When Human Meet Algorithm: the Adoption and Impact of Retail Algorithmic Trading

Pulak Ghosh* You Li† Jian Zhang‡

March 15, 2025

Abstract

We study the adoption and economic impact of artificial intelligence technology by retail investors in a developing economy. We document new facts to characterize the human-algorithm interaction in the context of retail investor trading using administrative account-level data of all individual investors from National Stock Exchange of India, the world's 8th largest stock exchange. While the retail algorithmic trading market is dominated by male investors, the relative share of female algorithmic participation increases steadily from 5% in 2012 to 10% in 2019. We find that algorithmic trades by male-young investors take up most of the overall increase in recent years and are highly procyclical to the market condition. Investors adapting to algorithmic trading experience better performance as measured by higher market-adjusted return and Sharpe ratio. The benefit is greater for less wealthy investors and those who are holding less diversified portfolio or exhibit more behavioral bias *ex ante*. We find evidence that improved performance is likely due to enhanced trading responsiveness to new market information and reduced behavioral biases. Consistent with "learning by algorithmic trading", unprofitable algorithmic traders are more likely to quit than profitable traders. Algorithmic trade size is also sensitive to past performance and retail algorithmic investors initially execute very small trades during the first few trials and increase trade size significantly after profitable trades.

JEL classification: D14, D83, G11, O33

Keywords: Algorithmic Trading, Retail Investors, Big data, Fintech, Learning

*IIM Bangalore, Department of Public Policy and Decision Sciences, Email: pulak.ghosh@iimb.ac.in
†School of Business, Macau University of Science and Technology, Email: youli@must.edu.mo
‡The University of Hong Kong, Faculty of Business and Economics, Email: zhangj1@hku.hk

1 Introduction

The past decade has seen a new technological shift with substantial developments in artificial intelligence (AI) and its wide-spread application (Furman and Seamans, 2019)[1]. As the prediction and automation technology, AI employs powerful algorithms that are capable to analyze large sets of information and self-adjust the trade execution with little human intervention, making it particularly adaptable to the investing world. As such, AI can reduce the extensive range of behavioral bias that can severely impact the investment outcome (Hirshleifer, 2015). The existing literature has been mostly focusing on characterizing the adoption of AI technology by financial professionals (Chen, Pelger and Zhu, 2022; Rossi and Utkus, 2024). However, little is known about how individual investors are interacting with AI, in particular algorithm trading, in their trading activity. In such settings, households may be better positioned to reap equity premium with the help of AI-related technology. This question is of central importance for household welfare and economic modeling (Campbell, 2006). To fill this gap, this paper provides an analysis of retail investor's uptake of algorithm in their trading activities and assesses what it implies for the investment outcome.

Combining human intelligence and AI can potentially generate considerable benefits to retail investors, given the extent to which human errors are ubiquitous. Despite the benefits of improving investment decisions, AI also have its own limitations, which potentially undercut its appeal to retail investors. First, although algorithms can be effective in identifying and responding to new patterns in the data, it cannot adapt well to rare situations and is not always convertible into profitable investment decisions. As such, humans have to play complementary roles making judgment within their capacities. Second, algorithmic trading requires traders to confer majority of autonomy from themselves to the algorithm. Given the well-known reluctance of individuals to cede all decision-making to algorithms, this can be

[1]For example, applications in financial setting include asset price prediction (Gu, Kelly and Xiu, 2020; Bai, Philippon and Savov, 2016), financial analysts (Grennan and Michaely, 2020; Abis and Veldkamp, 2020; Cao et al., 2024), Fintech innovation (Chen, Wu and Yang, 2019), loan underwriting (Jansen, Nguyen and Shams, 2020; Fuster et al., 2022), and robo-advising (D'Acunto, Prabhala and Rossi, 2019).

an additional bottleneck for investors with high level of algorithm aversion to embrace the technology (Dietvorst, Simmons and Massey, 2018; Pagliaro and Ansgar, 2022). Therefore, it remains an open question whether and how an average individual investor, who likely lacks technical or professional skills, can tap into the substantial enhancement from AI technology.

To date, the lack of micro-level data on retail investor's interaction with algorithms has posed the key challenges to understanding the adoption patterns and the economic impact of AI technologies. To circumvent this challenge, we turn to a unique and remarkably comprehensive database that contains the entire trading records of retail investors in the world's 8th largest stock exchange, National Stock Exchange of India (NSE). Most importantly, the regulator in India requires each trade emanating from algorithmic orders to be tagged with the unique identifier. This provides a rare data opportunity for researchers to distinguish trades with and without algorithm by the same investor.

Figure 1 shows the growing importance of algorithmic trading in India as well as how the share has changed over time. Specifically, as shown in Panel A, technological advances in the past decades have fueled the rapid growth of algorithmic trading and its aggregate trading value has nearly a five-fold rise over the past seven years. Furthermore, algorithmic trading as a proportion of total trading value has also evolved with marked difference, rising from 30 percent in 2012 to more than 50 percent in 2019. Since algorithmic trading is initially provided to institutional clients only, it's not surprising to see that the participation rate is much smaller among retail investors at a level below 10 percent. Still, we observe a steady increasing trend in the share of retail algorithmic trading value (to all retail trading value), which exhibits some short-term fluctuations in early phases.

We begin by depicting the investor composition of retail algorithmic trading in India and establishing some new empirical regularities. For this purpose, we explore the investor attributes in the data and categorize all retail algorithmic traders by age and gender into six different groups (i.e., male vs female, and young, mid-aged vs old). While male investors remain the dominant players in Indian retail algorithmic trading market, the relative share of female algorithmic trading participation in terms of market value is steadily rising from around 6 percent in 2012 to more than 10 percent in 2019. Across all six investor groups, male-young traders exhibit a dramatic increase in terms of both trading value and number of investors, and play the major role at the end of our sample period, possibly echoing the ease of their market accessibility as a result of Fintech disruption.

Examining the investor composition of algorithmic traders in the cross-section of stocks and market cycles also reveals interesting findings. Among investment in large stocks, we continue to observe similar pattern, whereas female traders make up a significantly higher proportion in algorithmic trading of small stocks. Investor's decision to adopt algorithm can be a function of her preferences and beliefs that likely vary with the macro economy and market conditions (Greenwood and Shleifer, 2014; Cohn et al., 2015). To assess the sensitivity of algorithmic trading participation to market movement, we distinguish booming market from bust market and document sharp differences across investor groups. Male and young traders tend to expand (contract) considerably in their usage of algorithmic trading when the market is in the up (down) status. By contrast, algorithmic trading participation by other investors is countercyclical with the market condition.

Algorithmic trading tool can presumably improve investment performance by simplifying the investing process and automating the execution. So we move to assess its effects on investment returns. Admittedly, investor's adoption decision can be endogenous and may coincide with other events that also affect their investment skills. To mitigate this concern, we conduct a within-investor analysis and compare the investment performance of both algorithmic and non-algorithmic trades by the same investor for the same period. Algorithmic trade is associated with better performance as measured by higher market-adjusted return and Sharpe ratio. Meanwhile, algorithmic trading is beneficial to investors by reducing the investment risk as the volatility of algorithmic trades is lower than that of non-algorithmic trades by the same investor. The performance improvement is not homogenous across investors. The benefit is higher for less wealthy investors and those who are holding less diversified portfolio or exhibit more behavioral bias *ex ante*. This indicates that reducing psychology-related human errors is one important channel through which algorithmic trading tool improves investor performance.

We further investigate potential drivers to better understand the sources of performance improvement following the adoption of algorithmic trading. Specifically, we examined if improved performance could be attributed to changes in two sets of investor behaviors: *enhanced trading responsiveness to new market information* and *reduced behavioral biases*. First, algorithmic trading accelerates information acquisition and automates the trading process, enabling retail investors to react swiftly to new information, such as earnings announcements. Indeed, we observe a significant increase in trading responsiveness among algorithmic trades in the post-adoption period, compared to non-algorithmic trades before adoption. Interestingly, this improvement also has a spillover effect on non-algorithmic trades after adoption, suggesting an overall enhancement in investor's ability to react to new market information.

