# A Tale of Two Zoos:

# Machine Learning Insights on Retail Investors

Pulak Ghosh*, Huahao Lu†, Hong Zhang‡, Jian Zhang§

This Draft: May 6, 2024

## Abstract

Both stock characteristics (the "factor zoo") and behavioral biases (the "bias zoo") may affect the returns of retail investors. But which factors/biases are more important? We address this question by utilizing machine-learning tools to analyze 15 million retail investor accounts in India. We observe that Neural Networks outperform other algorithms in uniquely predicting both good and bad out-of-sample performance, with (under)diversification, portfolio turnover, and momentum being the leading factors to influence total returns. For new trades, turnover, the disposition effect, and diversification emerge as the most important to affect returns. Overall, behavioral biases exert a larger influence than firm characteristics.

**Keywords: Machine Learning, behavioral biases, Retail Investors**
**JEL Codes: D90, G40,**

* Department of Decision Sciences and Centre of Public Policy, Indian Institute of Management Bangalore, Email: pulak.ghosh@iimb.ac.in
† PBC School of Finance, Tsinghua University, Email: luhh.21@pbcsf.tsinghua.edu.cn
‡ Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, Singapore 178899; E-mail: hongzhang@smu.edu.sg.
§ The University of Hong Kong, Faculty of Business and Economics, Email: zhangj1@hku.hk

# I. Introduction

Over the last few decades, the academic community has demonstrated that trading strategies based on a multitude of firm characteristics can yield striking returns (i.e., the “factor zoo”; see, e.g., McLean and Pontiff 2016; Harvey, Liu, and Zhu 2016; Hou, Xue, and Zhang 2020; Kelly and Pedersen 2022 for recent evidence). One important implication from this literature is that investing in stocks with the proper firm characteristics, purposefully or not, may allow retail investors to achieve better returns. A distinct yet interrelated literature investigates the various psychological heuristics and biases to which retail investors are susceptible (i.e., the “bias zoo”; see Barber and Odean 2013; Hirshleifer 2015; and Barberis 2018 for literature reviews). Important questions arise when we combine insights from the two streams of studies. Between behavioral heuristics and firm characteristics, which contribute more to the investment returns of retail investors? Moreover, within the realm of behavioral heuristics and firm characteristics, which factors exert more influence? These questions carry crucial normative implications. Retail investors globally utilize stock markets to build wealth, save for retirement, and achieve various financial objectives. However, vulnerability to biases and exposure to (wrong) characteristics may impede these goals and impact price efficiency. Hence, scrutiny of these questions is vital, not only for the investors themselves but also for academic researchers, policymakers, and financial institutions dedicated to promoting financial well-being and stability.

Our paper aims to shed light on these questions by employing various machine learning tools to a unique and large proprietary account-level dataset containing the daily trading activities of *all* retail investors on the National Stock Exchange of India (NSE). As the most populous country in the world, India provides an ideal testing ground to understand retail investors, with the NSE being its largest stock exchange in India and the seventh largest stock market worldwide by Dec 2023.[5]. From this dataset, we identify 15.4 million valid retail accounts and 1.523 billion investor-month return observations for the testing period of 2012-2020. We then construct 23 holding-weighted stock characteristics and 13 proxies of behavioral heuristics and characteristics for each account, which we term *stock characteristics* and *behavioral biases* when there is no confusion.

[5] https://www.cnbc.com/2023/12/12/india-overtakes-hong-kong-to-become-worlds-seventh-largest-stock-market.html

We employ the *Feedforward Neural Network* (*FNN*) and an enhanced *Residual Neural Network* (*ResNN*) as our main machine-learning models, while also examining traditional OLS and machine-learning models (e.g., LASSO, Ridge, and Random Forest). The benefit of Neural Network algorithms is that they can reliably estimate a complex functional relationship among a large set of predictors. Of the two Neural Network models, FNN is more traditional, while ResNN reflects the more recent development in convolution deep learning. Its key feature, "residual connections", is widely adopted in recent architectures such as BERT and ChatGPT. Later sections will delve into further details, demonstrating ResNN's superior capability for our purposes.

Our primary objective is twofold. First, we predict retail investors' overall investment returns using machine learning models based on ranks of behavioral biases and stock characteristics.[6] Second, upon establishing a model's reliable out-of-sample predictive power for investor returns, we then use the model to nail down the most important stock characteristics and behavioral biases affecting those returns.

Our return prediction analysis is employed as follows. In line with the literature (e.g., Kaniel et al., 2023), we divide our return prediction period from 2012 to 2020 into three equal-length subperiods. We then train a model on two subsets of the data and use the trained model to predict returns on the remaining subset. This approach ensures that we can test the out-of-sample predictive power of the model over the entire prediction period.[7] During the out-of-sample predicting period, we categorize retail investors into five quintiles according to the predicted returns of a particular model; all quintiles are rebalanced monthly. The *High* and *Low* groups comprise the top and bottom 20% of predicted winners and losers among investors, respectively. We then calculate the value-weighted out-of-sample returns of the high and low groups, along with

[6] We use ranks to normalize the distribution of all imputs so their importance can be more easily inferred. This method is widely used for machine learning models (e.g., Kelly, Pruitt, and Su, 2019; Freyberger, Neuhierl, and Weber, 2020). To calculate an investor's total investment returns, we calculate the daily return generated by her existing portfolios at the beginning of a given date and then compound her daily returns into monthly returns. This approach allows us to further decompose the investor's monthly total returns into two sources, the part generated by the holding at the beginning of the month (i.e., holding-based returns) and the part generated by the newly initiated trading during the month (i.e., new trading-based returns). As we will see shortly, behavioral biases and firm characteristics play different roles in affecting the two sources of returns.

[7] Our sample starts from 2010. We use the first two years of information to calculate the initial values of stock characteristics and behavioral biases. Hence, our return predicting test starts in 2012. We adopt the Kaniel et al., (2023) approach because it has the advantage of testing the model on every sample in the dataset, enhancing the robustness of model comparisons by mitigating the influence of specific periods. Additionally, within the training data, we randomly set aside 30% of the samples for validation purposes.

their return difference. We further use the local three-factor (Fama and French 1992, 1993) or four-factor models (Carhart, 1997) to adjust these investor portfolio returns.

We observe that the two Neural Network models outperform other models in predicting total returns for retail investors. Both FNN and ResNN identify investors who can generate significantly positive out-of-sample returns. Despite being retail investors, the top 20% of predicted winners can generate a monthly return of 1.3% and 1.7%. The economic magnitude remains approximately the same when risk-adjusted (e.g., 1.4% and 1.5% adjusted by four factors). In contrast, all other models fail to predict winners. Given how difficult it is to predict four-factor adjusted superior mutual fund performance in the US (e.g., Carhart 1997), the superior performance of retail investors strongly suggests that Neural Networks capture crucial characteristics of retail investors.

On the loser side, many models can select *Low* groups that deliver significant or marginally significant returns. The two neural network models still perform the best, with FNN and ResNN-predicted *Low* groups delivering a significant negative monthly return of -2.0% and -2.8%, respectively, allowing their *High* groups to outperform the *Low* groups by as much as 3.3% and 4.5% per month. The *Low* group retail investors selected by Lasso and Ridge can deliver marginally negative returns (at the 10% level), though the *High* group significantly outperforms the *Low* group by approximately 2.6%. In contrast, the *Low* groups selected by OLS and Random Forest models deliver insignificant returns, though the high-minus-low spread is still significant for OLS with a smaller economic magnitude (1.9%). Overall, *FNN* and, particularly, *ResNN* outperform other models in identifying true winners and losers among retail investors.

As Neural Network models exhibit superior predictive capabilities for investor returns, we next utilize them to explore the relative importance of behavioral biases and firm characteristics. While our previous analysis utilized both sets of predictors, we now examine each set independently to shed light on their relative importance. To be consistent with the literature (e.g., Kaniel, Lin, Pelger, and Van Nieuwerburgh 2023), we use the traditional FNN model to investigate and demonstrate this standalone predicting power.

When firm characteristics are used alone, FNN-predicted *High* and *Low* groups fail to deliver significantly positive or negative out-of-sample returns. Nor can FNN predict a significant *High*-minus-*Low* return spread. In contrast, using behavioral biases alone can predict a significant *High*-

minus-*Low* return spread, primarily driven by the negative returns on the loser side. This comparison highlights the relative importance of behavioral biases in affecting returns.

Moreover, we can conduct variable gradient analysis to demonstrate the relative importance of each variable when behavioral biases and stock characteristics are both used. In this case, the FNN identifies diversification, portfolio turnover, and momentum as the top three leading factors to influence overall retail returns. The first two variables could be related to the behavioral biases of under-diversification and overconfidence. Under-diversification is among the most common features of retail investors (e.g., Barber and Odean 2000; Benartzi and Thaler 2001) as many investors may not fully understand the benefits of diversification (Lusardi and Mitchell 2011). Overconfidence often causes investors to trade too aggressively, allowing their high portfolio turnover to reduce their welfare (Odean 1998; Barber and Odean 2000). The third denotes perhaps the most famous anomalies in the literature. It is interesting to observe that behavioral biases occupy two out of the top three factors affecting investment returns.[8]

Next, we notice that investors' total returns can stem from two distinct sources: holding an existing portfolio for a specified period, such as a month (i.e., holding returns), and initiating new trades to buy and sell stocks during the month (i.e., trading returns). Behavioral theories propose that the motivations for new trading may differ from those for continuation. For instance, the disposition effect (e.g., Shefrin and Statman, 1985) suggests that unrealized capital gains motivate investors to trade (i.e., sell), whereas loss aversion incentivizes investors to retain losing assets, thus influencing holding returns. Another example is the salience theory (Bordalo, Gennaioli, and Shleifer 2012; 2013; 2020), which suggests that salient information, such as extreme stock prices, may also attract investors' attention to initiate new trades. Hence, the next question we explore is how behavioral biases and firm characteristics affect the two sources of returns

To address this issue, we adjust our objective to predict trading or holding returns using Neural Networks. We first validate the predicting power by observing the predicted *High* (*Low*) group to

[8] Part of momentum effect may be related to investors' biases. For instance, the disposition effect may induce investors to load on loser's momentum. Since our goal is to identify the direct impact of predictors, we attribute such return to momentum. Additionally, observed portfolio diversification and turnover may also be related to alternative sources of biases, such as local bias, local information, and financial literacy. Since our purpose is to compare behavioral biases and stock characteristics, we do not further nail down the economic sources of behavioral bias.

generate significantly positive (negative) total returns.[9] This predicting power also gives rise to a significantly positive *High*-minus-*Low* total return spread. We then employ the variable gradient analysis to assess the relative importance of each variable in predicting each source of returns.

Our main finding is that behavioral biases play an even more important role in predicting trading returns. This observation is intuitive, as characteristics-related returns likely contribute more to less rebalanced portfolios, whereas new trading is often initiated by behavioral reasons. Consistent with this notion, we observed that portfolio turnover, the disposition effect, and the degree of portfolio diversification emerge as the three most important factors in predicting new trading returns. Notably, the disposition effect also emerges as a leading predictor for trading, confirming the importance of disposition-related preferences in initiating trading (e.g., realization utility or the utility predicted by prospect theory).[10] Turnover and diversification remain leading predictors for both total and trading returns.

The relative importance of behavioral biases and firm characteristics can also be quantified by the joint explanatory power of all predictors falling into each category. We observe that the joint explanatory power of behavioral biases slightly exceeds that of firm characteristics in predicting total returns. However, the former dominates the latter in predicting trades, with behavioral bias predictors jointly occupying almost 95% of the total predicting power. These observations confirm the importance of behavioral biases particularly for explaining the trading of retail investors.

We finally conduct robustness checks. In the main analysis, we exclude 30% of small stocks because these stocks are difficult to trade in emerging markets (Liu et al. 2019) and may overstate the predicting power of machine learning models (Avramov, Cheng, and Metzker 2023 and Cong et al., 2021). We show that the predicting power of both FNN and ResNN is robust to different thresholds of removal (e.g., 40%), suggesting that our conclusions are unlikely to be contaminated by small stocks.

---

[9] In other words, we observe that the predicted winners (losers) in one source of return can deliver positive (negative) total returns. Unreported results confirm that, when the objective is to predict trading (holding) based returns, Neural Networks can deliver significaht out-of-sample High-minus-Low spread in their respective return sources.

