an input layer for raw predictors, one or more hidden layers to model interactions, and an output layer to produce forecasts (Gu et al., 2020; Kaniel et al., 2023). Each layer is composed of artificial "neurons" that apply a transformation to signals and pass the result to the next layer through weighted connections.

To evaluate the impact of network depth, we consider networks with up to four hidden layers (i.e., ANN1, ANN2, ANN3, ANN4). For nonlinear activation functions, we consider the hyperbolic tangent (Tanh) and the rectified linear unit (ReLU):

$$Tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad \text{and} \tag{A.7}$$

$$ReLU(x) = \max(0, x). \tag{A.8}$$

The hyperbolic tangent in (A.7) helps center the data to facilitate learning for the next layer, and the rectified linear unit in (A.8) introduces sparsity by deactivating neurons with negative input, improving computational efficiency.

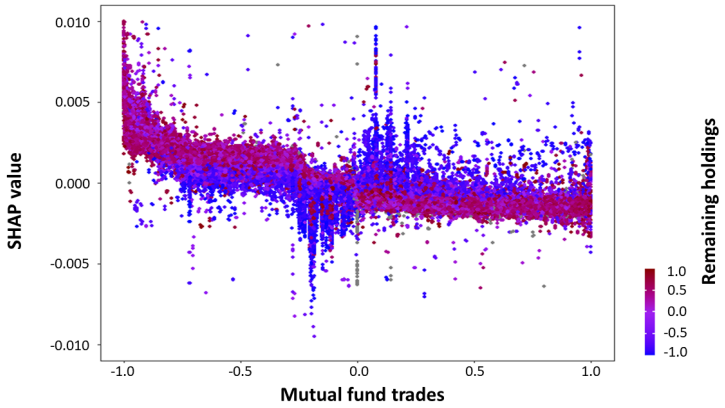
We consider learning rates in the set $\{0.00001, 0.0001, 0.001\}$. For the size of each hidden layer, we consider a grid of possible values. For ANN1, we consider a single hidden layer with 16, 8, or 4 neurons. For ANN2, we consider two hidden layers with number of neurons selected from (16, 8), (8, 4), or (4, 2). For ANN3, we consider three hidden layers, with number of neurons selected from (16, 8, 4) or (8, 4, 2). Finally, for ANN4 we consider four hidden layers with 16, 8, 4, and 2 neurons.

Appendix B Additional figures and tables

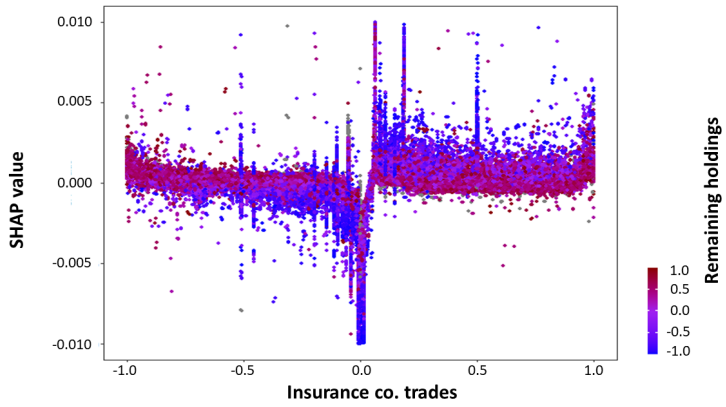
Figure B.1. Interactions between trading and holdings of remaining institutions

This figures illustrates the interactions between the trading of three types of institutions and the holdings of the remaining institutions. Panel A depicts the interaction between mutual fund trading and the aggregate holdings of institutional investors other than mutual funds, Panel B between insurance company trading and the holdings of the remaining institutions (other than insurance companies), and Panel C between wealth management firm trading and the holdings of the remaining institutions (other than wealth management firms). For each panel, the horizontal axis depicts the (standardized) trading variable and the vertical axis its SHAP value for each observation (dots). To visualize the interaction, we use color to display the original value of the holding of the remaining institutions (red for high value and blue for low value). We compute SHAP values for the last estimation window, which spans the period from December 2015 to November 2020.