Next, we analyze two notable behavioral biases in finance: the disposition effect and the rank effect. Single-difference tests suggest that algorithmic trading adoption is associated with decreases in both the disposition effect, the tendency to realize winners more than losers, and the rank effect, characterized by investors' inclination to sell the top-performing stocks in their portfolios. We also explore the cross-section of the effect and find that the benefits are particularly pronounced for investors who had lower reaction speeds and higher levels of bias *ex ante*. Algorithmic trading, which enables investors to automate their trades either partially or fully, has improved investment performance by reducing, though not eliminate, the frictions or biases encountered by retail investors.

Adapting to algorithmic trading is not a once-and-for-all solution and investors still need to learn to acquire their relative advantage. Our third set of results examines how investors learn by algorithmic trading. In order to understand the learning process in retail algorithmic trading, we first formulate a simple conceptual framework that allows us to derive implications for the empirical investigation. The investor decides whether to adopt algorithmic trade but is initially uncertain about her ability. She will infer her own ability by making algorithmic trades and observing the performance. The trade size during her first few trials is small and she adjusts the optimal trade sizes accordingly by increasing the trading size (decreasing size or even exiting) in response to success (failure).

We then exploit the trading records of India retail algorithmic investors to test these predictions and document evidences supporting the existence of learning process among algorithmic traders. Survival analysis suggests that some algorithmic investors quit relatively quickly (20 percent of those traders who take the first trial in algorithmic trading fail to ever trade again). Both the current and average past trading performance are significant determinants of the exit decision. Algorithmic trade size is also sensitive to past performance. They initially execute very small trades and increase (or decrease) their trade sizes after successful (unsuccessful) trades. These findings are in line with the notion that retail traders learn about their relevant ability by algorithmic trading.

Our paper contributes to at least three strains of literature. Since our paper studies investor algorithmic trading behavior, it is relevant to the broad literature on adoption of emerging FinTech tools in household finance, including digital payment (Agarwal et al., 2019; Choi and Loh, 2021), spending (D'Acunto, Rossi and Weber, 2019), credit scoring (Berg et al., 2020) and saving (Carlin, Olafsson and Pagel, 2019; Gargano and Rossi, 2022). In particular, we contribute to the literature on Fintech in the investment sphere, *robo-advisor*. Robo-advising could improve investor's welfare by allowing for easier access to financial advice at low cost and providing diversified and personalized portfolios. D'Acunto, Prabhala and Rossi (2019) is among the first few studies to examine the impact of robo-advisory on individual investors. They present evidence to highlight the benefits of adopting robo-advising, including better diversification, lower behavioral bias and improved performance. In line with this, Bianchi and Brière (2021) and Rossi and Utkus (2024) explore similar robo-adivsor platforms and confirm that investors achieve better investment performance after following robo-advice or fully delegating their asset management to robo-advisors. We examine a different type of Fintech solution, *algorithmic trading*, that can potentially help investors achieve better financial outcome.

We are also adding to the emerging literature that studies the use of algorithm in different settings (Cowgill and Tucker, 2022; Rambachan et al., 2020; Cao et al., 2024). For instance, algorithm can replace human in pricing goods and services (Calvano et al., 2020) or securities (Colliard, Foucault and Lovo, 2022). While algorithmic decision-making reduces face-to-face discrimination and bias, it is still insufficient to completely eliminate discrimination in the context of consumer lending (Bartlett et al., 2022) and criminal sentencing (Dressel and Farid, 2018). We complement these studies by examining the investor-algorithm interaction in the stock market.

Lastly, our study is related broadly to the literature on algorithmic trading. This strand of research has focused on the unique advantage brought by AT (or High-Frequency Trading) to the market participants, including faster speed (Budish, Cramton and Shim, 2015; Baron et al., 2019), information advantage (Biais, Foucault and Moinas, 2015) and better trading strategies (Hagströmer and Nordén, 2013; Brogaard, Hendershott and Riordan, 2014; O'hara, 2015; Van Kervel and Menkveld, 2019). AT can also have asset pricing implication and contribute to market liquidity (Hendershott, Jones and Menkveld, 2011; Brogaard et al., 2015; Brogaard, Hendershott and Riordan, 2019), and price efficiency (Chaboud et al., 2014; Conrad, Wahal and Xiang, 2015; Weller, 2018). Adding to these studies, we focus on a different perspective and exploit administrative data of account-level transaction records to examine how retail investors integrate algorithms into their trading activities. Within algorithmic trading literature, we are among the first few to study the adoption of AT by individual investors.

The article unfolds as follows. Section 2 provides the description of algorithmic trading and discusses the data. Section 3 presents several stylized facts about algorithmic trading, including investor composition and Section 4 discusses the implications of algorithmic trading for investment performance and trading behaviors. Section 5 sets up a simple framework and assesses the learning process in retail algorithmic trading. Section 6 concludes.

2 Institutional Details and Data

2.1 Algorithmic Trading

2.1.1 Overview

Algorithmic trading[2] is the use of an automated algorithm for the delivery and execution of trades in a pre-determined manner without any human intervention. The term "algorithm" is often defined at the most general level in describing its uses in trading. For example, algorithm refers to "a finite, deterministic and effective problem-solving method suitable for implementation as a computer program"[3]. The emphasis on the timing is also critical since algorithmic strategies are usually designed prior to the commencement of trading. Some of the algorithmic trading strategies explore the limit order book for millisecond arbitrage opportunities and operate at the ultra-high frequency, namely High-Frequency Trading (HFT). HFT are mostly accessible by institutional investors and impossible for retail investors to carry out.

Over the past 20 years, the financial world has witnessed the usage of algorithms by all types of players (i.e., funds, investment banks and other traders) to improve and execute, either entirely or partially, their trading strategies. In current era of global financial markets, algorithmic trading has become paramount to investment strategies for achieving financial goals and the market is expanding quickly. The latest Spherical Insights report estimates the value of global algorithmic trading market size to be 13.02 billion US dollars by 2021, and projects a compound annual growth rate of 13.6% for the next decade[4]. In relative terms,

---

[2]For brevity, we will use algorithmic and algo exchangeably thereafter, (e.g., algorithmic trading vs algo trading, algorithmic investor vs algo investor, etc.).
[3]https://www.sec.gov/files/Algo_Trading_Report_2020.pdf
[4]https://www.sphericalinsights.com/reports/algorithmic-trading-market

algorithmic trading accounts for around 60%–73% of the overall US equity trading as covered in the research report of Reportlinker.

The rapid expansion of algorithmic trading is fueled by the shift in both financial markets and regulatory regime. On the investors' side, there has been rising demand for quick, dependable, and efficient order execution, automatic market surveillance, and at the same time lower transaction costs. For regulators, they have introduced favorable governmental rules to regulate and monitor the algorithmic trading in the financial industry.

Compared to discretionary (or human-based) trading, algorithmic trading possesses numerous advantages that can significantly improve the trade execution. The primary benefit is that a fully automated system is involved in the execution procedure, making it substantially more efficient since little manpower is required to constantly monitor the price movement or check the news in the market. This clearly frees up the time of traders who would invest in carrying out research activities to revise and develop new strategies. Moreover, the automated execution can make real-time adjustment of leverage and risk factors in response to market dynamics, which is not possible in the case of pure human intervention. The automated strategy in algorithmic trading is typically ascertained by historical market data (via back-testing) in the first place and thus are relatively more transparent. The systematic approach to analyze the statistical properties allows for easier comparison across various strategies and determine the optimal allocation of capital.

Another obvious advantage of algorithmic trading is the immunity from human discretionary input. Since the algorithms are pre-designed and highly automated, the influence of any potential conflicts and bias can be minimized. For example, the behavioral finance literature has presented extensive empirical evidence that bias and cognitive limitation of traders (such as fear and greed) can be overwhelming during the investment process and erode the performance of a strategy (Hirshleifer, 2015). The automation feature leaves little space for the discretionary factors to distort the information processing and execution of trades.