[10] Since Shefrin and Statman (1985), the development of this literature has been extensive, though the causes and consequences of the disposition effect are still under debate (see, among others, Grinblatt and Han, 2005; Barberis and Xiong, 2009, 2012; Calvet, Campbell, and Sodini, 2009; Ivkovic and Weisbenner, 2009; Kaustia, 2010; Ben-David and Hirshleifer, 2012; Henderson, 2012; Li and Yang, 2013; Frydman et al., 2014; An, 2016; Chang, Solomon, and Westerfield, 2016; Fischbacher, Hoffmann, and Schudy, 2017; Frydman and Wang, 2020). DellaVigna (2009; 2018), Hirshleifer (2015), and Barberis (2018) provide recent surveys.

Our results are related to several strands of literature. A growing literature demonstrates that machine-learning models can help predict asset prices in different sectors of the markets, ranging from the equity premium to option pricing in the US and global markets.[11] Karolyi and Van Nieuwerburgh (2020) and Kelly and Xiu (2023) provide recent reviews. Our analysis is closely related to recent studies applying machine-learning models to predict the performance of institutional investors, such as mutual funds (e.g., Li and Rossi, 2020; DeMiguel, Gil-Bazo, Nogales, and Santos, 2023; Kaniel, Lin, Pelger, and Van Nieuwerburgh 2023) and hedge funds (Wu, Chen, Yang, and Tindall, 2021). We contribute by using a battery of machine-learning tools to scrutinize the performance of a large sample of retail investors. This extension is important, as the economic rationale guiding retail investors' investments can differ from that of institutional investors. It allows us to establish a more comprehensive understanding of investor decisions.

In doing so, we contribute to the literature on behavior finance. One important goal of this literature is to use psychological insights to explain many anomalies in individuals' financial decision-making. This effort leads to profound insight into how individual investors make decisions (see, e.g., Barber and Odean 2013, Hirshleifer 2015, Barberis 2018 for literature reviews). The multitude of proposed behavioral biases, however, also gives rise to a "lack of discipline" concern (Fama, 1998). Indeed, if the excessive return predictability of characteristics-based anomalies already imposes a "multidimensional challenge" (Cochrane 2011), the concern becomes even more prominent when examining retail investors, as both psychology-based and characteristics-based anomalies may influence their returns. To consolidate the multitude of behavioral biases, a few recent papers use survey-based methods to nail down their relative importance (e.g., Choi and Robertson, 2020; Liu, Peng, Xiong, and Xiong, 2022). Our novelty is to use machine learning tools to reduce the dimension of both types of anomalies, which can shed light on a more parsimonious conceptual framework of asset pricing and investor behavior.

[11] See, among others, Freyberger, Neuhierl, and Weber (2020), Gu, Kelly, and Xiu (2020), Bryzgalova, Pelger, and Zhu (2020), and Chen, Pelger, and Zhu (2023) for stock returns and characteristics, Jensen, et al. (2022) for trading-cost-adjusted portfolio optimization, Leippold, Wang, and Zhou (2022) for the Chinese equity market, Li et al (2023) on the spillover effect of the global supply chain, Bianchi, Büchner, and Tamoni (2021) for bond risk premium, Easley, López de Prado, O'Hara, and Zhang (2021) for market microstructure, Filippou et al. (2023) for currencies, Bali, Beckmeyer, Mörke, and Weigert (2023) for option pricing, and Van Binsbergen, Han, and Lopez-Lira, A. (2023) for the conditional biases in earnings expectations. Avramov, Cheng, and Metzker (2023) report that the predicted returns could drop substantially in magnitude when small firms are excluded.

In a closely related paper, Balasubramaniam, Campbell, Ramadorai, and Ranish (2023) also use a large sample of Indian retail accounts to shed light on investor attributes that can give rise to investor clientele effects for stock characteristics.[12] Our paper differs in that we first validate machine learning models based on their predicting power on investor returns and then use the most reliable tools (i.e., neural networks) to examine the importance of investor bias and stock characteristics. In other words, investor returns play a pivotal role in our analysis, which differs from their focus on investor holdings. A unique feature of our approach is to identify investor bias and stock characteristics that can directly impact the returns and thus welfare of retail investors.

Lastly, we also make methodological contributions by introducing Residual Neural Networks (*ResNN*) into financial analysis. Despite the popularity of Neural Networks in finance, a widely acknowledged challenge in deep learning is that deeper neural networks are more difficult to train (i.e., the vanishing gradient problem). *ResNN* addresses this difficulty by reformulating the output of a particular layer as a learning residual function plus the layer's input (He et al., 2015).[13] The key feature of *ResNN*—"residual connections", or the addition of the original input to the output of a deeper layer within a neural network—is also widely used in Transformer models such as BERT and ChatGPT. This feature allows *ResNN* to be trained deeper and more easily optimized. Our results confirm that *ResNN* serves as a suitable tool for comprehensive financial tasks, such as analyzing retail investors.

The remaining article is organized as follows. Section II describes the data and machine learning models. Section III provides baseline tests for predicting retail investors' returns. Section IV examines the importance of behavioral heuristics and firm characteristics. Section V provides additional tests and robustness checks, followed by a short conclusion with policy implications.

## II. Data, Main Variables, and Machine Learning Models

This section describes the data and explains how we construct our main variables. We then briefly describe the machine learning models used in our later analysis.

[12] In the literature, researchers have also used Scandinavian account-level household data to examine the attributes that affect investors' investment decisions. A common finding is that the decisions of households are strongly influenced by behavioral biases (see, among others, Massa and Simonov 2006; Døskeland and Hvide 2011; Grinblatt et al. 2016). We instead systematically explore a list of attributes to determine their relative importance.

[13] Residual Neural Networks were originally developed to improve image recognition and won the *ImageNet* 2015 competition. As of now, the seminar work of He et al., 2015 has garnered more than 189,026 Google citations.

## A. Data

We collected data from multiple sources. To characterize the impact on investors' trading behavior, we obtain a comprehensive database of all trading records on the NSE of India for the period 2010-2020. The NSE is the leading exchange in India and the world's 9th-largest stock exchange as of May 2021.[14] For each transaction, we can observe the anonymized permanent account number (PAN) of the individual[15], the transaction date, the ticker of the security, the number of shares purchased or sold, and the execution price. We require all transactions to be associated with stocks included in the Prowess Database (similar to CRSP in the US) maintained by the Centre for Monitoring Indian Economy (CMIE). Additionally, we retain only securities that are common shares of domestic stocks and exclude trading activities related to ETFs and foreign stocks.

The initial sample consists of the entire sample of 19.36 million retail accounts at NSE. For each retail investor, we further obtain sociodemographic data including gender, age, and, most importantly, geographic identifier (i.e., India PIN code), which allows us to identify the district of residence for each investor. We exclude accounts that have a negative balance, as such accounts could incur missing information or short selling. Our final sample includes 15.418 million valid individual accounts and approximately 1.52 billion investor-month portfolio-return observations.

We obtain stock returns and characteristics from the CMIE Prowess database maintained by CMIE, Center for Monitoring the India Economy. Previous studies on Indian firms have utilized this dataset, including works by Bertrand, Mehta, and Mullainathan (2002), Gopalan, Nanda, and Seru (2007), Lilienfeld-Toal, Mookherjee, and Visaria (2012) and Gopalan, Mukherjee, and Singh (2016). The detailed firm characteristics are summarized in Table 1. In addition, we employ the Fama-French three-factor model (Fama and French 1992, 1993) and Carhart's four-factor model (Carhart, 1997) to adjust investor returns. The data for these local factors are downloaded from Global Factor Data (Jensen, Kelly, and Pedersen, 2023).[16]

## B. Main Variables

[14] https://www.world-exchanges.org/our-work/statistics

[15] The PAN is a unique identifier issued to all taxpayers by the Income Tax Department of India. The trading data are at the individual level so that it is not a concern if a given individual investor may hold multiple accounts.

[16] We thank the authors for maintaining a comprehensive global factor dataset and making is easily accessible on: https://jkpfactors.com/

Our objective is to use machine-learning tools to predict investors' total returns. To construct the time series for an investor's total investment returns, we calculate the daily return generated by her existing portfolios at the beginning of a given date and then compound her daily returns into monthly returns. This approach is in line with the literature (e.g., Odean, 1998; Barber and Odean, 2000), which also allows us to further decompose the investor's monthly total returns into two sources: the part generated by the holding at the beginning of the month (i.e., holding-based returns) and the part generated by the newly initiated trading during the month (i.e., new trading-based returns). Empirically, the new trading-based return of a month is calculated as the difference between the monthly total return and the holding-based return. As we will see shortly, behavioral biases and firm characteristics play different roles in affecting the two sources of returns.

In constructing portfolio returns, we follow the literature on international stock returns—e.g., Liu et al.'s (2019) analysis of the Chinese stock market—and exclude 30% of small stocks. These small stocks are not only difficult to trade in emerging markets (Liu et al. 2019) but also may overstate the predicting power of machine learning models (Avramov, Cheng, and Metzker 2023; Cong et al., 2021). Due to the presence of some extreme values in the distribution of investors' monthly returns, we applied a winsorizing procedure at the 1st and 99th percentiles to mitigate the impact of outliers. Later sections will show that our results are robust to these data screening processes.

We resort to the recent behavioral and asset pricing literature to construct the list of predators. This data enables us to construct 13 investor characteristics, most of which are proxies for behavioral biases. Below we describe how we construct these variables.

**The Disposition Effect:** Many studies have demonstrated the behavioral bias of investors to sell stocks that have gained profits while choosing to continue holding stocks that have incurred losses (Shefrin and Statman, 1985; Odean, 1998).

Following Sui and Wang (2023), we estimate the disposition effect through the following model:

$$Sell_{i,j,t} = \alpha + \beta_i Gain_{i,j,t-1} + \epsilon_{i,j,t}$$

where i, j, and t represent investor i, stock j, and time t, respectively. $Sell_{i,j,t}$is a dummy variable that equals 1 if investor i sells stock j at time t and 0 otherwise. $Gain_{i,j,t-1}$is also a dummy variable that equals 1 if investor i gains a profit on stock j at time t-1 and 0 otherwise. To avoid look-ahead bias, we estimate this model using a rolling window approach, where $\beta_i$ represents the disposition effect of the investor.

**Diversification:** We measure the degree of diversification for each investor based on the number of stocks held in their portfolio. Specifically, we calculate the daily count of stocks in the investor's investment portfolio and subsequently take the monthly average.

**Turnover:** We employ investor turnover as a proxy for their trading activity. Prior research has consistently shown that increased trading frequency is often associated with inferior performance (Odean 1998; Barber and Odean 2000). We calculate the daily turnover as the ratio of the trading amount to the total value of the investor's portfolio, followed by monthly averaging.

**Local Bias:** Many studies indicate that investors often exhibit a preference for companies located in close geographic proximity (Ivkovic and Weisbenner, 2005; Massa and Simonov, 2006). We utilized geographical location data for company headquarters and matched it with investors' registered addresses based on postal codes. Employing the Google Maps API, we obtained latitude and longitude coordinates for both the headquarters of each company and the registered addresses of investors. Subsequently, using the Haversine formula, designed for computing surface distances between any two points on a sphere, we calculated the km distances from each investor to every company. We then performed a weighted summation of distances for companies held in each investor's portfolio, considering the weights associated with each holding.

**Extrapolation:** Extrapolation refers to the tendency of investors to preferentially purchase stocks that have exhibited superior performance in the recent past. Consequently, we initially computed the excess returns of each stock over the market return in the preceding three months. Subsequently, we aggregated these excess returns, considering the investor's portfolio weights for the respective stocks.

**Lottery Preference:** We employed three variables to represent the lottery-like characteristics of stocks: the relative size of prices (using open, close, high, and low prices), idiosyncratic volatility, and idiosyncratic skewness, following the definition outlined in Kumar (2009). The estimation of idiosyncratic volatility and idiosyncratic skewness is based on the CAPM Model. Similar to our summation of extrapolation, we aggregated these proxy variables for the three lottery preferences at the investor level using a value-weighted approach.

**Past Performance:** We employed the investor's portfolio returns over the preceding three months as a proxy variable for their investment capability.

**Portfolio Value:** To capture the potential wealth effect, we also calculate the total market value of the stocks held by investors in the previous month as an investor characteristic.

To ensure equal power of these proxies, each proxy is ranked in the cross-section (between 0 and 1). We then use the ranks as preditors presentating investors' behavior biases.

In addition to these investor behavioral proxies, we computed the 23 most important stock characteristics based on their holdings. Since these stock characteristics are common in the literature, we do not explain them in details. All stock characteristics underwent rank normalization across the cross-section. Subsequently, we weighted these ranked characteristics based on the investor's holdings of each stock.