Panel A. Mutual trading vs. remaining holding



Panel B. Insur. trading vs. remaining holding



Panel C. Wealth mgmt. trading vs. remaining holding

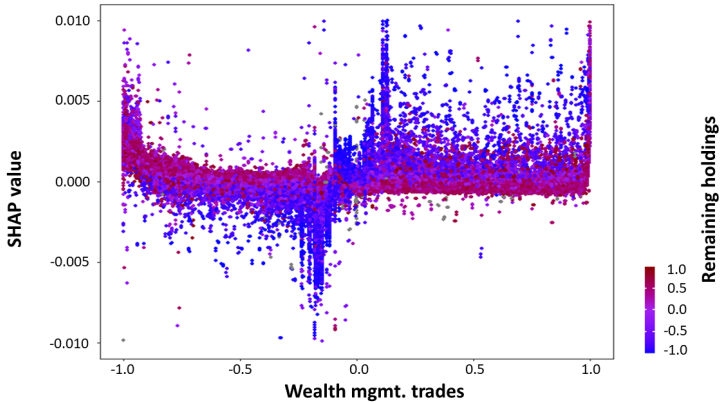


Table B.1. Excess return of univariate portfolios of stocks sorted by trades or lagged holdings

This table reports the monthly excess return of univariate portfolios obtained by sorting stocks based on the trades (or lagged holdings) of a single type of market participant. Panels A and B report the results for the univariate portfolios constructed using trades and lagged holdings, respectively. For each panel, the first column reports the name of the sorting variable, the second to seventh columns report the alpha of the low, second, third, eighth, ninth, and high decile portfolios, respectively, and the eighth column reports the alpha for the long-short portfolio, which goes long stocks in the high decile and short stocks in the low decile. At the end of each month, we rank all sample stocks based on either the trades or the lagged holdings of each market participant and then sort them into decile portfolios. We calculate stock excess returns by subtracting the one-month risk-free rate from raw returns. We then compute equal-weighted excess returns on each decile portfolio, as well as on the long-short portfolio. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. To facilitate the comparison, we evaluate performance in the period from 2013 to 2020, which coincides with the out-of-sample period for the evaluation of the composite return predictors.