Algorithmic trading is not necessarily readily accessible to all investors. There are at least three features that potentially deter retail investors from integrating algo-trading into their trading routine. First, developing algorithmic trading strategies goes hand in hand with proficiency in programming and scientific modeling. Fortunately, the emergency of various algo-trading platforms aims to break these barriers and enables novice traders to customize trading ideas based on common strategies (i.e., trend following, mean reversion, etc.). Still, basic knowledge and understanding of the strategies are necessary and could incur additional learning costs. Second, relative to discretionary trading, algo-trading generally comes with higher monetary costs. For example, the automated trading platform charges monthly fees for accessing data fed into quantitative strategies. For certain types of trading strategy, one is faced with account minimum requirement[5]. Third, up-taking algo-trading requires traders to confer full autonomy from themselves to the algorithm. Given the well-known reluctance of individuals to cede all decision-making to algorithms(Dietvorst, Simmons and Massey, 2018), this can be an additional bottleneck, especially for investors with high level of algorithm aversion.

The details of algo-trading strategies vary substantially across different platforms and traders. Examples include systematic trading, market making, inter-market spreading, arbitrage, or pure speculation. Though the details may be different, retail algo-traders usually begin their design of strategy by formulating hypothesis that can be tested using data observations. For example, one hypothesis may be "does the spread between two ETFs have mean-reverting behavior". To approve or disapprove the prediction is subject to back-testing based on historical data. Alternatively, retail algo-traders can refer to the machine learning approach and incorporate a large quantity of parameters or "indicators" into the trading strategy design. There exists no perfect strategy once for all. It's possible to reevaluate and revise it when the strategy "breaks down" after a period of profitability.

[5]SEC requires pattern day traders to maintain a minimum equity of $25,000 in their margin account on a daily basis.

2.1.2 Algorithmic Trading in India

The setting we study focuses on Indian equity market (National Stock Exchange, NSE) and targets the interaction of individual investors and algorithmic trading. Algorithmic trading in the Indian equity markets was restricted to arbitrage related strategies prior to 2008. On April 3rd, 2008, the Securities & Exchange Board of India (SEBI) formally introduced algorithmic trading by allowing a Direct Market Access facility to institutional clients[6]. At the beginning, the adaption to algorithm trading is relatively slow and the algo-trading volume remains low. The situation changes significantly after the introduction of co-location facilities at NSE, allowing traders to place their servers near the exchange premises[7].

Though DMA facility was provided only to institutional clients, algorithmic trading gained popularity eventually among retail traders. The rising trend in retail participation is driven by the emergence of various Fintech and broker companies that provide application programming interfaces (API), including Zerodha, 5Paisa, Alice Blue, etc. The Indian retail algo-trading market we study displays similarities with respect to that in the United States. Retail investors can choose the platform that offers access to certain markets and are free to create their own strategies or select from pre-existing ones. While some platforms are free, others charge monthly fees.

Figure 2 provides an example of the steps investors take to implement a simple trading strategy based on moving averages at Streak, an algorithmic trading platform partnered with Zerodha. The process begins by selecting a few stocks from the Nifty 50, such as Adani Enterprises Ltd, with a one-hour candle interval, and deciding to trade 100 shares. The investor then designs the entry and exit strategy, specifying an entry position when the closing stock price crosses the moving average from below, and exiting at a stop loss or target profit of 5%. After running a backtest on applicable historical data, the investor can deploy the algorithm on the trading platform.Streak bots monitor stock movements to

[6]Foreign institutional investors were allowed to use DMA facility through nominated managers from February, 2009.

[7]After the availability of co-location service, there is a substantial increase in algorithmic trading volume with latency dropping from 10–30 ms to 2–6 ms.

generate alerts when a buy or sell signal occurs. With fully automated algorithms, they can even automate the order placement entirely.

2.2 Data

2.2.1 Stock Trading Data

Our primary data source is a unique and remarkably comprehensive database that contains the entire trading records on the National Stock Exchange of India (NSE) during the period of 2004 to 2020. NSE is the leading exchange in India and the world's 8th-largest stock exchange based on the market capitalization as of September 2021[8]. We can observe the anonymized Permanent Account Number (PAN) of the investor[9] and the trading data corresponds to aggregation at the individual level. Therefore, our setting is free of the concern that a given individual investor may hold multiple accounts[10]. For each transaction, we can also observe the date of transaction, the ticker of the security, the number of shares purchased or sold, and the execution price. We require that all transactions are associated with stocks included in the Prowess Database (like CRSP in the U.S.) maintained by the Centre for Monitoring Indian Economy (CMIE). In addition, we retain only securities that are common shares of domestic stocks and exclude trading activities related to ETFs and foreign stocks. Since the data includes the complete trading records, we are able to reconstruct the portfolio of stocks held by each investor on a daily bias. For each retail investor, we further obtain their demographics information, e.g., age and gender, which is crucial to identify the composition of retail algo traders and examine the link with trading preferences and patterns.

The initial sample includes equity trading transactions for 19 million unique investors across the country. Panel A of Figure A1 shows the geographical distribution of retail investors for all districts in India. Not surprisingly, the country's economic centers, such as

[8]https://www.statista.com/
[9]PAN is the unique identifier issued to all taxpayers by the Income Tax Department of India.
[10]This differs from some of the previous literature that cannot combine trades made by the same trader from different accounts.

the state of Maharashtra (where the National Stock Exchange (NSE) is located) and Tamil Nadu, host the most number of investors. Panel B of Figure A1 plots the evolution of number of investors who trade in a particular year as well as the total turnover by year since 2004.

[Insert Figure A1]

Most importantly for us to identify algorithmic trade, NSE requires all algorithmic orders emanating from the system to be tagged with the unique identifiers starting from 2012. This provides a rare opportunity for researchers to observe, for each trade, whether it's originating from an algorithm or not. Since the rule applies for all traded securities and traders, we can have a complete tracking of all investors' algorithmic trading, including entry, exit, as well as their daily trading activities. In the analysis, we limit our sample to the period between 2012 and 2019 and drop the year 2020 to exclude any potential confounding factors stemming from COVID-19. As the trader code enables us to distinguish retail investors from institutional traders (i.e., corporations, investment companies), we focus on trades by retail investors and explore how they interact with and use algorithmic trading in their investment activities.

Figure 1 shows the increasingly importance of algorithm in the India market. The aggregate trading value has greatly increased (Panel A) and the shares of algorithmic trading value among all trades on NSE rises significantly from 30% in 2012 to more than 50% in 2019 (Panel B). The pattern is similar on the retail front as Panel C suggests a steady increasing trend in the share of retail algorithmic trading value among all trades made by retail investors from 6% in 2012 to around 8.5% in 2019.

[Insert Figure 1]

2.2.2 Descriptive Statistics

Table 1 tabulates the mean, median, standard deviation, and quantile distribution for demographics and key variables that describe investor trading behavior. Panel A includes all traders that place at least one trade during the selected sample period (2012–2019), while Panel B presents descriptive statistics of algo investors in our analysis. The variables are winsorized at the 1st and 99th percentiles. All variables have a reasonable distribution in our sample. Comparing Panels A and B, it is apparent that algo trader is remarkably different from the entire population. Algo trader is younger (the average ages for algo traders and all traders are 31.72 and 33.29, respectively) and more likely to be male (only 17% of algo traders are female while the ratio is 24% for all traders) than the average investor in India. Algo traders also trade more actively (placing 150 trades on average per year) than other investors (44 trades on average per year for all traders). The difference is larger when we consider the median: the median number of trades for algo and all investors is 75 and 7, respectively. Moreover, algo investors trade larger amounts of money and in a wider selection of securities.