Table 1 tabulates these variables as well as their detailed definitions. The Online Appendix (Table IN1) presents the summary statistics of our main variables. We can see all portfolio-level variables have a reasonable distribution. Based on these summary statistics, it is reasonable to examine further how behavioral biases and firm characteristics affect retail investors' investment returns. We will undertake this task in the next section.

## C. Machine-Learning Models

To examine further how behavioral biases and firm characteristics affect retail investors' investment returns, we employ a list of machine learning models, including Lasso, Ridge, Random Forests, and Neural Networks. Below we describe their main algorithms.

### C.1 Lasso and Ridge

When the number of predictors in a model is substantial, simple linear models may struggle to effectively fit the data, potentially leading to overfitting issues. Lasso and Ridge are both models grounded in the linear assumption, yet, in comparison to simple linear models, the distinction lies in the incorporation of regularization into the model's objective function. The objective of estimating model parameters is no longer solely minimizing the error between fitted values and observed values; rather, it involves introducing penalties for the magnitude of linear model parameters. Lasso penalizes the first moment of model parameters, denoted as "l1" parameter penalization, whereas Ridge penalizes the second moment, known as "l2" parameter penalization. Specifically, a model with regularization can be expressed in the following form:

$$\min_{\beta \in R^p} \{ \frac{1}{N} ||y - X\beta||_2^2 + \lambda ||\beta||_1 + \gamma ||\beta||_2^2 \}$$

Where $\beta$ is the model parameters, $\lambda$ and $\gamma$ are regularization coefficients, and in the context of the Lasso model, $\gamma$=0, while for the Ridge model $\lambda$=0.

### C.2 Random Forests

Random Forest is an ensemble learning method that operates by constructing a multitude of decision trees during the training phase. Decision trees are commonplace in machine learning, offering a non-linear modeling approach in contrast to linear models. Notably, decision trees are non-parametric models. A tree is constructed by iteratively splitting the dataset into subsets, forming successive child nodes. The splits are based on predictor variables that most effectively discriminate among potential outcomes.

Random Forests employ an ensemble strategy by averaging multiple deep decision trees, each trained on different segments of the same training set. This approach aims to mitigate variance, offering a robust modeling technique.

### C.3 Neural Network

Neural networks are currently highly popular models in various application domains, having achieved tremendous success in fields such as natural language processing and computer vision. According to the universal approximation theorem (Kurt et al. 1989), neural networks can approximate any function between input x and output y. In this context, we employed a multi-layer perceptron (MLP) network, also known as a feed-forward network (FNN), which is a standard and widely applicable neural network model in the financial market.

A multi-layer perceptron network consists of an input layer, an output layer, and one or more hidden layers. In each layer of the multi-layer perceptron, the input undergoes a linear transformation followed by an element-wise non-linear transformation (activation function). For the l-th layer of the MLP, its computational process can be expressed as follows:

$$X^l = g\left(W^{(l)T} X^{(l-1)} + b^{(l)}\right),$$

where $X^{(l-1)} \in R^{D^{l-1}}$ is the input to the $l$-th layer of the network, $W^{(l)} \in R^{D^{l-1} \times D^{l}}$ and $b^{(l)} \in R^{D^{l}}$ are the learnable parameters for the $l$-th layer, and $g(*)$ is the non-linear activation function. It is noteworthy that the output layer does not utilize a non-linear activation function. Instead, it directly aggregates the output from the previous layer through a linear mapping to form predictions for future returns, i.e.,

$$X^{output} = W^{(output)T} X^{(-1)} + b^{(output)}$$

For the choice of the activation function, we employ the most common rectified linear unit function (ReLU) in this context.

$$g(z) = ReLU(z) = \max(z, 0)$$

### C.4 Residual Learning

A Residual Neural Network (ResNN) is a deep learning model in which the weight layers learn residual functions concerning the layer inputs. It is characterized by skip connections, termed "residual connections," which perform identity mappings and are combined with the layer outputs through addition. This architecture facilitates the training of deep learning models with tens or hundreds of layers, leading to improved accuracy as the depth of the network increases. Notably, the concept of identity skip connections, or residual connections, extends beyond Residual Networks and finds application in various other models such as Transformer models (e.g., BERT, and GPT models like ChatGPT).

Following the seminal work of He et al., (2015), the computation for each layer can be expressed in the following form under the paradigm of residual connections:

$$X^{l} = g\left(W^{(l)T} X^{(l-1)} + b^{(l)}\right) + X^{(l-1)},$$

where $X^{(l-1)}$ denotes the original input, which is added back (e.g., through concatination) to the output of the layer, $X^{l}$.

This design allows *ResNN* to be trained deeper and more easily optimized due to three beneficial features when compared to traditional neural networks. First, instead of finding the optimal function or true information for optimization, which often lead to overparameterization when the optimization process is complex, each layer of neural in the residual learning framework only needs to figure out the additional information—compared to the inputs—that helps to improve maximization. In other words, the goal now becomes to augment the initial data by providing additional information. This design helps to simplify the task for each layer of neurals and also give it better information to achieve this simplified task.

Second, residual learning often combine several layers of neural into a block to facilitate residual connections. In this case, $X^{(l-1)}$ and $X^{l}$ in the above equation become the original input and the output of the block. The benefit of this design is that it allows information flows both within and aside the block, which shortens the gradient path of the blocks to speed up the training. This feature can significantly mitigate the issue of vanishing gradients in deep learning. Lastly, residual learning often adopts the modularity principle, which means to build the neural network based on blocks with similar structure. Modularity allows for more blocks and deeper learning (see, e.g., Sun and Guyon 2023 for a recent survey).[17]

Among the three features, the first cound be especially applicable to the financial market. This is because many information in the financial market is expressed vis-à-vis a benchmark. For instance, mutual funds are often benchmarked against an index. In this case, the excess return of a fund provides the most important information about its operation. For another example, investment returns are typically adjusted by risk factors, allowing risk-adjusted returns to reveal important properties of the investment strategy. Roughly speaking, this feature of *ResNN* also allows machine learning to be benchmarked against some economically important inputs. As we will see shortly, this feature can help further improve the performance of neural networks.

### D. Data sampling and Optimization

We employed the cross-validation method to train and assess the performance of our model. Following the approach outlined by Kaniel et al. (2023), we uniformly divided the entire dataset into three parts. In each iteration, we trained the model on two of the folds and tested its performance on the remaining fold. This approach offers the advantage of testing the model on every sample in the dataset, enhancing the robustness of model comparisons by mitigating the influence of specific periods. Additionally, within the training data, we randomly set aside 30% of the samples for validation purposes.

[17] https://arxiv.org/abs/2310.01154

After partitioning the data, we employed a gradient-based approach to train the neural network model. There are various training strategies for neural networks, and a common solution is to utilize the Adam optimizer. To enhance the optimization speed and performance of the model, the Adam optimizer randomly selects a subset of samples (batch) from the training data for gradient updates in each iteration.

A key parameter of the Adam optimizer is the learning rate, which dictates the step size for updates along the gradient direction. A well-chosen learning rate involves a trade-off between convergence speed and avoiding overshooting. Thus, it is essential to dynamically adjust the learning rate based on the training process's state. Therefore, we implemented a learning rate scheduler during training. A learning rate scheduler is a predefined framework that modifies the learning rate between epochs or iterations as the training advances. In this context, we employed a learning rate decay strategy, gradually reducing the learning rate as the training progresses.

Neural networks often exhibit strong expressive power and the ability to fit any arbitrary function, but they are also susceptible to the issue of overfitting. Overfitting occurs when a neural network performs well on the training data but poorly on unseen test data. Previous research generally attributes overfitting to the model memorizing the noise and details of the training data excessively while neglecting the overall distribution of the data, resulting in a decrease in the model's generalization ability.

To mitigate overfitting, we employed EarlyStopping and Dropout. EarlyStopping is a regularization technique in model training. If the model's performance on the validation dataset does not improve consistently, training is halted to prevent the model from excessively fitting the training data. Dropout involves ignoring the output of certain hidden layer nodes during training, setting these nodes' output values to zero. This approach reduces interactions between hidden layer nodes, thereby minimizing overfitting in neural networks (Hinton et al., 2012).

For the specific parameters of the model, we employed a three-layer fully connected neural network with 32, 16, and 8 neurons in each layer, respectively. The learning rate is set at 0.001, and the maximum training epochs are set at 150. Additionally, we implement an Early Stopping mechanism with a patience setting of 3, meaning that the training would be terminated if there was no improvement in the model's performance for more than three epochs.

# III. Predicting Total Returns for Retail Investors

We now use all the models to predict retail investors' total investment returns.

## A. The Portfolio Analysis Approach

Our baseline tests involve a machine-learning-based portfolio analysis. We first use all the models to predict retail investors' total investment returns. We then sort retail investors into five quintiles according to predicted returns, with the *High* and *Low* groups consisting of 20% of predicted winners and losers among investors, respectively. Finally, we report the value-weighted out-of-sample returns of the high and low groups as well as their return difference. we also use the locally estimated three-factor and four-factor models to adjust these returns.

## B. The Performance of Model Selected Investors

Table 2 tabulates the predicted returns of investor quintiles. Columns 1-3 present average monthly returns and alpha adjusted through local FF-3 and Carhart-4 models. Columns 4-6 depict results for the high group, while columns 7-9 detail outcomes for the high minus low return.

We observe that the two Neural Network models outperform other models in predicting retail investors' total returns. In particular, both FNN and ResNN identify investors who can generate significantly positive out-of-sample returns. Indeed, column (4) reports that the top 20% of retail investors of the two models can generate a monthly return of 1.5% and 1.2%, which remains highly significant with a similar economic magnitude when risk-adjusted (e.g., 1.7% and 1.4% adjusted by four factors, as reported in column 6).

In contrast, all other models fail to predict winners. Given how difficult it is to predict four-factor adjusted superior mutual fund performance in the US (e.g., Carhart 1997), the superior retail performance strongly suggests that Neural Network models capture important properties of retail investors.

On the loser side, ResNN performs the best. Column (1) reports that the ResNN-predicted *Low* group delivers a significantly negative monthly return of $-3.1\%$, allowing the *High* group to outperform the *Low* group by 4.4% per month in column (7). FNN-predicted *Low* group delivers a slightly lower and marginally significant return of $-2.5\%$. Although the low group's risk-adjusted performance becomes insignificant, its *High* group outperforms the *Low* group by 4.0% per month and remains highly significant after the risk adjustments.

Among other models, Lasso and Ridge can select retail investors that deliver marginal negative returns, whereas OLS and Random Forest do not exhibit significant predicting power on the negative return side. As a result, the *High* group of Lasso and Ridge can significantly outperform the *Low* group by 1.6% and 2.5%, as reported in column (7). Although the *Low* groups selected by OLS fail to deliver significant returns, its high-minus-low spread remains significant at 2.5%.

Collectively, we observe that the two Neural Network models outperform other models in predicting total returns for retail investors. In particular, both FNN and ResNN can identify retail investors who can consistently deliver positive returns. This predictive power is striking given how difficult it is for professional investors—such as mutual funds—to deliver out-of-sample performance. Of course, the difficulty in predicting risky adjusted superior mutual fund performance is typically based on traditional OLS methods (e.g., Carhart 1997), whereas machine-learning tools are typically more powerful to predict the performance of mutual funds (e.g., Li and Rossi, 2020; DeMiguel, Gil-Bazo, Nogales, and Santos, 2023; Kaniel, Lin, Pelger, and Van Nieuwerburgh 2023) and hedge funds (Wu, Chen, Yang, and Tindall, 2021).

In our setup, Neural Networks outperform even other machine-learning models. Such evidence strongly suggests that Neural Networks capture crucial characteristics of retail investors that contribute to their returns. As a result, Neural Networks provide a reliable tool to further analyze how behavioral biases and firm characteristics affect retail investors' investment returns.

## IV. The Importance of Predictors

We now employ Neural Network models to investigate the relative importance of behavioral biases and firm characteristics in affecting retail investors' investment returns.

### A. The Stand-alone Power of Behavioral Biases and Firm Characteristics

Our previous analysis uses both behavioral biases and firm characteristics as predictors of the Neural Network models. However, we can use each set of predictors alone, which can shed light on the relative importance of these predictors in predicting retail investors' investment returns. To be consistent with the literature (e.g., Kaniel, Lin, Pelger, and Van Nieuwerburgh 2023), we mainly use the traditional FNN model to investigate and demonstrate this standalone predicting power. But we will also illustrate the difference when the more advanced RedNN algorithm is used.