	L	2	3	8	9	H	**H - L**
Panel A: Trade sorts							
Banks	1.59%**	1.54%**	1.47%**	1.38%**	1.35%**	1.13%*	-0.45%**
	(2.32)	(2.47)	(2.44)	(2.29)	(2.12)	(1.68)	(-2.30)
Firms	1.43%*	1.42%**	1.73%***	1.05%*	1.45%***	1.21%**	-0.23%
	(1.79)	(2.09)	(2.76)	(1.78)	(2.65)	(2.00)	(-0.49)
Hedge funds	1.27%*	1.37%**	1.39%**	1.47%**	1.61%**	1.57%**	0.30%*
	(1.76)	(2.06)	(2.33)	(2.46)	(2.49)	(2.37)	(1.68)
Insurance co.	1.60%**	1.40%**	1.39%**	1.28%**	1.15%*	1.31%**	-0.29%*
	(2.33)	(2.27)	(2.30)	(2.05)	(1.91)	(2.05)	(-1.85)
Mutual funds	1.84%**	1.57%**	1.62%**	1.20%*	1.17%*	1.27%*	-0.57%***
	(2.55)	(2.45)	(2.60)	(1.98)	(1.81)	(1.91)	(-2.64)
Other inst.	1.40%*	1.45%**	1.29%**	1.46%**	1.45%**	1.77%**	0.36%*
	(1.91)	(2.24)	(2.14)	(2.46)	(2.27)	(2.57)	(1.69)
Short sellers	1.04%	1.28%*	1.35%**	1.69%***	1.29%*	1.18%	0.15%
	(1.31)	(1.90)	(2.23)	(2.70)	(1.98)	(1.61)	(0.59)
Wealth mgmt.	1.71%**	1.58%**	1.52%***	1.36%**	1.19%*	1.40%*	-0.31%*
	(2.30)	(2.45)	(2.71)	(2.35)	(1.94)	(1.90)	(-1.88)
Retail	0.99%*	1.32%**	1.33%**	1.33%*	1.60%**	1.87%***	0.88%***
	(1.97)	(2.36)	(2.24)	(1.88)	(2.17)	(2.85)	(2.72)
Panel B: Lagged-holdings sorts							
Banks	1.05%	2.10%***	1.93%***	1.23%**	1.33%**	1.15%**	0.10%
	(1.40)	(2.93)	(2.96)	(2.04)	(2.31)	(2.02)	(0.21)
Hedge funds	1.07%	1.83%***	1.48%**	1.42%**	1.47%**	1.34%**	0.28%
	(1.41)	(2.72)	(2.44)	(2.27)	(2.29)	(2.09)	(0.72)
Insurance co.	2.03%*	2.11%**	1.66%**	1.23%**	1.28%**	1.25%**	-0.35%
	(1.72)	(2.40)	(2.61)	(2.10)	(2.34)	(2.06)	(-0.86)
Mutual funds	0.86%	2.24%***	1.71%***	1.28%**	1.25%**	1.22%*	0.36%
	(1.10)	(3.28)	(2.68)	(2.16)	(2.07)	(1.89)	(0.73)
Other inst.	1.07%	1.57%***	1.38%**	1.41%**	1.60%**	1.50%**	0.42%
	(1.41)	(2.71)	(2.29)	(2.28)	(2.58)	(2.12)	(0.96)
Short sellers	1.25%	1.18%	1.10%	1.50%***	1.49%**	1.95%***	0.71%
	(1.54)	(1.62)	(1.62)	(2.70)	(2.55)	(3.73)	(1.56)
Wealth mgmt.	1.02%	2.19%***	1.85%***	1.27%**	1.25%*	1.25%*	0.24%
	(1.34)	(2.94)	(2.98)	(2.00)	(1.90)	(1.83)	(0.49)

Table B.2. Alpha (q5+MOM) of univariate portfolios sorted by trades or lagged holdings

This table reports the alpha with respect to the Hou et al. (2021) and momentum factors of univariate portfolios obtained by sorting stocks based on the trades (or lagged holdings) of a single type of market participant. Panels A and B report the results for the univariate portfolios constructed using trades and lagged holdings, respectively. For each panel, the first column reports the name of the sorting variable, the second to seventh columns report the alpha of the low, second, third, eighth, ninth, and high decile portfolios, respectively, and the eighth column reports the alpha for the long-short portfolio, which goes long stocks in the high decile and short stocks in the low decile. At the end of each month, we rank all sample stocks based on either the trades or the lagged holdings of each market participant and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile as well as on the long-short portfolio and report their alphas. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. To facilitate the comparison, we evaluate performance in the period from 2013 to 2020, which coincides with the out-of-sample period for the evaluation of the composite return predictors.