Figure 3 displays the geographical distribution of algo traders based on (1) total algo trading value and (2) number of algo investors in the sample. We can see that they are widely dispersed across districts in India. As expected, the retail algo trading is relatively concentrated at surroundings of mega cities (like Mumbai, Bangalore, and Kolkata).

[Insert Table 1, Figure 3]

2.2.3 Measuring Trading Performance

We are primarily interested in how investment performance (of both algo and non-algo trades) is correlated with investor's adoption of algorithmic trading. Nevertheless, performance measurement is a challenging task for at least two reasons. First, the holding period following each trade is essentially unobservable to researchers. Second, the holding period may vary from case to case and comparing the performance of trade with different holding periods generates a new problem. We take a straightforward approach by forcing the length of holding period for the trade (by investor i for stock j at day t) to be truncated at the cutoff, h, and calculate the return earned by each trade in the following h trading days, $\text{Return}_{i,j,t}$, as follows:

$$\text{Return}_{i,j,t} = \frac{\text{Closing Price}_{j,t+h}}{\text{Execution Price}_{i,j,t}} - 1 \tag{1}$$

where Execution $\text{Price}_{i,j,t}$ denotes the transaction price at day t and Closing $\text{Price}_{j,t+h}$ denotes the stock's closing price adjusted for splits and dividend at day $t + h$ ($h = 5, 10, 20, 30$). We choose the length of holding period, h, to be 5 or 10 in the main analysis because the median gap between two trades in the data is 9 trading days. Our approach aims to capture the short-term signals that the investor may have received after each trade. The findings remain unchanged if we use alternative time windows of 20- and 30-day holding period. For robustness check, we report the tests of longer horizons in the appendix.

3 Trends in Retail Algorithmic Trading

In this section, we present new stylized facts about retail algorithmic trading based on the administrative data. Specifically, we first show how the relative behavior of algo trading among different basic investor groups evolves over time and make comparisons across stock types and during different market circumstances.

We start with a discussion of the investor group classification and present some basic empirical regularities about how algo trading varies across different investor groups. Theory of life-cycle portfolio choice emphasizes the age-profiles of household stock market participation and risky portfolio share. For example, with a simple life-cycle model, Cocco, Gomes and Maenhout (2005) predicts households with borrowing constraints and undiversifiable income risk should lower risky portfolio share as they age. Following their predictions, we consider three basic investor groups[11] categorized based on age (as of 2012, the beginning of our sample period): (1) young (below 35); (2) mid-aged (aging from 35 to 60); (3) old

[11]The choice of three age groups is to account for the possibility that algo trading participation might be a hump-shaped function of household age (Gomes and Smirnova, 2021).

(above 60) investors. In addition to investor's age, the data also provides information on gender, so we further refer to the female and male counterparts of the investor-age groups as "female young", "female mid-aged", etc. Ultimately, we have six categories of investors with different age-gender profiles.

We are interested in how algorithmic trade participation varies across investors and over time. For this purpose, we track and aggregate all algo trades (in terms of trading value and number of investors) across investors at year t and obtain the annual measure of algo trading value for investor group g, $AT_{g,t}$. Similarly, we quantify the annual participation by computing the total number of algo traders at each investor group. Then we construct the relative share of algo trading across investor group g at year t:

$$\text{Relative Share}_{g,t} = \frac{AT_{g,t}}{\sum_{k=1}^{6} AT_{k,t}} \tag{2}$$

General Trend — Given these definitions, we can examine the time-varying properties of the relative algo-trading activities by each investor group. Panel A of Figure 4 shows the relative share based on the algo trading value, while Panel B shows the composition of the number of algo investors. Both figures offer a visualization of distinct differences and their dynamics among the groups. While male investors remain the dominant players in Indian retail algorithmic trading market, the proportion of female trading value has been steadily increasing throughout the sample period from around 6 percent in 2012 to more than 10 percent in 2019. This trend may mirror the declining gender gap in financial inclusion and growing financial liberalization among females in India. Relative to mid-aged and old traders, the share of young investors has increased dramatically as reflected in both the trading value and investor numbers[12]. For example, the relative algo trading share by the youngster is lower than 20 percent in 2012 but it jumps by three times to over 60 percent by 2019. The trends hold for both gender groups and perhaps reflect their increased market

[12]Since we fix the benchmark year to compute the age of investor, the trend reflects the composition of different cohorts rather than the fact that investors become aged as they trade in the market.

accessibility due to financial innovation (i.e., the popularity of online discount broker).

[Insert Figure 4]

Across Stocks — The pattern in Figure 4 considers the dynamic changes in algo trading shares across investor groups and over all periods. However, the composition may be varying depending on stocks with certain characteristics. Indeed, prior literature has highlighted the gender difference in investment style and asset allocation. Therefore, we focus on two dimensions that are most frequently used to categorize stocks, large vs small and value vs growth, and examine the extent to which the observed composition of different investor groups may change among stocks with various attributes[13].

We consolidate the relative share plots of algorithmic trading value in Figure 5 (with Panel A for market capitalization and Panel B for market-to-book ratio) and delegate similar plots of shares of algo investor numbers to Figure A2 in the Appendix. Among investment in large stocks, we continue to observe similar composition across investor groups as in Figure 4, whereas relative share plot shows marked differences in algo trading across small-cap stocks. For example, female traders play a larger role in investing in small stocks relative to large stocks and the trading value rises to nearly 40 percent of total trading value in 2017. When we break down between value and growth stocks, this pattern is similar to that in the aggregate data.

[Insert Figure 5, Figure A2]

Different Circumstances — As we've showed above, the importance of different investor group in algo trading participation varies over time and across stocks. One natural follow-up question is how algo trading across investor group responds to changes in stock market conditions. In fact, both investor's preferences and beliefs are likely to be varying with the macroeconomic cycles and the argument has been confirmed by recent experimental

[13]Since the algo trading is strategy-based, the evidence presented here is more likely to indicate differences in investment strategy rather than preference for specific stocks.

and field studies (Greenwood and Shleifer, 2014; Cohn et al., 2015). For example, following the 2008 crisis, both qualitative and quantitative measures of risk aversion increased substantially (Guiso, Sapienza and Zingales, 2018). Therefore, we explore the sensitivity of investor group holdings to market cycles (i.e., boom and bust periods)[14].

To assess whether the market is in a down or up status, we follow Daniel and Moskowitz (2016) and define, for each month, boom market if the excess cumulative market return in the past 24 months is positive and bust market otherwise. Figure 6 shows the relative algorithmic trading share for each investor group during booming and bust markets. It highlights interesting patterns of how algo trading participation by different investors moves with market conditions. For example, young and male traders, who are expanding considerably in the retail algo trading market as shown in Figure 4, decrease their algo trading during bust periods. Algo trade participation by mid-aged and elder investors moves countercyclical with the market and plays a larger role in algorithm adoption during the crisis period.

[Insert Figure 6]

4 Algorithmic Trading, Performance and Trading Behaviors

In the second part of the analysis, we focus on exploring the effect of up-taking algorithm on investor's trading performance. When algorithmic trading tool is not available, stock investing involves a complicated set of tasks as investors have to decide over a large pool of securities to allocate their wealth among the chosen stocks. Moreover, they need to actively monitor the market during trading hours and make active adjustment to rebalance portfolios in response to price movement. To do so, investors will unavoidably apply heuristics in the decision making, which can lead to suboptimal investment outcome. The algo-trading presumably can help simplifies the process and automate the trade execution.

[14]The sample period 2012–2019 is not long enough for us to study the responses of algo trading participation to periods of economic expansion and contraction.

Since the strategy is pre-determined, it can also mitigate the incidence of behavioral biases for adopters and result in better investment performance.

4.1 Comparing Performance of AT vs Human Trades

The standard method to evaluate the benefits of adopting the Fintech-driven investment tool (i.e., robo-advisor) is to contrast the performance before and after the subscription. For instance, D'Acunto, Prabhala and Rossi (2019) studies the introduction of a wealth-management robo-advisor by an India brokerage firm and reports a positive change in market-adjusted trade performance after using the portfolio optimizer. Since the adoption decision is mostly endogenous, any documented effect on performance might be driven instead by changes in investor's ability. Such self-selection bias is possible if any event (unobservable to econometricians) improves investor's investment skills and at the same leads to higher Fintech adoption.