The results are tabulated in Table 3. The first row reports the results when the FNN is trained only based on the information of firm characteristics. Using this set of predictors, the Neural Network model fails to select the *High* and *Low* groups that can deliver significantly positive or negative out-of-sample returns. Nor can firm characteristics alone predict a significant *High*-minus-*Low* return spread.

The second row reports the results when behavioral biases are used alone by FNN. Different from the first line, using behavioral biases alone can predict a significant *High*-minus-*Low* return spread of 3.3% per month. Moreover, its power mostly arises from the negative return (loser) side, with the Low group delivering a -2.5% return. Both observations hint at the relative importance of behavioral biases in affecting returns.

The third row reports outcomes from simultaneously incorporating behavioral biases and firm characteristics by the FNN model. Although this result has also been reported in the previous table, the side-by-side comparison between this and the behavioral-only result can help reveal more properties of the FNN estimation. We first observe that the simultaneous use of both behavioral biases and firm characteristics enables FNN to predict a significant High-minus-Low return spread. This return spread (4.0%) is larger than the case when the FNN algorithm is trained only by behavioral biases.

However, the High-minus-Low return spread of the third row is primarily driven by the positive returns generated by the High group, not by the Low group. In contrast to the second-row case when FNN can select Low-group investors to deliver significantly negative returns when only behavioral biases are used, we find that the inclusion of more predictors (i.e., firm characteristics) diminishes the model's ability to identify the Low-group investors. Conceptually speaking, including more preditors should not impede an optimization algorithm, because it can always replicate the results based on a more restrictive set of predictors. Our counterintuitive empirical observation, however, suggests that the increased complexity of parameter space may subject the neural network optimization to common issues such as overparameterization and vanishing gradients, leading to a partial loss of predictive capacity (He et al., 2015). In particular, the potential efficiency loss concentrates on the impact of behavioral bias on the Low-group investors.

The above issue motivates us to adopt the residual learning framework proposed by He et al. (2015) to facilitate better training of neural networks. Specifically, when passing through a hidden layer (we have three in total), we concatenate the initial set of economically important predictors (i.e., the behavioral biases) with the output from the layer as the combined inputs for the subsequent layer. As mentioned earlier and demonstrated by He et al. (2015), such an approach can significantly enhance the optimization efficiency of neural networks by mitigating common issues, such as overparameterization and vanishing gradients, across the hidden layers.[18]

The last row affirms that our designed Residual Neural Network can effectively predict future returns for both loser and winner groups. For the loser group, it yields a substantial -3.1% excess return and a -2.6% alpha after adjusting with the Carhart four-factor model. Conversely, for the winner group, it generates a notable 1.2% excess return and a 1.4% four-factor adjusted alpha. In other words, ResNN can identify investors with both good and bad performance. Its out-of-sample High-minus-Low return spread is also approximately 10% higher in relative terms than that of FNN.

## B. Variable Gradient Analysis

We next follow the methodologies proposed by Sadhwani et al. (2020) and Horel and Giesecke (2020) and conduct variable gradient analysis to demonstrate the relative importance of each variable when behavioral biases and stock characteristics are both used. More explicitly,

$$Importance(x) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_t}\sum_{i=1}^{N_t}\left(\frac{\partial R_{i,t+1}^{pred}}{\partial x_{i,t}}\right)^2,$$

where $T$ represents the number of periods in the data, and $N_t$ denotes the total number of investors in the t-th period. The partial derivative measures the gradient of the model's predicted output with respect to each variable.

It is noteworthy to discuss two features of the estimation. First, the gradient can be positive or negative, so its square term is used to gauge its importance. For linear regression models, the partial derivative is simply the regression coefficient. Intuitively, a larger partial derivative implies a

[18] Our ResNN septically preserve the important predictors on Low-group investors. On the High-group side, we also note that, when used alone, neither firm characteristics nor behavioral biases would allow the FNN to successfully select the High group of investors to deliver superior out-of-sample returns. Hence, the interactions between investor behavioral biases and holding-based firm characteristics provide the source for FNN to predict good returns for investors. Such potential interaction effects remain effective in the residual network framework.

greater influence of a variable on the model's output, indicating greater importance in predicting future returns.

Second, the traditional neural network algorithm is more appropriate for this gradient analysis. Indeed, although the residual neural network can help improve model performance, the concatenation of the output from a layer with the initial set of economically important predictors (i.e., the behavioral biases) may introduce a selection bias on the importance of these variables. As a result, we compute the *Importance(x)* for each predictor when we use FNN to predict the total return of investors based on both firm characteristics and behavioral biases.

The results are plotted in Panel A of Figure 2. The FNN identifies diversification, portfolio turnover, and momentum as the top three leading factors to influence overall retail returns. The first two variables are related to the behavioral biases of under-diversification and overconfidence. Under-diversification is among the most common features of retail investors (e.g., Barber and Odean 2000; Benartzi and Thaler 2001) as many investors may not fully understand the benefits of diversification (Lusardi and Mitchell 2011). Overconfidence often causes investors to trade too aggressively, allowing their high portfolio turnover to reduce their welfare (Odean 1998; Barber and Odean 2000). The third denotes perhaps the most famous anomalies in the literature. It is interesting to observe that behavioral biases occupy two out of the top three factors affecting investment returns.[19]

### C. Predicting Trading or Holding-based Returns As an Alternative Objective

Next, we recognize that investors' total returns may originate from two different sources: from holding an existing portfolio for a given period of, for instance, a month (i.e., holding returns) and from newly initiated trading during the month (i.e., trading returns). Behavioral theories suggest that the motivations to initiate trading may differ from those of continuation. For instance, the well-documented disposition effect (e.g., Shefrin and Statman, 1985) suggests that unrealized capital gains motivate investors to trade (i.e., sell), whereas unrealized capital losses incentivize investors to hold onto losing assets and thus affect holding returns. For another example, salient information, such as extreme stock prices, may also attract investors' attention to initiate new

[19] Note that observed portfolio diversification and turnover may also be related to alternative sources of biases, such as local bias, local information, and financial literacy. Since our purpose is to compare behavioral biases and stock characteristics, we do not further nail down the economic sources of behavioral bias.

trades according to the salience theory (Bordalo, Gennaioli, and Shleifer 2012; 2013; 2020), even though traditional financial theory suggests investors should pay more attention to returns rather than the level of prices. Hence, the next question we ask is how behavioral biases and firm characteristics affect the two sources of returns.

To address this issue, we shift the predicting goal, using behavioral biases and firm characteristics to predict holding returns or trading returns in the Neural Network. We then use the variable gradient analysis to assess the relative importance of each variable in predicting each source of returns. The results are plotted in Panels B and C of Figure 2.

Our main finding is that behavioral biases still play a relatively more important role in predicting trading returns than holding returns. We observed that portfolio turnover, the disposition effect, and the degree of portfolio diversification emerge as the three most important factors in predicting new trading returns, followed by the opening and closing price of stocks. In other words, the disposition effect emerges as one of the leading predictors for trading, in addition to turnover and diversification.

Interestingly, (under)diversification, portfolio turnover, and momentum are still the three leading factors to influence holding-based returns. Different from the case of total returns, however, momentum becomes more prominent and surpasses turnover in terms of importance. The observation is intuitive: given the return predicting power of momentum, its influence should be stronger for holdings.

## V. Additional Analyses

This section provides additional analysis to shed light on the economic interpretation and robustness of our existing results.

### A. The Joint Power of Behavioral Biases and Firm Characteristics

The relative importance of behavioral biases vis-à-vis firm characteristics can also be expressed as the joint explanatory power of all predictors falling into each category. Figure 3 illustrates the Relative Importance of Behavioral Bias vs. Firm Characteristics.

Based on our earlier delineation, where predictors are categorized into Investor Behavioral Bias and Firm Characteristics, we define the variable importance measure of a group by computing the average of the importance measures within that group. Without loss of generality, we normalize the variable importance to sum up to 1.

Generally, the relative importance of Stock features gradually decreases as the prediction target shifts from Holding Return to Trading Return. Specifically, for Holding Return, the relative importance of Stock features and Investor features is 47.8% and 52.2%, respectively. In predicting total returns, we observe that the joint explanatory power of behavioral biases slightly exceeds that of firm characteristics, where the relative importance is 36.5% for Stock features and 63.5% for investor features. However, when it comes to predicting trading returns, behavioral bias predictors take the lead, accounting for nearly 95% of the total predictive power. These observations confirm the importance of behavioral biases particularly for the trading and its related return of retail investors.

### B. Using Alternative Objective-trained Algorithms to Predict Total Return

We next investigate whether neural network-selected investor groups based on alternative training goals—i.e., to predict trading and holding-based returns—can also help predict total returns. Take the training goal of predicting investors' trading-based returns as an example. In this case, we sort investors into quintiles by their predicted trading-based returns. It is perhaps not surprising that such a trading goal would allow FNN and ResNN to predict trading-based returns—unreported tests indeed confirm this predicting power. However, what total returns do these investors receive? Do investors poor at generating trading-based returns also receive poor total returns, which hurt their welfare? Such a predicting power, if observed, can further illuminate the economic source of retail investors' returns.

The results of this test are tabulated in Table 4. Panel A tabulates the VW out-of-sample monthly total returns generated by the Low group and High group, as well as the High-minus-Low spread. In columns (1) to (3) and (4) to (6), the training goal is to predict the trading and holding-based returns of investors, respectively. In addition to training goals, we also differentiate the impact of different predictors. Similar to the previous table, the first three rows demonstrate the prediction power of FNN when different sets of predictors (i.e., firm characteristics or behavioral biases) are used. The last row reports the results for the residual neural network.

Across columns (1) to (3), we observe that trading-based returns are mainly associated with the poor performance of Low-group investors (but not the good performance of High-group investors). Furthermore, the difference between the first and remaining rows suggests that behavioral biases seem to provide the most important ground to generate such poor performance.

Conversely, columns (4) to (6) suggest that focusing on holding-based returns allows neural networks to identify High-group investors who can deliver superior out-of-sample returns. Interestingly, the inclusion of behavioral biases as additional predictors does not enhance the predicting power on the winner side. Instead, behavioral biases again seem to contribute mostly to the identification of Low-group investors and, as a result, enhance the High-minus-Low spread.

Panels B and C present the risk-adjusted performance of the Low group and High group, as well as the High-minus-Low spread. The layout of these panels is the same as Panel A. We observe that our results remain highly robust under risk adjustments.

Collectively, behavioral biases seem to hurt investors' welfare, whereas proper holding-based firm characteristics may contribute to investor welfare. It is especially interesting to see that poor trading-based returns are mostly induced by behavioral biases, which also predict poor total returns. Behavioral biases, in this perspective, harm investor welfare by inducing poor judgments in investors' buying and selling decisions.

### C. Removing More Microcap Stocks

In the main analysis, we exclude 30% of small stocks because these stocks are difficult to trade in emerging markets (Liu et al. 2019) and may overstate the predicting power of machine learning models (Avramov, Cheng, and Metzker 2019 and Cong et al., 2020).

Table 5 examines whether our main conclusions are robust to different thresholds of removal (e.g., 20% or 40%). Columns 1-3 tabulate the raw or risk-adjusted *High*-minus-*Low* return spread when 20% of small stocks are removed. Columns 4-6 report similar results when 40% of small stocks are removed. Neural Network models (FNN and ResNN) continue to generate significantly out-of-sample returns, which dominate all other alternative machine-learning or OLS models. These patterns suggest that our conclusions derived from Neural Network models are robust to the less or more strict control of small stocks.

It is also worth noting that ResNN outperforms FNN in both cases. For instance, the two models deliver, respectively, 4.4% and 3.4% value-weighted *High*-minus-*Low* return spread when 40% microcap stocks are removed. Hence, ResNN outperforms FNN by almost 30% in generating out-of-sample returns. Indeed, across all empirical specifications, ResNN outperforms FNN in its overall performance. This difference confirms that ResNN may provide a particularly useful tool for return predictive analysis.

### D. Model Comparisons in Predicted Performance

We lastly examine the differences in predicted performance across various models. Since the two neural network models outperform other models, we first investigate the difference between FNN and other non-neural network models. We then move on to tabulate the difference between the two neural network models. Our analysis focuses on the High-minus-Low return spread as the overall performance measure of each model. To assess the robustness of our results, we also systematically remove 20%, 30%, and 40% of small stocks and use the local three or four factors to adjust the High-minus-Low return spread.

The results are tabulated in Table 6. The first four rows compare FNN to other non-neural network models. Regardless of the sample we use or the factor models we use to adjust returns, the High-minus-Low return spread of FNN significantly outperforms other non-neural network models across all empirical specifications. These observations reveal the striking predictive power of neural networks in general.