	L	2	3	8	9	H	**H - L**
Panel A: Trade sorts							
Banks	0.46%***	0.35%**	0.39%***	0.24%***	0.17%	-0.03%	-0.49%**
	(2.72)	(2.44)	(2.92)	(2.64)	(1.27)	(-0.29)	(-2.42)
Firms	0.24%	0.23%	0.63%***	0.37%*	0.36%***	0.19%	-0.04%
	(0.55)	(1.03)	(3.78)	(1.78)	(2.86)	(1.17)	(-0.09)
Hedge funds	0.22%*	0.18%	0.31%**	0.36%***	0.41%***	0.56%***	0.34%*
	(1.79)	(1.27)	(2.08)	(3.40)	(3.90)	(2.95)	(1.80)
Insurance co.	0.38%***	0.36%***	0.39%***	0.13%	0.06%	0.18%**	-0.19%
	(3.60)	(2.95)	(2.83)	(0.94)	(1.15)	(2.05)	(-1.31)
Mutual funds	0.62%***	0.43%***	0.46%***	0.13%	0.10%	0.23%***	-0.39%*
	(3.90)	(2.77)	(3.72)	(1.02)	(0.91)	(2.76)	(-1.89)
Other inst.	0.30%*	0.33%***	0.19%*	0.51%***	0.45%**	0.61%***	0.32%*
	(1.80)	(2.93)	(1.66)	(3.63)	(2.58)	(3.52)	(1.82)
Short sellers	-0.27%	0.18%	0.32%**	0.52%***	0.14%	-0.05%	0.22%
	(-1.18)	(0.98)	(2.17)	(4.62)	(0.87)	(-0.27)	(1.18)
Wealth mgmt.	0.59%***	0.46%***	0.58%***	0.32%***	0.22%*	0.15%	-0.44%***
	(4.36)	(4.34)	(3.74)	(2.80)	(1.87)	(0.96)	(-2.72)
Retail	0.39%**	0.45%***	0.27%**	0.03%	0.39%*	1.08%***	0.69%**
	(2.61)	(3.45)	(2.50)	(0.14)	(1.68)	(4.31)	(2.25)
Panel B: Lagged-holding sorts							
Banks	0.14%	0.87%**	0.97%***	0.03%	0.22%***	-0.03%	-0.17%
	(0.31)	(2.18)	(3.89)	(0.40)	(2.66)	(-0.37)	(-0.37)
Hedge funds	-0.03%	0.75%***	0.47%***	0.40%***	0.27%***	0.33%***	0.36%
	(-0.07)	(2.99)	(3.10)	(3.13)	(2.67)	(2.67)	(0.87)
Insurance co.	0.81%*	0.41%	0.84%***	0.06%	0.10%	0.06%	-0.74%*
	(1.89)	(1.01)	(2.76)	(0.70)	(1.53)	(0.66)	(-1.75)
Mutual funds	-0.28%	1.20%***	0.90%***	0.18%**	0.02%	0.04%	0.32%
	(-0.55)	(3.46)	(4.06)	(2.24)	(0.25)	(0.33)	(0.60)
Other inst.	0.00%	0.69%***	0.43%***	0.25%**	0.44%***	0.31%**	0.31%
	(0.00)	(3.10)	(2.88)	(2.60)	(4.08)	(2.05)	(0.70)
Short sellers	0.01%	-0.05%	-0.05%	0.41%***	0.69%***	1.53%***	1.52%***
	(0.03)	(-0.28)	(-0.39)	(4.46)	(3.26)	(4.89)	(4.70)
Wealth mgmt.	0.02%	0.68%*	0.69%***	0.19%*	0.31%***	0.34%***	0.32%
	(0.04)	(1.71)	(4.51)	(1.70)	(3.24)	(3.82)	(0.77)