The setting we examine provides a unique chance for us to construct valid counterfactuals to mitigate such concern. Specifically, the data allows for tracking the investment performance of both algo and non-algo trades by the same investor for the same period, thus the improvement in trading performance after adoption is unlikely contaminated by unobserved confounding factors. We consider three dimensions, return, risk and Sharpe ratio, to measure the performance of these two types of trading methods. We calculate the returns earned by each trade in the next 5 or 10 trading days following each trade given the median (mean) gap between two trades in our data is 9 (2) trading days. Our findings remain unchanged if we use alternative time windows of 20- and 30-day holding period[15].

All measures are first constructed for each trade and then aggregated using the trading value as the weight at a given month. For example, the monthly return of algo/non-algo trades is the value-weighted average of market-adjusted return (5 or 10-day) of all algo/non-algo trades submitted by investor i in month t. To measure the risk associated with algo/non-

[15] We provide details of the robustness check in Figure A3 of the appendix.

algo trades, for each trade submitted in month t, we first calculate the standard deviation of daily return over the next 5 or 10 trading days and compute the value-weighted average across all trades of the same type in a given month. Sharpe ratio is defined similarly: we first obtain the trade-level Sharpe ratio as the ratio of market-adjusted return to standard deviation of daily return for each algo/non-algo trade at month t and then take the value-weighted average across all trades of the same type in this month.

Figure 7 shows the dynamics of these performance measures for algo vs non-algo trades over the period of twelve months before and after the adoption. Time 0 indicates the month when the investor starts to integrate algorithm in her trading. We find that algorithmic trading is associated with better performance as measured by higher market-adjusted return and Sharpe ratio. In both cases the trading outcome dominates that of non-algo trades in every month after the adoption. Meanwhile, algo trading is beneficial to investors by lowering the investment risk as the volatility of return of algo trades is lower than that of non-algo trades by the same investor. Overall, these results portrait the positive influence of using algorithm in stock trading on both unadjusted and risk-adjusted basis.

Table 2 summarizes the various performance measures during the time windows before and after adopting algorithmic trading and formally tests the difference between them. Panel A and Panel B display the results using windows of 5- and 10-day, respectively. Consistent with what we observe in Figure 7 and Figure A3, both panels in Table 2 suggest a significant improvement of algo participation on investment outcomes. For example, Panel B shows that the 10-day market-adjusted return observed in the twelve months prior to and after algo adoption is -0.30% and -0.26%, whereas the return from algo trades in Column (3) is -0.15%. The extent of performance improvement from adapting to algorithmic trading is economically large: the gaps in 10-day market-adjusted return are 0.15% and 0.11%, which amounts to an annual return of 3.75% and 2.75%.

At the same time, the improved performance is not accompanied with higher exposure to risk as the volatility of algo trade is significantly lower that of non-algo trades both before and after the adoption. Thus the investment performance improvement, measured by Sharpe ratio, is achieved through both high return and lower volatility. The results are generally consistent with findings in prior literature that Fintech-oriented investment tools (i.e., robo-advisory) are beneficial to investors and perform better than their own portfolios.

[Insert Figure 7, Figure A3, Table 2]

4.2 Performance Improvement and Investor Characteristics

Given that algorithmic trading aims to improve performance by automating trades and eliminating bias and cognitive limitation of human traders, the superior performance may vary across investors. Therefore, the estimate of average effect in Table 2 may mask substantial heterogeneity based on investor's demographics or level of sophistication and behavioral biases *ex ante*. To quantify the performance improvement across investors, we compute the change in the market-adjusted Sharpe ratio of algo and non-algo trades by investor i during the twelve months after each user's first algo trade and regress this change on a list of investor's attributes in the cross section to assess how the effect of algorithmic trading is associated with investor characteristics.

As for investor-level variables, we first consider investor demographics, including age, gender and wealth. The information on investors' wealth is not directly available, so we calculate the average value of portfolio during the 12-month period prior to algo adoption as a proxy for their relative level of wealth. Second, we use two portfolio-level outcomes before adoption to proxy for portfolio diversification *ex ante*, that is, (i) number of stocks held in their portfolio, and (ii) the Herfindahl index (HHI) of the portfolio based on the value share of each stock in their portfolio.

The last set of investor features we study relates to a series of well-known behavioral bias attributed to retail investors. We focus on two types of behavioral bias in the literature: (1) salience (or availability) bias and (2) extrapolation bias (or trend chasing). Salience bias refers to investors' propensity to buy the best-performing stocks that are attention-grabbing

(Bordalo, Gennaioli and Shleifer, 2012). We first calculate the percentage of investor i's purchase of attention-grabbing securities among all her trading in month t and compute the average during the 12-month period prior to her algo adoption. Extrapolation bias measures the extent to which investors purchase stocks after a sequence of positive return, expecting the superior performance to continue afterwards (Cassella and Gulen, 2018). Similarly, we compute the percentage of the investor's purchase of momentum stocks among all her trading and calculate the average during the 12-month period prior to the adoption.

Table 3 reports the results that link the algo-trade-induced performance improvement with investor characteristics. On average, the benefits accrue more to mid-aged and elderly and female investors. The coefficient estimate on investor's wealth is negative, indicating that algo trading leads to more performance improvement to less wealthy investors who are relatively less financial literate, experienced, or have limited access to information source or financial advisory services. Columns (2) and (6) show that the impact on performance by algo trading is dependent on the extent of their *ex ante* diversification. We find that the performance improvement is even larger for investors who hold fewer stocks or more concentrated portfolio. Moving to the heterogenous effect across different levels of bias, we show evidence consistent with the notion that algo trading improves investment performance via mitigating investor behavioral bias. The positive and significant coefficients in Columns (3)-(4) and (7)-(8) indicate that investors who display higher level of investment biases benefit more after adapting to algo trading.

[Insert Table 3]

4.3 Source of Performance Improvement: Trading Responsiveness

It is widely recognized that retail investors often face cognitive limitations when it comes to reacting to new information in a timely manner. With limited time and resources to monitor the markets continuously, they often find themselves trailing behind institutional investors to capitalize on market inefficiencies (Barberis, 2018)[16]. In this context, algorithmic trading becomes a game-changer. With its cutting-edge technologies, it accelerates the process of information acquisition and automates the trading process. This allows investors to execute trades based on predefined rules and conditions, freeing them from the need for constant manual intervention so that investors can react swiftly to new information.

In this section, we explore the extent to which the adoption of algorithmic trading helps level the playing field for retail and institutional investors, particularly in terms of improving their trading responsiveness to new market information. We specifically examine whether algorithm trading adoption allows retail investors to gain faster access to earnings information after their release.

Following Bhattacharya, Cho and Kim (2018), we calculate the speed of trading response speed to earnings announcements as the total volume of shares of a firm traded by investor i during the three-day period centered on the announcement date (t=-1 to t=+1), divided by the total volume of shares of the same firm traded by the same investor over the seven-day period starting from the day before the announcement date (t=-1 to t=+5). This speed measure aims to capture what proportion of total announcement period trades occur within one day surrounding the release date. A higher value indicates that trades are closely clustered around earning announcements.

Figure 8 illustrates the average speed measure as previously defined. The left bar of the figure represents the average speed across retail investors before adoption, while the middle and right bars depict the average trading responsiveness in non-algo and algo trades after adoption. Two key findings emerge from this figure. First, it reveals a significant increase in trading responsiveness surrounding earnings announcement periods, particularly from non-algorithmic trades in the pre-adoption period to algorithmic trades in the post-adoption period. Second, this improvement exhibits a spillover effect to non-algorithmic trades post-adoption, as the trading responsiveness is also higher *ex post* than pre-period with the same

[16] A key example of this is the well-documented phenomenon of post–earnings-announcement drift.

trading technology (i.e. non-algo).