As striking is the observation in the last row, which compares the two neural network models: it reports that ResNN significantly outperforms FNN across all empirical specifications. Indeed, ResNN can generate between 9 to 13 basis points (bps) more in monthly four-factor adjusted alpha than FNN. These results confirm the usefulness, if not the potential superiority, of the residual neural network for analyzing our large sample of retail investors.

## Conclusions

This paper employs various machine learning models to analyze the returns for millions of retail investors in India. We observe that Neural Network outperforms other models, including traditional linear OLS models, in predicting investor returns. In particular, the more recently developed Residual Neural Network (*ResNN*) exhibits superior power in identifying both good and bad out-of-sample performance. Such a predicting power suggests that Neural Network models comprehend important information about investors that contributes to their returns.

We further conduct variable gradient analysis, which indicates that behavioral biases in general play a more important role than holding-weighted stock characteristics. This analysis enables us to identify the most important behavioral biases and stock characteristics that affect retail investors' investment returns. Indeed, we identify diversification, portfolio turnover, and momentum as the leading factors to influence overall retail returns. Additionally, turnover, the disposition effect, and diversification emerge as the three most important factors in predicting new trading-generated returns.

Our results call for further research, potentially utilizing state-of-the-art machine learning tools, to comprehensively understand the relative importance of behavioral biases and firm characteristics to retail investors.

## References


Acharya, V.V., Steffen, S., 2020. The risk of being a fallen angel and the corporate dash for cash in the midst of COVID. *The Review of Corporate Finance Studies*, 9(3), pp.430-471.

An, L., 2016. Asset pricing when traders sell extreme winners and losers. *Review of Financial Studies* 29, pp.823-861.

An, L., Engelberg, J., Henriksson, M., Wang, B. and Williams, J., 2019. The portfolio-driven disposition effect. *PBCSF-NIFR Research Paper.*

Avramov, D., Cheng, S., and Metzker, L. (2023). Machine learning vs. economic restrictions: Evidence from stock return predictability. Management Science, 69(5):2587–2619.

Avramov, D., Cheng, S., Metzker, L., and Voigt, S. (2023). Integrating factor models. The Journal of Finance, 78(3):1593–1646.

Balasubramaniam, V., Campbell, J. Y., Ramadorai, T., & Ranish, B. (2023). Who owns what? A factor model for direct stockholding. The Journal of Finance, 78(3), 1545-1591.

Bali, T. G., Beckmeyer, H., Morke, M., and Weigert, F. (2023). Option Return Predictability with Machine Learning and Big Data. The Review of Financial Studies, 36(9):3548–3602.

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(2):773–806.

Barber, B. M. and Odean, T. (2013). Chapter 22 - the behavior of individual investors. volume 2 of Handbook of the Economics of Finance, 1533–1570. Elsevier.

Barberis, N., & Xiong, W. (2009). What drives the disposition effect? An analysis of a long‐standing preference‐based explanation. the Journal of Finance, 64(2), 751-784.

Barberis, N. (2018). Chapter 2 - psychology-based models of asset prices and trading volume., Handbook of Behavioral Economics - Foundations and Applications 1, 79–175. North-Holland.

Ben-David, I., & Hirshleifer, D. (2012). Are investors really reluctant to realize their losses? Trading responses to past returns and the disposition effect. The Review of Financial Studies, 25(8), 2485-2532.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. The American Economic Review, 91(1):79–98.

Bertrand, M., Mehta, P., & Mullainathan, S. (2002). Ferreting out tunneling: An application to Indian business groups. The Quarterly Journal of Economics, 117(1), 121-148.

Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. The Review of Financial Studies, 34(2), 1046-1089.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and asset prices. American Economic Review, 103(3):623–28.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. Quarterly Journal of Economics, 135(3):1399–1442.

Bryzgalova, S., Pelger, M., & Zhu, J. (2020). Forest through the trees: Building cross-sections of stock returns. Available at SSRN 3493458.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the financial sophistication of households. American Economic Review, 99(2), 393-398.

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

Chang, T. Y., Solomon, D. H., & Westerfield, M. M. (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. The Journal of Finance, 71(1), 267-302.

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

Choi, J. J., & Robertson, A. Z. (2020). What matters to individual investors? Evidence from the horse's mouth. The Journal of Finance, 75(4), 1965-2020.

Cong, L., Tang, K., Wang, J. & Zhang, Y., 2021, AlphaPortfolio: Direct Construction Through Deep Reinforcement Learning and Interpretable AI, working paper.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. Journal of Economic literature, 47(2), 315-372.

DellaVigna, S. (2018). Structural behavioral economics. In Handbook of Behavioral Economics: Applications and Foundations 1 (Vol. 1, pp. 613-723). North-Holland.

Cochrane, J. H. (2011). Presidential address: Discount rates. The Journal of Finance, 66(4):1047–1108.

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(3):103737.

Døskeland, T., and Hvide, H.K., 2011, Do individual investors have asymmetric information based on work experience?, Journal of Finance 66, 1011-1041.
Easley, D., Lopez de Prado, M., O'Hara, M., and Zhang, Z. (2021). Microstructure in the Machine Age. The Review of Financial Studies, 34(7):3316–3363.
Fama, E. F. and French, K. R. (1998). Value versus growth: The international evidence. The Journal of Finance, 53(6):1975–1999.
Filippou, I., Rapach, D., Taylor, M. P., & Zhou, G. (2023). Out-of-sample exchange rate prediction: A machine learning perspective. Available at SSRN 3455713.
Fischbacher, U., Hoffmann, G., & Schudy, S. (2017). The causal effect of stop-loss and take-gain orders on the disposition effect. The Review of Financial Studies, 30(6), 2110-2129.
Freyberger, J., Neuhierl, A., & Weber, M. (2020). Dissecting characteristics nonparametrically. The Review of Financial Studies, 33(5), 2326-2377.
Frydman, C., & Rangel, A. (2014). Debiasing the disposition effect by reducing the saliency of information about a stock's purchase price. Journal of economic behavior & organization, 107, 541-552.
Frydman, C., & Wang, B. (2020). The impact of salience on investor behavior: Evidence from a natural experiment. The Journal of Finance, 75(1), 229-276.
Gopalan, R., Nanda, V., & Seru, A. (2007). Affiliated firms and financial support: Evidence from Indian business groups. Journal of Financial Economics, 86(3), 759-795.
Grinblatt, M., & Han, B. (2005). Prospect theory, mental accounting, and momentum. Journal of financial economics, 78(2), 311-339.
Grinblatt, M., Ikäheimo, S., Keloharju, M., Knüpfer, S., (2016) IQ and Mutual Fund Choice. Management Science 62(4):924-944.
Gu, S., Kelly, B., and Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. The Review of Financial Studies, 33(5):2223–2273.
Harvey, C. R., Liu, Y., and Zhu, H. (2015). . . . and the Cross-Section of Expected Returns. The Review of Financial Studies, 29(1):5–68.
He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778.
Henderson, V. (2012). Prospect theory, liquidation, and the disposition effect. Management Science, 58(2), 445-460.
Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. ArXiv, abs/1207.0580.
Hirshleifer, D. (2015). Behavioral finance. Annual Review of Financial Economics, 7(1):133–159.
Horel, E. and Giesecke, K. (2020). Significance tests for neural networks. J. Mach. Learn. Res., 21(1).
Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5):359–366.
Hou, K., Xue, C., and Zhang, L. (2020). Replicating anomalies. Review of Financial Studies, 33(5):2019–2133.
Ivoovic, Z. and Weisbenner, S. (2005). Local does as local is: Information content of the geography of individual investors' common stock investments. The Journal of Finance, 60(1):267–306.
Ivković, Z., & Weisbenner, S. (2009). Individual investor mutual fund flows. Journal of Financial Economics, 92(2), 223-237.
Jensen, T. I., Kelly, B. T., Malamud, S., & Pedersen, L. H. (2022). Machine learning and the implementable efficient frontier. Swiss Finance Institute Research Paper, (22-63).
Jensen, T. I., Kelly, B., and Pedersen, L. H. (2023). Is there a replication crisis in finance? The Journal of Finance, 78(5):2465–2518.
Kaniel, R., Lin, Z., Pelger, M., and Van Nieuwerburgh, S. (2023). Machine-learning the skill of mutual fund managers. Journal of Financial Economics, 150(1):94–138.
Karolyi, G. A., & Van Nieuwerburgh, S. (2020). New methods for the cross-section of returns. The Review of Financial Studies, 33(5), 1879-1890.
Kaustia, M. (2010). Prospect theory and the disposition effect. Journal of Financial and Quantitative Analysis, 45(3), 791-812.
Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics, 134(3), 501-524.
Kelly, B., & Xiu, D. (2023). Financial machine learning. Foundations and Trends® in Finance, 13(3-4), 205-363.
Kumar, A. (2009). Who gambles in the stock market? The Journal of Finance, 64(4):1889–1933.

Leippold, M., Wang, Q., and Zhou, W. (2022). Machine learning in the chinese stock market. Journal of Financial Economics, 145(2):64–82.

Li, B., & Rossi, A. G. (2020). Selecting mutual funds from the stocks they hold: A machine learning approach. Available at SSRN 3737667.

Li, B., A. Rossi, X. Yan and L. Zheng (2023). Real-Time Machine Learning in the Cross-Section of Stock Returns, working paper.

Li, Y., & Yang, L. (2013). Prospect theory, the disposition effect, and asset prices. Journal of Financial Economics, 107(3), 715-739.

Lilienfeld‐Toal, U. V., Mookherjee, D., & Visaria, S. (2012). The distributive impact of reforms in credit enforcement: Evidence from Indian debt recovery tribunals. Econometrica, 80(2), 497-558.

Liu, H., Peng, C., Xiong, W. A., and Xiong, W. (2022). Taming the bias zoo. Journal of Financial Economics, 143(2):716–741.

Liu, J., Stambaugh, R. F., and Yuan, Y. (2019). Size and value in china. Journal of Financial Economics, 134(1):48–69.

Lusardi, A. and Mitchell, O. S. (2011). Financial literacy around the world: an overview. Journal of Pension Economics & Finance, 10(4):497–508.

Massa, M. and Simonov, A. (2006). Hedging, familiarity and portfolio choice. Review of Financial Studies, 19(2):633–685.

Mclean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? The Journal of Finance, 71(1):5–32.

Odean, T. (1998). Are investors reluctant to realize their losses? The Journal of Finance, 53(5):1775–1798.

Sadhwani, A., Giesecke, K., and Sirignano, J. (2020). Deep Learning for Mortgage Risk*. Journal of Financial Econometrics, 19(2):313–368.

Shefrin, H. and Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. The Journal of Finance, 40(3):777–790.

Sui, Pengfei and Wang, Baolian, Stakes and Investor Behaviors (August 27, 2023). Available at SSRN: https://ssrn.com/abstract=4219861 or http://dx.doi.org/10.2139/ssrn.4219861

Van Binsbergen, J. H., Han, X., & Lopez-Lira, A. (2023). Man versus machine learning: The term structure of earnings expectations and conditional biases. The Review of Financial Studies, 36(6), 2361-2396.

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

## Table 1: Variable Names and Explanations

This table tabulates the list of firm characteristics and behavioral biases that we use in our analysis. In Panel A, the first six categories represent firm characteristics and the last consists of behavioral biases. Panel B summarizes the literature on behavioral biases.