Table B.3. Performance of individual machine-learning methods

This table reports the out-of-sample return and alpha of the long-short portfolios based on each of the twelve machine-learning methods. The first column reports the name of the machine-learning method, and the second, third, and fourth columns report the out-of-sample return, alpha with respect to the five Fama and French (2015) and momentum factors (FF5+MOM), and alpha with respect to the five Hou et al. (2021) and momentum factors (q5+MOM) of the long-short portfolio based on each of the 12 predictive methods: ordinary least squares (OLS), principal component regression (PCR), partial least squares (PLS), adaptive least absolute shrinkage and selection operator (ALasso), ridge regression (Ridge), elastic net (ENet), random forest (RF), gradient boosting regression trees (GBRT), and artificial neural networks with one to four hidden layers (ANN1, ANN2, ANN3, ANN4). At the end of each month, we rank all sample stocks based on one of the predictors, and then sort them into decile portfolios. We then compute the equal-weighted returns on the long-short portfolio and report the excess returns and alphas. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	Return	FF5 + MOM	q5 + MOM
OLS	1.20%***	1.79%***	1.66%***
	(3.52)	(10.17)	(6.23)
PCR	0.95%**	1.38%***	1.14%**
	(2.38)	(6.02)	(2.54)
PLS	1.16%***	1.77%***	1.62%***
	(3.54)	(10.75)	(7.01)
Ridge	1.21%***	1.60%***	1.67%***
	(4.69)	(12.36)	(6.87)
Enet	1.15%***	1.56%***	1.61%***
	(4.55)	(9.91)	(6.03)
Alasso	1.14%***	1.73%***	1.61%***
	(3.33)	(8.80)	(5.96)
RF	2.03%***	2.29%***	2.18%***
	(5.69)	(5.62)	(6.05)
GBRT	1.96%***	2.36%***	2.08%***
	(8.14)	(7.95)	(5.83)
ANN1	1.59%***	1.97%***	1.69%***
	(3.87)	(7.49)	(4.56)
ANN2	1.90%***	2.33%***	2.15%***
	(5.43)	(8.57)	(6.25)
ANN3	1.88%***	2.19%***	1.98%***
	(5.80)	(7.46)	(5.75)
ANN4	1.91%***	2.19%***	1.99%***
	(5.85)	(7.48)	(5.81)

Table B.4. Performance of composite return predictors obtained excluding retail trades

This table reports the alpha of the portfolios of stocks sorted by the linear (LCP) and nonlinear (NLCP) composite return predictors obtained after excluding the retail trades from the input variables used to train the machine-learning methods. The first column reports the label for each decile portfolio and the long-short portfolio, the second and third columns report the alpha of each portfolio based on the LCP with respect to the five Fama and French (2015) and momentum factors (FF5+MOM) and the five Hou et al. (2021) and momentum factors (q5+MOM), the fourth and fifth columns for the portfolios based on the NLCP, and the sixth and seventh columns for the difference between the returns of the portfolios based on NLCP and LCP (NLCP minus LCP). At the end of each month, we rank all sample stocks based on either the LCP or NLCP, and then sort them into decile portfolios. We then compute the equal-weighted returns on each decile as well as on the long-short portfolio and report the alphas. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	LCP		NLCP		NLCP minus LCP	
	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM	FF5 + MOM	q5 + MOM
L	-1.04%***	-0.55%**	-1.35%***	-0.84%**	-0.31%	-0.30%
	(-3.74)	(-2.07)	(-3.86)	(-2.28)	(-1.58)	(-1.38)
2	-0.44%**	-0.10%	-0.39%**	0.00%	0.05%	0.10%
	(-2.49)	(-0.48)	(-2.29)	(0.01)	(0.34)	(0.63)
3	-0.18%	0.17%	0.17%	0.51%**	0.36%**	0.34%*
	(-1.05)	(0.89)	(1.01)	(2.57)	(2.10)	(1.95)
4	0.41%***	0.70%***	-0.01%	0.34%*	-0.41%***	-0.36%**
	(2.69)	(3.91)	(-0.05)	(1.84)	(-2.75)	(-2.08)
5	0.02%	0.38%**	0.07%	0.29%**	0.05%	-0.09%
	(0.18)	(2.29)	(0.58)	(2.22)	(0.33)	(-0.52)
6	0.23%*	0.38%***	-0.12%	0.10%	-0.35%**	-0.28%*
	(1.98)	(2.73)	(-1.35)	(1.06)	(-2.59)	(-1.91)
7	0.35%***	0.55%***	0.31%***	0.49%***	-0.04%	-0.06%
	(3.05)	(4.89)	(3.41)	(5.33)	(-0.32)	(-0.44)
8	0.52%***	0.78%***	0.51%***	0.72%***	-0.01%	-0.06%
	(3.84)	(4.93)	(3.05)	(4.20)	(-0.06)	(-0.35)
9	0.41%***	0.67%***	0.72%***	1.05%***	0.32%**	0.38%**
	(2.93)	(4.36)	(4.07)	(5.92)	(2.04)	(2.36)
H	0.72%***	1.14%***	1.08%***	1.46%***	0.35%	0.32%
	(3.66)	(5.16)	(4.17)	(4.83)	(1.65)	(1.14)
H-L	1.76%***	1.69%***	2.43%***	2.30%***	0.67%**	0.61%*
	(9.03)	(6.07)	(10.10)	(6.29)	(2.30)	(1.75)