Regarding the economic magnitude of the improvement, the average value of the speed measure is about 49.02 percentage points for non-algorithmic trades before the adoption. Changes in the speed measure post-adoption are about 4.45 and 3.10 percentage points for both algorithmic and non-algorithmic trades, respectively. Thus, the adoption of algorithmic trading is associated with a proportionate increase in trading responsiveness by an economically significant 9.08% and 6.32%. We calculate a second measure of response speed in the same manner, using the the dollar volume instead of number of shares traded and obtain virtually identical results. In Table 4, we also formally test whether the speed measure changed systematically before and after adoption. We reject the null hypothesis that the differences between (1) non-algorithmic trades before and after, and (2) non-algorithmic trades before and algorithmic trades after, equal zero both statistically and economically at conventional levels.

Trading Responsiveness Ex ante: As an extension of our baseline estimate, we examine how improvements in trading responsiveness interact with the cross-sectional variation in reaction speed *ex ante* (as a proxy for cognitive limitation). We categorize all investors into four groups, from low to high levels of trading responsiveness, based on their average value prior to adoption. We then calculate the percentage of investors who improve (i.e., increase) their trading responsiveness after adoption. Again we compare the metric between (1) non-algorithmic trades before and after adoption, as well as (2) non-algorithmic trades before and algorithmic trades after adoption. Figure 9 presents these percentages using bar graphs which represent the number of shares traded and the dollar volume of shares, respectively.

The improvement in trading speed decreases monotonically with the *ex ante* trading responsiveness for both comparisons. Moving from left to right, the percentage of investors who improve their trading responsiveness drops from approximately 87% for those with a low reaction speed to only 13% for those with a high reaction speed. Furthermore, an interesting pattern emerges when we examine these two types of comparisons for each group of investors. The improvement in trading speed is higher among non-algorithmic trades for investors with the lowest reaction speed, whereas for investors who already react quickly to new information before adoption, the improvement is more prominent in their algorithmic trades.

[Insert Table 4, Figure 8, 9]

4.4 Source of Performance Improvement: Behavioral Bias

The second set of outcomes we examine to understand performance improvement relates to investor behavioral biases. Past literature has confirmed that these biases, either preference- or belief-based, can result in suboptimal investment decisions and eventually financial losses for individual investors (Barber and Odean, 2013). Algorithmic trading, designed to make data-driven decisions following pre-set rules, can potentially reduce the influence of human emotions and behavioral biases on investment choices. This is consistent with the results we present in the section 4.2, which demonstrate a more significant performance improvement among investors who previously exhibited more behavioral biases.

We focus on two types of behavioral biases established in the literature: (1) *the disposition effect*, where investors are more likely to realize gains than losses on their positions; and (2) *the rank effect*, where investors are more likely to sell the best-performing and worst-performing stocks in their portfolios compared with the other stocks. In practice, one can use a mixed strategy in their purchase and selling decision (e.g., buying decisions based on pre-defined rules while selling involves human intervention), and thus it's inaccurate to distinctly categorize portfolios into algorithmic trading and non-algorithmic trading. As both measures are portfolio-based and require specific cost of inventory definitions for their calculations, one caveat with the analysis is that we can't construct bias measures seperately for algo and non-algo trades *ex post* as we do in Section 4.3. We address this issue by using single-difference tests, where we compare trading biases within individuals before and after the adoption of algorithmic trading.

4.4.1 Algorithmic Trading and Disposition Effect

To assess the disposition effect in our sample, we follow Odean (1998) to calculate the difference between the proportion of gains realized (PGR) and the proportion of losses realized (PLR) for all investors before and after adopting algorithmic trading:

$$
\begin{aligned}
\text{PGR} &= \frac{\text{Realized Gains}}{\text{Realized Gains + Paper Gains}} \\
\text{PLR} &= \frac{\text{Realized Losses}}{\text{Realized Losses + Paper Losses}}
\end{aligned}
\tag{3}
$$

A large value of the difference between PGR and PLR indicates a stronger tendency to realize gains more than losses, implying a higher disposition effect. Each bar in Figure 10 displays the average disposition effect before (left) and after the adoption (right). We find that the disposition effect at the investor level decreases after incorporating algorithmic trading in their decisions. The bias doesn't completely disappear as the post-adoption level is still significantly different from zero.

Regarding the economic magnitude of the effect, the average difference between PGR and PLR is 26.14 percentage points. The change in this measure post-adoption compared to beforehand is 2.48 percentage points, which translates to about 9.5% of the average bias extent before the adoption. We also conduct a formal test in Panel B of Table 4 to determine whether the change is statistically significant and reject the null hypothesis that the within-investor change equals zero.

As in the procedure described in the previous section, we examine how bias reduction interacts with the average extent of bias before adoption. The results in Panel B of the Figure suggest that the decrease in the disposition effect is more prevalent among investors with a higher initial level of bias. Specifically, 76% and 61% of investors in the top two highest bias groups experienced a decrease, compared to only 56% and 25% in the two lowest bias groups.

4.4.2 Algorithmic Trading and Rank Effect

The second behavioral bias we consider is the rank effect, which captures the tendency of investors to sell the best- and worst-performing stocks in their portfolios, while ignoring stocks with intermediate performance. Following Hartzmark (2015), we first compute the proportion of best-, worst-, and middle-performing stocks investors sell. We then compute two measures of the rank effect as the difference between Best-Middle and Worst-Middle. A positive and statistically significant value signifies the presence of the rank effect, as observed among retail investors in U.S. and other countries.

$$\begin{aligned} \text{Best} &= \frac{\text{Best Sold}}{\text{Best Sold} + \text{Best not Sold}} \\ \text{Middle} &= \frac{\text{Middle Sold}}{\text{Middle Sold} + \text{Middle not Sold}} \\ \text{Worst} &= \frac{\text{Worst Sold}}{\text{Worst Sold} + \text{Worst not Sold}} \end{aligned} \tag{4}$$

In the left and right panels of Figure 11, we present the average difference of *Best-Middle* and *Worst-Middle* before (left bar) and after (right bar) adopting the algorithm. Consistent with previous findings, we observe the rank effect among Indian retail investors as their tendency to sell the best-performing stocks is significantly higher than for other stocks. However, after incorporating an algorithm into their trading decisions, the extent of this bias reduces. In terms of the magnitude, the share of best-performing stocks sold on average is about 13.24 percent points, whereas the size of the change after using algo-trading compared with before is about 4.23 percent points. The extent of reduction in rank effect translates into about 32%, which is substantially higher than the effect on disposition effect we show in Section 4.4.1 and statistically significant. The reduction in bias increases monotonically with the level of initial bias as shown in Panel B of Figure 11. The percentage of investors for whom the rank effect decreases is 78% for those with the highest level of bias and 24% for those with the lowest.

In contrast, we find little signs of the Worst-Middle rank effect within our sample as the investors have a lower tendency to sell their poorest performing stocks than their mid-performing ones. Collectively, the findings suggest that adopting algorithmic trading can improve investor performance by enhancing trading speed and overcoming behavioral biases.

[Insert Figure 10, 11]

5 Learning by Algorithmic Trading

Given the heterogeneity of algorithmic trading performance in Section 3, it is natural to ask whether investors learn from trading using algorithms over time. This section conducts a series of additional analyses to shed light on the potential learning process underlying the retail algorithmic trading.

5.1 Conceptual Framework

To guide the empirical investigation of the learning process in retail algorithmic trading, we formulate a simple framework in which investors are unsure about their abilities and learn as they make algorithmic trading. Consider the case of a retail investor who decides to adopt algorithms in her trading. She could devote time and efforts to grasping core technical skills (i.e., programing and statistical modeling) and become skilled in developing profitable trading strategies on her own. Alternatively, she may consider the methods available at the algo platform or consult for the advice of brokers, neighbors, and friends. Either way, the investor must learn to identify which of the various options deserves most attention and determine how much weight to allocate to each possible strategy. Such skills can only be improved as investors make actual algo trading and learn from the investment outcome.