Panel A: Variable Names and Explanations

| Name | Explanation | Name | Explanation |
|---|---|---|---|
| ***Profitability*** | | ***Investor Behavioral Bias*** | |
| ROA | Return on assets | Distance | Local Bias |
| NSOLA | Net Sales Over Lagged Assets | Port_Value | Investor’s Portfolio Value |
| COGS | Cost of Goods Sold over lagged assets | Extrapolation | Extrapolation |
| SaleGrow | Sales Growth | Disp | Disposition Effect |
| | | Month_Diver | Diversification |
| ***Past Returns*** | | | |
| R1_0 | Last month return | Investor_Tvr | Investor Turnover |
| R2_1 | Return from t-2 to t-1 | Past_Perform | Investor Past Performance |
| R12_7 | Intermediate momentum | IVOL | Idiosyncratic Volatility (Proxy for Lottery Preference) |
| R12_2 | Momentum | Low_Price | Low Price Rank (Proxy for Lottery Preference) |
| | | High_Price | High Price Rank (Proxy for Lottery Preference) |
| ***Investments*** | | | |
| DPI2A | Change in property, plants, and equipment | Open_Price | Open Price Rank (Proxy for Lottery Preference) |
| NI | Net Share Issues | Close Price | Close Price Rank (Proxy for Lottery Preference) |
| | | Skew | Idiosyncratic Skewness (Proxy for Lottery Preference) |
| ***Intangibles*** | | | |
| NIA | Net Intangible Asset | | |
| ***Value*** | | | |
| TobinQ | Tobin’s Q | | |
| Div_Yield | Dividend Yield | | |
| EPS | Earnings Per Share | | |
| BVPS | Book Value Per Share | | |
| PE | Price to Earnings | | |
| PB | Price to Book Value | | |
| ***Trading Frictions*** | | | |
| TA | Total Asset | | |
| Size | Market Equity | | |
| Turnover | Monthly Turnover | | |
| TradVol | Monthly Trading Volume | | |
| Leverage | Financial leverage | | |
| NOE | Number of Employees Growth | | |

Panel B: Reference of the behavioral bias used in our analysis.

| Bias | Proxy | Papers |
|---|---|---|
| The disposition effect | Regression coefficient | Shefrin and Statman (1985), Odean (1998), Ben-David and Hirshleifer (2012) |
| Lottery preference | Ivol<br>Iskew<br>Stock price | Kumar (2009), Harvey and Siddique (2000), Bordalo, Gennaioli, and Shleifer (2012; 2013; 2020) |
| Extrapolation | Excess return of holding stocks | Barber and Odean (2013) |
| Underdiversification | Number of stocks in an investor's portfolio | Barber and Odean (2000), Benartzi and Thaler (2001), Lusardi and Mitchell (2011) |
| Local bias | Average distance between an investor's location and the headquarters of the stocks the investor bought | Ivkovic and Weisbenner (2005), Massa and Simonov (2006) |
| Turnover | The frequency of trading for the investors | Odean (1998), Barber and Odean (2000) |

## Table 2. Model Comparison: Predict Total Return

This table reports the performance comparison of various models, including Linear, Lasso, Ridge, Random Forest, the standard Feedforward Neural Network (FNN), and the enhanced Residual Neural Network (ResNN). We first use each model to predict investor returns and then sort retail investors into five quintiles according to predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among all investors, respectively. Finally, we report the value-weighted out-of-sample returns of the high and low groups as well as their return difference. we also use the locally estimated three-factor and four-factor models to adjust these returns. The OLS t-statistics are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | Low Group | | | High Group | | | High Minus Low | | |
| | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 |
| Linear | -0.018* | -0.017 | -0.014 | 0.007 | 0.007 | 0.008 | 0.025*** | 0.024*** | 0.022*** |
| | (-1.77) | (-1.63) | (-1.40) | (1.37) | (1.33) | (1.52) | (3.34) | (3.14) | (2.95) |
| Lasso | -0.008 | -0.006 | -0.004 | 0.008 | 0.008 | 0.010* | 0.016** | 0.014* | 0.013* |
| | (-0.75) | (-0.53) | (-0.34) | (1.56) | (1.61) | (1.90) | (2.06) | (1.80) | (1.68) |
| Ridge | -0.018* | -0.017 | -0.014 | 0.007 | 0.007 | 0.008 | 0.025*** | 0.024*** | 0.022*** |
| | (-1.76) | (-1.61) | (-1.38) | (1.36) | (1.32) | (1.52) | (3.32) | (3.10) | (2.91) |
| Random Forest | -0.015 | -0.014 | -0.012 | 0.011 | 0.009 | 0.007 | 0.026 | 0.023 | 0.019 |
| | (-1.24) | (-1.17) | (-1.03) | (1.08) | (0.90) | (0.70) | (1.47) | (1.59) | (1.63) |
| FNN | -0.025* | -0.018 | -0.015 | 0.015*** | 0.015*** | 0.017*** | 0.040*** | 0.033*** | 0.031*** |
| | (-1.74) | (-1.62) | (-1.43) | (2.66) | (2.70) | (3.08) | (3.38) | (3.24) | (3.11) |
| Residual Neural Network | -0.031** | -0.029** | -0.026** | 0.012** | 0.013** | 0.014** | 0.044*** | 0.042*** | 0.041*** |
| | (-2.38) | (-2.17) | (-2.00) | (2.00) | (2.08) | (2.33) | (4.57) | (4.38) | (4.26) |

**Table 3. Information Set Comparison: Holding-Based Characteristics vs. Investor Behavioral Biases**

This table reports the performance of Neural Network algorithms when different subsets of predictors (i.e., firm characteristics or investor behavioral biases) are used alone or jointly to predict returns. The first (second) line uses the standard Feedforward Neural Network (FNN) to predict investor returns based on firm characteristics (behavioral biases) only. The third line allows FNN to use both firm characteristics and investor behavioral biases to predict returns. The last line utilizes the Residual Neural Network (ResNN) using both sets of predictors. For each model-predictor combination (e.g., FNN using only firm characteristics), we sort retail investors into five quintiles according to their predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among investors, respectively. Finally, we report the value-weighted out-of-sample returns of the high and low groups as well as their return difference. We also use the locally estimated three-factor and four-factor models to adjust these returns. The OLS t-statistics are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | | Low Group | | | High Group | | | High Minus Low | |
| | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 |
| Stock Characteristics | -0.003 | -0.001 | 0.001 | 0.006 | 0.006 | 0.008 | 0.009 | 0.007 | 0.007 |
| | (-0.23) | (-0.08) | (0.09) | (1.01) | (1.02) | (1.32) | (0.99) | (0.79) | (0.72) |
| Behavioral Biases | -0.025** | -0.024** | -0.022** | 0.008 | 0.009 | 0.010* | 0.033*** | 0.033*** | 0.032*** |
| | (-2.62) | (-2.43) | (-2.25) | (1.46) | (1.55) | (1.80) | (5.63) | (5.41) | (5.29) |
| Stock Chars + Behavioral | -0.025* | -0.018 | -0.015 | 0.015*** | 0.015*** | 0.017*** | 0.040*** | 0.033*** | 0.031*** |
| | (-1.74) | (-1.62) | (-1.43) | (2.66) | (2.70) | (3.08) | (3.38) | (3.24) | (3.11) |
| Residual Neural Network | -0.031** | -0.029** | -0.026** | 0.012** | 0.013** | 0.014** | 0.044*** | 0.042*** | 0.041*** |
| | (-2.38) | (-2.17) | (-2.00) | (2.00) | (2.08) | (2.33) | (4.57) | (4.38) | (4.26) |

**Table 4. Predictor Comparison when Predicting Holding and Trading-based Returns**

This table reports the performance of Neural Network algorithms when different subsets of predictors (i.e., firm characteristics or investor behavioral biases) are used alone or jointly to predict returns. In each panel, the training objectives are to predict total return (columns 1-3) and its two components: trading return (columns 4-6) and holding-based return (columns 7-9). The first (second) line uses the standard Feedforward Neural Network (FNN) to predict investor returns based on firm characteristics (behavioral biases) only. The third line allows FNN to use both firm characteristics and investor behavioral biases to predict returns. The last line utilizes the Residual Neural Network (ResNN) using both sets of predictors. For each model-predictor-training goal combination (e.g., FNN using only firm characteristics to predict trading returns), we sort retail investors into five quintiles according to their predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among investors, respectively. Panel A reports the value-weighted out-of-sample *total returns* of the high and low groups as well as their return difference. Panels B and C further report the risk-adjusted total returns based on the locally estimated three-factor and four-factor models. The OLS t-statistics are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Excess Return

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Trading Return | | | Holding Return | | |
| | Low | High | High Minus Low | Low | High | High Minus Low |
| Stock Characteristics | -0.000 | -0.006 | -0.006 | -0.005 | 0.012** | 0.017* |
| | (-0.05) | (-0.71) | (-1.11) | (-0.38) | (1.99) | (1.96) |
| Behavioral Biases | -0.031*** | 0.003 | 0.034*** | -0.022** | 0.010* | 0.031*** |
| | (-4.01) | (0.42) | (17.75) | (-2.00) | (1.69) | (4.32) |
| Stock Chars + Behavioral | -0.034*** | 0.003 | 0.037*** | -0.019 | 0.012** | 0.031** |
| | (-4.10) | (0.43) | (13.81) | (-1.16) | (2.18) | (2.35) |
| Residual Neural Network | -0.028*** | 0.003 | 0.031*** | -0.023 | 0.012** | 0.035*** |
| | (-3.67) | (0.39) | (14.69) | (-1.61) | (2.25) | (3.04) |

Panel B: Fama-French Three Factor Adjusted Alpha

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Trading Return | | | Holding Return | | |
| | Low | High | High Minus Low | Low | High | High Minus Low |
| Stock Characteristics | 0.001 | -0.005 | -0.006 | -0.003 | 0.013** | 0.016* |
| | (0.09) | (-0.60) | (-1.13) | (-0.26) | (2.12) | (1.85) |
| Behavioral Biases | -0.029*** | 0.004 | 0.034*** | -0.020* | 0.010* | 0.031*** |
| | (-3.73) | (0.58) | (17.00) | (-1.82) | (1.79) | (4.13) |
| Stock Chars + Behavioral | -0.032*** | 0.004 | 0.036*** | -0.019 | 0.013** | 0.032** |
| | (-3.82) | (0.59) | (13.23) | (-1.12) | (2.27) | (2.34) |
| Residual Neural Network | -0.027*** | 0.004 | 0.031*** | -0.021 | 0.013** | 0.034*** |
| | (-3.40) | (0.56) | (14.20) | (-1.48) | (2.33) | (2.90) |

Panel C: Carhart Four Factor Adjusted Alpha

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Trading Return | | | Holding Return | | |
| | Low | High | High Minus Low | Low | High | High Minus Low |
| Stock Characteristics | 0.003 | -0.004 | -0.006 | 0.000 | 0.015** | 0.015* |
| | (0.35) | (-0.41) | (-1.17) | (0.03) | (2.57) | (1.68) |
| Behavioral Biases | -0.028*** | 0.006 | 0.033*** | -0.018 | 0.012** | 0.030*** |
| | (-3.55) | (0.84) | (16.79) | (-1.65) | (2.09) | (4.04) |
| Stock Chars + Behavioral | -0.031*** | 0.006 | 0.036*** | -0.015 | 0.014** | 0.029** |
| | (-3.65) | (0.84) | (13.12) | (-0.90) | (2.62) | (2.17) |
| Residual Neural Network | -0.025*** | 0.006 | 0.031*** | -0.020 | 0.014*** | 0.034*** |
| | (-3.22) | (0.82) | (14.07) | (-1.35) | (2.74) | (2.87) |

## Table 5. Robustness Checks on Removing 20% or 40% of Small Stocks

This table reports the performance comparison of various models, including Linear, Lasso, Ridge, Random Forest, the standard Feedforward Neural Network (FNN), and the enhanced Residual Neural Network (ResNN). Different from our main analysis, we remove the bottom 40% of small stocks in this robustness check. We first use each model to predict investor returns and then sort retail investors into five quintiles according to predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among all investors, respectively. Finally, we report the value-weighted out-of-sample returns of the high and low groups as well as their return difference. we also use the locally estimated three-factor and four-factor models to adjust these returns. The OLS t-statistics are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | High Minus Low (Removing 20%) | | | High Minus Low (Removing 40%) | | |
| | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 |
| Linear | 0.014 | 0.012 | 0.013 | 0.013 | 0.021 | 0.018 |
| | (1.71) | (1.61) | (1.57) | (1.52) | (1.21) | (2.00) |
| Lasso | 0.022* | 0.017* | 0.015 | 0.022* | 0.024* | 0.026 |
| | (1.87) | (1.78) | (1.22) | (1.87) | (1.78) | (1.22) |
| Ridge | 0.023** | 0.019* | 0.017* | 0.021** | 0.021** | 0.020** |
| | (2.27) | (2.00) | (1.88) | (2.57) | (2.27) | (2.22) |
| Random Forest | 0.023 | 0.017 | 0.016 | 0.024 | 0.012 | 0.012 |
| | (1.15) | (1.04) | (1.00) | (1.02) | (1.31) | (1.15) |
| FNN | 0.034*** | 0.032*** | 0.029*** | 0.034*** | 0.032*** | 0.028*** |
| | (3.67) | (3.53) | (3.18) | (3.86) | (3.47) | (3.12) |
| Residual Neural Network | 0.045*** | 0.040*** | 0.038*** | 0.044*** | 0.042*** | 0.041*** |
| | (4.59) | (4.48) | (4.32) | (4.78) | (4.66) | (4.37) |