Table B.5. Value-weighted portfolio performance and firm size

This table reports the alpha of the value-weighted portfolios of stocks sorted by the LCP (or NLCP) across sub-samples of stocks categorized by firm size. At the end of each month, we rank all sample stocks based on firm size and sort them into tercile portfolios. We then rank stocks within each tercile based on either the LCP or NLCP, and sort them into decile portfolios. We then compute the value-weighted returns on each decile as well as on the long-short portfolio and report the alphas with respect to the five Fama and French (2015) and momentum factors (FF5+MOM). The first column reports the label for each decile portfolio and the long-short portfolio, the second and third columns report the alphas for small firms, the fourth and fifth columns for medium firms, and the last two columns for large firms. The t-statistics are Newey-West adjusted and shown in parentheses, with ***, **, and * indicating statistical significance at the 1%, 5%, and 10% level, respectively. The out-of-sample period for the evaluation of the composite return predictors spans from 2013 to 2020.

	Small		Medium		Large	
	LCP	NLCP	LCP	NLCP	LCP	NLCP
L	-2.04%***	-2.41%***	-0.41%*	-0.71%***	-0.27%	-0.13%
	(-4.39)	(-4.06)	(-1.76)	(-2.85)	(-1.64)	(-0.77)
2	-0.75%*	-1.11%**	-0.22%	-0.05%	0.13%	-0.23%
	(-1.95)	(-2.53)	(-1.24)	(-0.35)	(0.75)	(-1.09)
3	-0.11%	0.01%	-0.08%	0.29%*	0.31%*	0.08%
	(-0.31)	(0.02)	(-0.48)	(1.85)	(1.80)	(0.49)
4	0.34%	0.41%	-0.08%	0.09%	0.08%	0.07%
	(1.05)	(1.41)	(-0.61)	(0.71)	(0.61)	(0.62)
5	0.28%	0.04%	0.05%	-0.20%	-0.10%	0.10%
	(1.25)	(0.19)	(0.28)	(-1.49)	(-0.72)	(0.79)
6	0.82%***	0.57%***	0.12%	0.21%	-0.08%	0.19%
	(2.79)	(2.81)	(0.80)	(1.26)	(-0.85)	(1.65)
7	0.30%	0.37%	0.03%	-0.21%	-0.04%	-0.06%
	(1.30)	(1.25)	(0.16)	(-1.20)	(-0.30)	(-0.63)
8	0.59%**	0.59%***	0.27%*	0.05%	-0.14%	-0.10%
	(2.04)	(2.76)	(1.90)	(0.25)	(-0.80)	(-1.06)
9	0.55%*	1.22%***	0.02%	0.29%**	-0.19%	-0.07%
	(1.78)	(4.11)	(0.14)	(2.62)	(-1.63)	(-0.44)
H	0.96%***	0.89%***	0.28%	0.17%	0.40%***	0.23%
	(3.12)	(2.65)	(1.51)	(0.88)	(3.20)	(1.56)
H-L	3.00%***	3.30%***	0.69%**	0.87%***	0.67%***	0.36%
	(6.70)	(6.11)	(2.55)	(3.04)	(3.26)	(1.22)