Consider an individual investor who maximizes her utility over terminal wealth at T. The investor can decide whether to adopt algo-trade (and by how much) any time prior to the terminal date, $t = 1,\ 2,\ 3\ \cdots,\ T-1$. The investors will receive signals (through observing the trading outcome) about their "true" ability as they adapt to making more algo trades. If the signal is positive (with probability p), when her investment gains in value, the trader will infer positive news about her skill and choose to invest more in subsequent trades. If the signal is negative (with probability $1 - p$), in which the initial investment incurs losses, she will infer negative information about her skills and decides to trade less or quit.

The signal is only observable as the investor actually makes algo trades. The investor has a prior belief about p and uses Bayes' rule to update her beliefs after receiving new information on algo trading outcome. Assuming the investor's prior beliefs about p are normally distributed with mean p_0 and variance σ_0^2. The investor observes T independent signals about p, $s_t = p + \epsilon_t$, where ϵ_t is normal with zero mean and known variance σ_2. The individual's posterior beliefs based on Bayes' rule are normally distributed with mean $\hat{p}_T$ and variance $\hat{\sigma}_T^2$, where

$$\begin{aligned}\hat{p}_T &= p_0 \frac{\frac{1}{\sigma_0^2}}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}} + \bar{s} \frac{\frac{T}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}} \\ \hat{\sigma}_T^2 &= \frac{1}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}}\end{aligned} \tag{5}$$

and $\bar{s} = \frac{1}{T} \sum_{t=1}^{T} s_t$ is the average signal value. The posterior mean $\hat{p}_T$ is a precision-weighted average of the prior mean and average signal, while the posterior variance does not depend on the realization of the signals. Instead, the uncertainty about p, denoted by the variance $\hat{\sigma}_T^2$, decreases as the number of signals T increases.

We can derive a few empirical implications that can be summarized as follows. First, investors who are uncertain about their ability start to trade small amounts using algorithm during the first few trials and infer their own ability by observing the trading performance. Second, they will adjust gradually to the optimal trade sizes accordingly by increasing their trading size (decreasing trading size or exiting) in response to successful (unsuccessful) algo trades. Third, as uncertainty about p decreases with the number of signals T, the impact of signal weakens in the later stage of investor's algo trading experience.

5.2 Survival Rate of Retail Algorithmic Trading

In this section we start with showing empirical evidence on attrition and examine the rate at which individual investors, who are doing algorithmic trading in our sample, quit over time. As the emergence of retail algorithmic trading occurs in 2012, for each retail individual investor we are able to observe their algorithmic trading history and track the exact entry and exit records. Specifically, we can identify the first trade when an individual begins algorithmic trading as their entry trade. The corresponding exit trade is defined when we observe her placing no further algorithmic trades for the next 12 months. Due to this requirement, we restrict our analysis to investors who begin algorithmic trading before January 2019, for whom we are confident in providing more reliable estimates of true exit trades.

Figure 12 presents a plot of the Kaplan-Meier survival function for algorithmic trading status. In fact, attrition is a substantial aspect for algo traders: approximately 20% of those traders who take the first trial in algorithmic trading fail to ever trade using algorithm again. The survival rate drops at a fast speed to only 50% at the 8th trade, after which the curve becomes relatively flatter. The pattern in Figure 12 provides the first piece of suggestive evidence that most of the retail algo traders attempt to learn about their ability by making several algo trades as the first few trials and quit when the signal observed in not encouraging.

[Insert Figure 12 about here]

5.3 Learning about One's Ability

In the next step, we set down to test the implication of the learning model directly and focus on the interaction between retail algo trading activity and past performance. As mentioned in Section 4.1, learning investors perceive that their skills, pertaining to algo trading, are positively related with actual profit/loss and thus will infer their ability accordingly. If the algo trading outcome is poor, they realize that relying on algorithms fails to improve investment performance and refrain from active algo trading.

To test whether an investor continues to make algo trade when her performance is good and ceases trading once she gets a few bad draws, we model the decision to quit algorithmic trading as a function of past performance and estimate the following Cox proportional hazard rate model:

$$q(x,t) = q_0(t)e^{\beta X_t} \tag{6}$$

where $q(x,t)$ gives the hazard function at time t for each investor conditional on the covariate vector X_t. $q_0(t)$ denotes the baseline hazard rate when all covariates equal zero and are identical across investors. The impact of differentiated outcomes on the hazard rate can be obtained without estimating $q_0(t)$.

The key time-varying covariate in X_t includes proxies for the investor's algo trading performance. Specifically, for n-th algo trade, we consider both (1) the future return following each trade and (2) the average performance of all past trades from 1 to $n-1$, namely Trade Average Return. We calculate the returns earned in the next 5 or 10 trading days[17] following each algo trade. To account for the fact that investors may be quitting due to capital constraint, we also include the logarithmic size of the n-th algo trade in the estimation. Year-month fixed effects are included to control for any market-level movement that may affect investor's decision to exit. Again, we track the algo-trading records of all retail investors from the first entry until quitting, which is defined as the last trade after which we observe no algo trading in the next 12 consecutive months. Such filtering excludes the case that investors re-enter the algo-trading market after a long break and thus offers a cleaner setting for us to explore

[17]Given that the median (mean) lapse between an investor's algo trades is 9 (2) days, the choice of performance measurement should be reasonable to mimic investor's perception of algo trading outcome.

the learning process for the novice investors.

The results are tabulated in Table 5. We find that past performance, either measured using the most recent trade (trade n) or all past trades (trade 1 to trade $n-1$), is significant determinant of the investor's decision to exit algo trade. For example, the corresponding hazard ratio in Column (1) indicates that every 1% increase in the future 5-day return of most recent trade lowers the investor's exit rate of algo trade by 0.81%. Turning to all past trades, the investment performance also exhibits a negative impact on hazard rate: 1% increase in the future 5-day return reduces the hazard rate by 0.67%. The slightly small magnitude reflects the fact that performance of most recent trade is relatively more salient to the investor (Bordalo, Gennaioli and Shleifer, 2012). The exit decision is also sensitive to trade size, but the impact on survival rate is relatively small.

To visualize the impact of past performance on algo trading exit, we construct 16 dummy variables indicating each of the 5-basis-point intervals: (–inf, –50 bps), (–50 bps, –45 bps), $\cdots$, (20 bps, inf). We then estimate the Cox proportional hazard rate model in Equation 6 where we treat the interval (–5 bps, 0) as the benchmark category and include the remaining 15 as covariates in our estimation. As control variables we include the logarithmic size of n-th algo trade and year-month fixed effects.

Figure A4 presents the plot of results from the estimation. The decision to quit algo trading is quite sensitive to past average performance, especially to the extreme negative returns. For example, when past average performance range is moving from just unprofitable, (–5 bps, 0), to the low range of losses (–25 bps, –20 bps), the hazard rate rises by 150 percentage points (from 1.00 to 2.50). In contrast, if we move by the same magnitude to the range of gains, (15 bps, 20 bps), the hazard ratio rises by just 60 percentage points (from 1.00 to 1.60).

Overall, the above tests suggest evidence consistent with learning model: traders do not initially know their own ability to do algo trading and will infer their abilities by observing the actual investment performance. They respond to the feedback quickly and cease their algo trading activities in response to discouraging performance.

[Insert Table 5, Figure A4]

5.4 Size of Retail Algorithmic Trade

To this point, we find evidence consistent with the learning model that poor performance is more likely to result in investor's decision to quit algo-trade. If retail investors indeed exhibit learning in algo trading, their trade sizes are also supposed to be sensitive to past performance. As predicted in Section 4.1, investors who are uncertain about their ability start to trade small amounts using algorithm during the first few trials and learn about their own ability by observing the trading performance. Algo trading becomes more attractive after success and less attractive after failure. With this in mind, we further assess the impact of past performance on size of algo trade by retail investors.