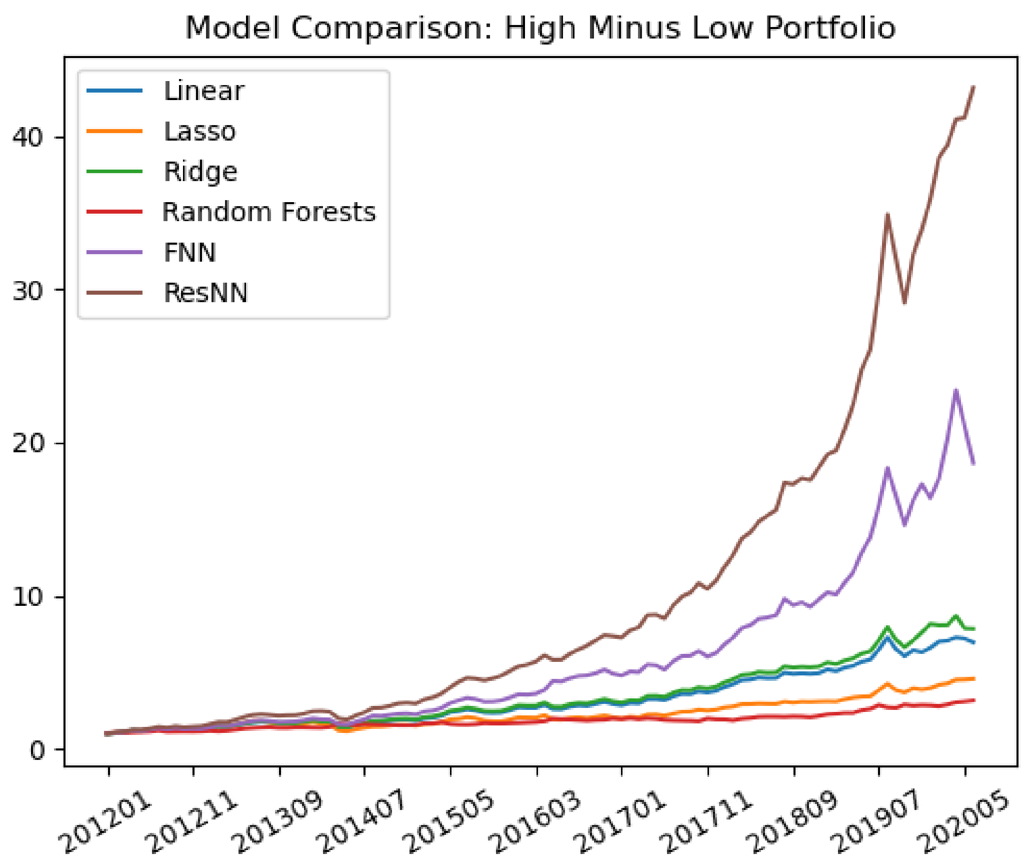
**Table 6. Model Comparisons Across Models**

This table reports the difference in performance across various models, including Linear, Lasso, Ridge, Random Forest, the standard Feedforward Neural Network (FNN), and the enhanced Residual Neural Network (ResNN). In conducting the comparison, we remove 20%, 30%, and 40% of small stocks. Each model is used to predict investor returns and sort retail investors into five quintiles according to predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among all investors, respectively. The value-weighted out-of-sample High-minus-Low return spread is calculated for each model. The first four rows report the difference between the High-minus-Low return spread of FNN and that of Linear, Lasso, Ridge, and Random Forest models. The last row reports the difference between ResNN and FNN. we also use the locally estimated three-factor and four-factor models to adjust these returns. The OLS t-statistics are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | High Minus Low (Excluding 20% Small Stocks) | | | High Minus Low (Excluding 30% Small Stocks) | | | High Minus Low (Excluding 40% Small Stocks) | | |
| | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 | Mean | FF-3 | Carhart-4 |
| FNN - Linear | 0.020*** | 0.020*** | 0.016*** | 0.015** | 0.015*** | 0.016*** | 0.021*** | 0.020*** | 0.016*** |
| | (2.98) | (2.86) | (2.80) | (2.37) | (2.77) | (2.84) | (3.03) | (2.94) | (2.89) |
| FNN - Lasso | 0.012** | 0.015*** | 0.014*** | 0.024*** | 0.019*** | 0.017*** | 0.012** | 0.011** | 0.008** |
| | (2.43) | (2.97) | (2.82) | (2.95) | (3.11) | (3.17) | (2.37) | (2.28) | (2.22) |
| FNN - Ridge | 0.011** | 0.013*** | 0.012** | 0.015** | 0.009** | 0.010** | 0.013** | 0.011** | 0.008** |
| | (2.37) | (2.80) | (2.48) | (2.37) | (2.50) | (2.41) | (2.43) | (2.38) | (2.27) |
| FNN - Random Forest | 0.011** | 0.015*** | 0.013*** | 0.014** | 0.010** | 0.013*** | 0.010** | 0.020*** | 0.016*** |
| | (2.24) | (2.98) | (2.85) | (2.33) | (2.48) | (2.85) | (2.02) | (2.91) | (2.83) |
| ResNN - FNN | 0.011** | 0.008** | 0.009*** | 0.004* | 0.009** | 0.010*** | 0.010** | 0.010** | 0.013*** |
| | (2.17) | (2.32) | (2.92) | (1.79) | (2.26) | (2.83) | (2.08) | (2.25) | (2.88) |

**Figure 1: The Cumulative Returns of High-minus-Low Return Spread from Various Models**

This figure plots the cumulative returns generated by the High-minus-Low investor groups predicted by various models, including Linear, Lasso, Ridge, Random Forest, the standard Neural Network algorithms (FNN), and the enhanced Residual Neural Network (ResNN). We first use each model to predict investor returns and then sort retail investors into five quintiles according to predicted returns, with the High and Low groups consisting of 20% of predicted winners and losers among all investors, respectively. We next calculate the value-weighted out-of-sample returns of the high and low groups. Finally, we plot the cumulative returns of the High-minus-Low spread for the period from Jan 2012 to June 2020.



## Figure 2: Top Variable Importance of Behavioral Bias vs. Stock Characteristics

This table reports the importance of each predictor when the Feedforward Neural Network (FNN) is used to predict retail investors' returns based on all predictors, including behavioral biases and stock characteristics. The marginal importance of a variable is computed from the variable gradient analysis: $Importance(x) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_t}\sum_{i=1}^{N_t}\left(\frac{\partial R_{i,t+1}^{pred}}{\partial x_{i,t}}\right)^2$, where T represents the number of periods in the data, and $N_t$ denotes the total number of investors in the t-th period. The partial derivative measures the gradient of the model's predicted output with respect to each variable. Intuitively, a larger partial derivative implies a greater influence of a variable on the model's output, indicating greater importance in predicting future returns. We computed the Importance(x) separately for the training objectives of total return, trading return, and holding return, as described in Table 4.

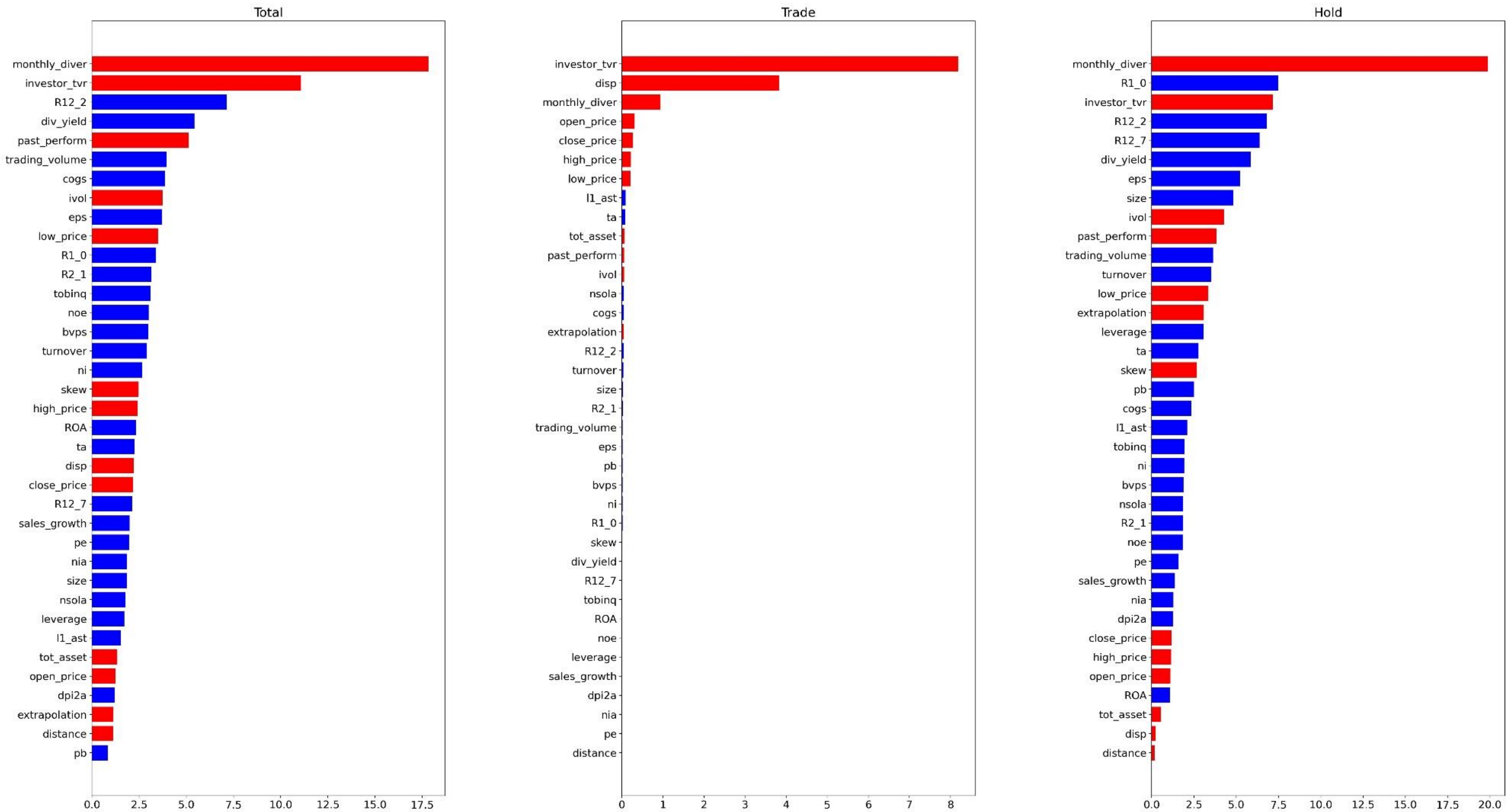
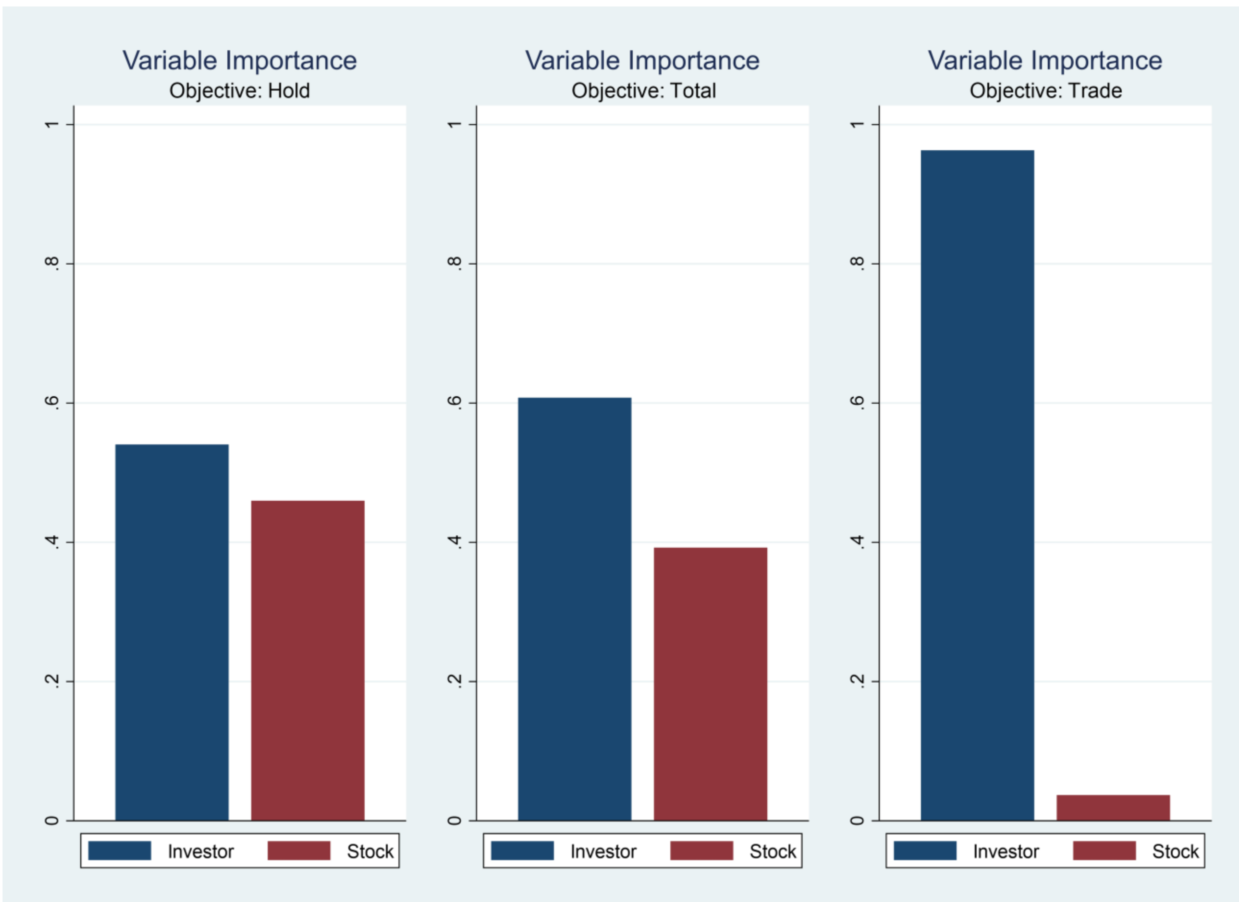

Total
monthly_diver
investor_tvr
R12_2
div_yield
past_perform
trading_volume
cogs
ivol
eps
low_price
R1_0
R2_1
tobinq
noe
bvps
turnover
ni
skew
high_price
ROA
ta
disp
close_price
R12_7
sales_growth
pe
nia
size
nsola
leverage
l1_ast
tot_asset
open_price
dpi2a
extrapolation
distance
pb
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5
Trade
investor_tvr
disp
monthly_diver
open_price
close_price
high_price
low_price
l1_ast
ta
tot_asset
past_perform
ivol
nsola
cogs
extrapolation
R12_2
turnover
size
R2_1
trading_volume
eps
pb
bvps
ni
R1_0
skew
div_yield
R12_7
tobinq
ROA
noe
leverage
sales_growth
dpi2a
nia
pe
distance
0 1 2 3 4 5 6 7 8
Hold
monthly_diver
R1_0
investor_tvr
R12_2
R12_7
div_yield
eps
size
ivol
past_perform
trading_volume
turnover
low_price
extrapolation
leverage
ta
skew
pb
cogs
l1_ast
tobinq
ni
bvps
nsola
R2_1
noe
pe
sales_growth
nia
dpi2a
close_price
high_price
open_price
ROA
tot_asset
disp
distance
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5 20.0

**Figure 3: The Relative Importance of Behavioral Bias vs. Firm Characteristics**

This figure plots the relative importance of behavioral biases vis-à-vis firm characteristics. Based on our earlier delineation, where predictors are categorized into Investor Behavioral Bias and Firm Characteristics, we define the variable importance measure of a group by computing the average of the importance measures within that group, which can also be expressed as the joint explanatory power of all predictors falling into each category. Without loss of generality, we normalize the variable importance to sum up to 1.



# Online Appendix

## Appendix Table 1: Summary Statistics of Main Variables.

Panel A: Holding based firm characteristics

| Variable | N | Mean | SD | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|---|---|
| Dpi2a | 1.523e+09 | -0.324 | 0.750 | -1.342 | -0.893 | -0.339 | 0.185 | 0.684 |
| R1_0 | 1.523e+09 | 0.0630 | 0.719 | -0.930 | -0.417 | 0.0980 | 0.566 | 0.994 |
| R2_1 | 1.523e+09 | 0.0460 | 0.722 | -0.949 | -0.441 | 0.0780 | 0.553 | 0.984 |
| R12_7 | 1.523e+09 | 0.0330 | 0.738 | -0.987 | -0.473 | 0.0560 | 0.568 | 0.994 |
| R12_2 | 1.523e+09 | 0.0230 | 0.747 | -1.019 | -0.496 | 0.0450 | 0.575 | 0.999 |
| Net Issue | 1.523e+09 | 1.192 | 0.533 | 0.467 | 0.974 | 1.365 | 1.582 | 1.671 |
| Nsola | 1.523e+09 | -0.371 | 0.742 | -1.376 | -0.952 | -0.394 | 0.134 | 0.615 |
| Cogs | 1.523e+09 | -0.404 | 0.738 | -1.428 | -0.966 | -0.423 | 0.0710 | 0.580 |
| ROA | 1.523e+09 | 0.0530 | 0.702 | -0.839 | -0.441 | 0.0500 | 0.526 | 0.995 |
| Sales Growth | 1.523e+09 | 0.124 | 0.711 | -0.870 | -0.301 | 0.159 | 0.578 | 1.050 |
| Nia | 1.523e+09 | 0.725 | 0.744 | -0.336 | 0.321 | 0.861 | 1.318 | 1.559 |
| PB | 1.523e+09 | 0.171 | 0.733 | -0.856 | -0.330 | 0.241 | 0.709 | 1.079 |
| PE | 1.523e+09 | 0.120 | 0.663 | -0.760 | -0.294 | 0.132 | 0.536 | 0.989 |
| BVPS | 1.523e+09 | 0.323 | 0.739 | -0.746 | -0.158 | 0.419 | 0.882 | 1.213 |
| EPS | 1.523e+09 | 0.253 | 0.802 | -0.960 | -0.295 | 0.367 | 0.867 | 1.239 |
| Leverage | 1.523e+09 | 0.0250 | 0.719 | -0.917 | -0.456 | 0.00400 | 0.533 | 1.030 |
| TobinQ | 1.523e+09 | 0.163 | 0.725 | -0.840 | -0.354 | 0.198 | 0.695 | 1.090 |
| Div Yield | 1.523e+09 | 0.497 | 0.771 | -0.651 | 0.0500 | 0.640 | 1.085 | 1.404 |
| TA | 1.523e+09 | 1.190 | 0.516 | 0.471 | 0.985 | 1.363 | 1.557 | 1.650 |
| Size | 1.523e+09 | 1.141 | 0.514 | 0.411 | 0.898 | 1.283 | 1.529 | 1.657 |
| Turnover | 1.523e+09 | 1.192 | 0.506 | 0.498 | 0.973 | 1.349 | 1.563 | 1.669 |
| Trading Vol | 1.523e+09 | 1.227 | 0.499 | 0.560 | 1.025 | 1.389 | 1.585 | 1.677 |
| NOE | 1.523e+09 | 0.839 | 0.734 | -0.198 | 0.442 | 1.012 | 1.418 | 1.625 |

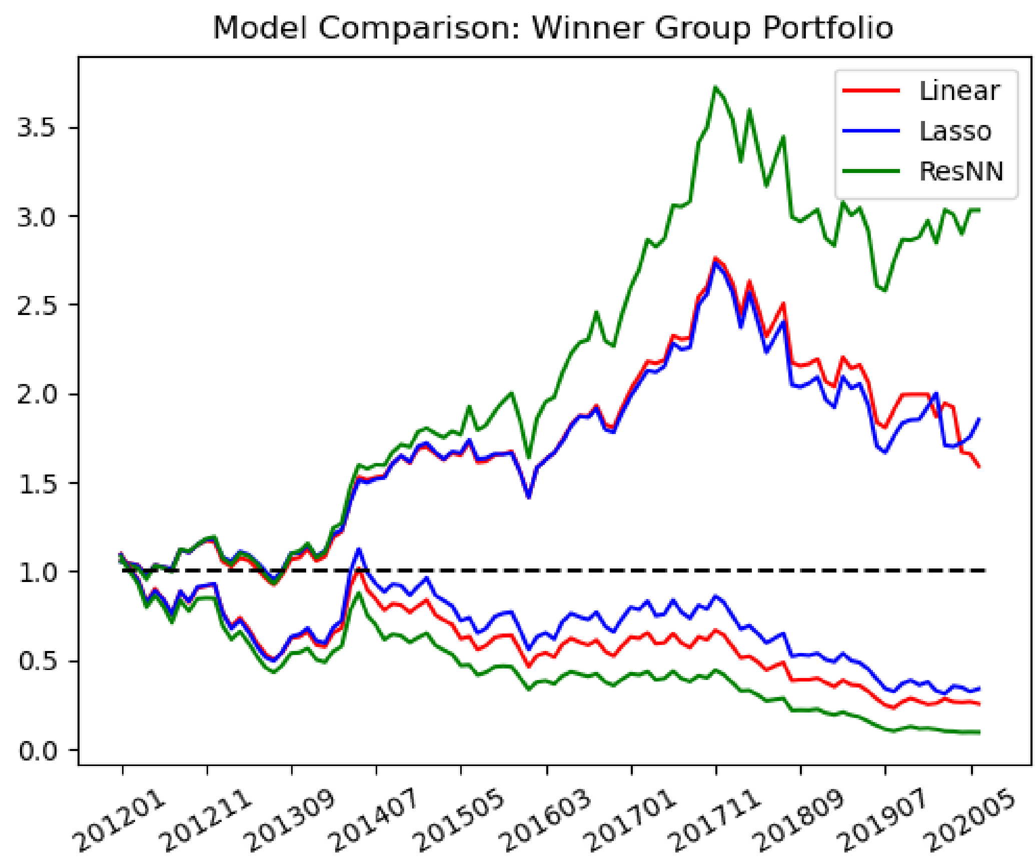
Appendix Table 1: Summary Statistics of Main Variables.

Panel B: Investor behavioral biases

| Variable | N | Mean | SD | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|---|---|
| Port. Value | 1.523e+09 | 1.187 | 0.521 | 0.462 | 0.979 | 1.363 | 1.560 | 1.651 |
| Diver | 1.523e+09 | 9.669 | 17.013 | 1.000 | 2.000 | 5.000 | 11.000 | 22.000 |
| Disp | 1.523e+09 | 0.001 | 0.008 | -0.004 | -0.001 | 0.002 | 0.003 | 0.004 |
| ivol | 1.523e+09 | -0.0570 | 1.018 | -1.416 | -0.944 | -0.157 | 0.877 | 1.360 |
| iskew | 1.523e+09 | -0.0150 | 1.031 | -1.408 | -0.924 | -0.126 | 0.915 | 1.399 |
| Distance | 1.523e+09 | 877.144 | 498.822 | 258.547 | 528.240 | 858.926 | 1185.493 | 1484.389 |
| Investor Tvr | 1.523e+09 | -0.128 | 0.909 | -0.662 | -0.627 | -0.589 | -0.534 | 1.498 |
| Open Price | 1.523e+09 | 0.317 | 0.773 | -0.852 | -0.181 | 0.426 | 0.910 | 1.258 |
| High Price | 1.523e+09 | 0.315 | 0.773 | -0.854 | -0.184 | 0.424 | 0.907 | 1.257 |
| Low Price | 1.523e+09 | 0.318 | 0.773 | -0.849 | -0.180 | 0.428 | 0.911 | 1.259 |
| Close Price | 1.523e+09 | 0.317 | 0.773 | -0.851 | -0.181 | 0.427 | 0.910 | 1.258 |
| Extrapolat~n | 1.523e+09 | 0.0380 | 0.174 | -0.155 | -0.0510 | 0.0430 | 0.135 | 0.228 |
| Past Perform | 1.523e+09 | 1.009 | 0.124 | 0.897 | 0.947 | 1.004 | 1.063 | 1.130 |

## Figure IN1: High vs. Low Group Returns from Selected Models.

Figure 1 shows the value-weighted out-of-sample returns of the high and low groups. We first use all the models to predict retail investors' total investment returns. We then sort retail investors into five quintiles according to predicted returns, with the *High* and *Low* groups consisting of 20% of predicted winners and losers among investors, respectively.



# Figure IN2: Top Variable Importance of Behavioral Bias vs. Stock Characteristics

In Figure 2, we delve deeper into the analysis of the direction of influence each variable has on the prediction outcomes. Specifically, we define the directional impact of a variable as:

$$Direction(x) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_t}\sum_{i=1}^{N_t}\frac{\partial R_{i,t+1}^{pred}}{\partial x_{i,t}}$$