As the first step to evaluate whether this implication is manifested in the data, we plot the distribution of trade sizes among investors who make algo trade for the first time in Panel A, Figure 13. It suggests that the distribution of initial trade size is negatively skewed with the majority of observations concentrated at the lower end. For example, roughly 40% of initial trades are smaller than 10,000 rupees (approximately 120 USD). In Panel B of Figure 13, we study how the average trade size evolves as investors continue to trade beyond the initial trial. The size of each algo trade is normalized by the initial trade and thus indicates the relative change. The plot suggests that algo investors make a significantly bigger trade as she stays and makes more trades. For instance, compared to her first algo trade, the average investor who goes on to trade for the second and fifth times increase the trade size by 4% and 30% respectively. Interestingly, the relation is not linear and changes in algo trade size decelerate as the investor becomes more experienced over time. Note that the increasing algo trade size patterns are estimated only for "alive" algo investors and don't account for the endogenous attrition as unsuccessful investors chooses to quit at some points. Overall, the above tests suggest evidence consistent with predictions of learning model.

To examine directly how algo trade sizes are sensitive to past performance, we regress the log size of n-th algo trade on the performance of trade $n-1$. Because trade size may be correlated over time even in the absence of learning, we also include the size of trade $n-1$ (in logarithm) as an additional control:

$$\begin{aligned}\log(\text{Trade size})_{i,n,t_1} = \alpha_i + \gamma_n + \theta_{t_1} + \beta_1 \times \text{Performance}_{i,n-1,t_2} + \\ \beta_2 \times \log(\text{Trade size})_{i,n-1,t_2} + \epsilon_{i,n,t_1}\end{aligned} \tag{7}$$

where the investor, year-month and trade time fixed effects are denoted by α_i, γ_n and θ_{t_1}. The performance is measured by the returns earned in the next five or ten trading days following each algo trade. We exclude the last algo trade before their exit so that our estimate focuses on the impact of past performance on trade size changes.

Table 6 presents the results for the trade size regression and suggests that algo trading outcomes influence trade sizes significantly. Column (2) reports the impact of past performance as measured by 5-day return and shows that the investor increases the size of n-th trade by 8.37% when 5-day return in trade $n-1$ is higher by one percentage point. As indicated in Figure 12, retail investor's participation in algo trading changes significantly over time. To account for time-invariant unobserved factors at investor-level, we control for investor fixed effects in Columns (2) and (4). We find that the learning estimates are slightly larger when we account for the attrition effect of low-ability traders.

[Insert Figure 13, Table 6]

6 Conclusions

In this study, we explore detailed equity trading records for the universe of retail investors in the National Stock Exchange of Indian to provide new evidence on how individual investors are integrating algorithms into their trading activity. The upfront question is: who are adapting to algorithmic trading among retail investors? We address this question by decomposing the retail participation in algorithmic trading into different investor groups. We present a few interesting and stylized facts. First, male investors occupy a substantially large share in Indian retail algorithmic trading market, whereas the relative share of female algorithmic trading participation (in terms of market value) gradually increases from 2012 to 2019. Exploring the investor composition over market cycles, we find that young and male traders are most responsive to stock market condition and their usage of algorithms are expanding (contracting) considerably during booming (bust) periods.

Next, we show that adapting to algorithmic trading is associated with better investment performance as algo trades deliver higher market-adjusted return and Sharpe ratio relative to non-algo trades by the same investor at the same time. The performance improvement is not homogeneous across investors but more pronounced among investors who are less wealthy, hold less diversified portfolio or exhibit more behavioral bias *ex ante*. Exploring the source of the improved performance, we show that they are likely attributed to changes in two sets of investor behaviors: enhanced trading responsiveness to new market information and reduced behavioral biases.

Finally, we explore further to uncover the adaption process and test whether retail investors rationally learn about their ability by trading using algorithms over time. They are initially unsure about their skill and trade small amounts during the first few trials. In response to success (or failure), they adjust the optimal trade sizes accordingly by increasing (or decreasing) their trading size. The investor stops algo trading after realizing that her ability is insufficient.

The concept of algorithmic trading has been in existence since 1990s and exclusively used by investment funds and institutional traders. However, the rapid development of Fintech and big data analytics during the past decade have spurred exponential demand for algorithmic trading technology from the retail front. Our results contribute to enhancing the understanding of the algorithm usage in retail investment decisions. We examine a large-sample administrative data and confirm the effectiveness of algo-trading as an alternative investment tool that changes financial behaviors and outcomes of households. The finding may also have policy implications for regulator to formulate policies to ensure retail investors' suitability and protect their interest in algorithm trading participation.

References

Abis, Simona, and Laura Veldkamp. 2020. "The changing economics of knowledge production." *Available at SSRN 3570130*.

Agarwal, Sumit, Wenlan Qian, Bernard Y Yeung, and Xin Zou. 2019. "Mobile wallet and entrepreneurial growth." Vol. 109, 48–53.

Bai, Jennie, Thomas Philippon, and Alexi Savov. 2016. "Have financial markets become more informative?" *Journal of Financial Economics*, 122(3): 625–654.

Barber, Brad M, and Terrance Odean. 2013. "The behavior of individual investors." In *Handbook of the Economics of Finance*. Vol. 2, 1533–1570. Elsevier.

Barberis, Nicholas. 2018. "Psychology-based models of asset prices and trading volume." In *Handbook of behavioral economics: applications and foundations 1*. Vol. 1, 79–175. Elsevier.

Baron, Matthew, Jonathan Brogaard, Björn Hagströmer, and Andrei Kirilenko. 2019. "Risk and return in high-frequency trading." *Journal of Financial and Quantitative Analysis*, 54(3): 993–1024.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace. 2022. "Consumer-lending discrimination in the FinTech era." *Journal of Financial Economics*, 143(1): 30–56.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri. 2020. "On the rise of fintechs: Credit scoring using digital footprints." *The Review of Financial Studies*, 33(7): 2845–2897.

Bhattacharya, Nilabhra, Young Jun Cho, and Jae B Kim. 2018. "Leveling the playing field between large and small institutions: Evidence from the SEC's XBRL mandate." *The Accounting Review*, 93(5): 51–71.

Biais, Bruno, Thierry Foucault, and Sophie Moinas. 2015. "Equilibrium fast trading." *Journal of Financial Economics*, 116(2): 292–313.

Bianchi, Milo, and Marie Brière. 2021. "Augmenting Investment Decisions with Robo-Advice." *Université Paris-Dauphine Research Paper*, , (3751620).

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012. "Salience theory of choice under risk." *The Quarterly Journal of Economics*, 127(3): 1243–1285.

Brogaard, Jonathan, Björn Hagströmer, Lars Nordén, and Ryan Riordan. 2015. "Trading fast and slow: Colocation and liquidity." *The Review of Financial Studies*, 28(12): 3407–3443.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. "High-frequency trading and price discovery." *The Review of Financial Studies*, 27(8): 2267–2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2019. "Price discovery without trading: Evidence from limit orders." *The Journal of Finance*, 74(4): 1621–1658.

Budish, Eric, Peter Cramton, and John Shim. 2015. "The high-frequency trading arms race: Frequent batch auctions as a market design response." *The Quarterly Journal of Economics*, 130(4): 1547–1621.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolo, and Sergio Pastorello. 2020. "Artificial intelligence, algorithmic pricing, and collusion." *American Economic Review*, 110(10): 3267–97.

Campbell, John Y. 2006. "Household finance." *The Journal of Finance*, 61(4): 1553–1604.

Cao, Sean, Wei Jiang, Junbo L Wang, and Baozhong Yang. 2024. "From man vs. machine to man+ machine: The art and AI of stock analyses." *Columbia Business School Research Paper*.

Carlin, B, A Olafsson, and M Pagel. 2019. "Fintech and financial fitness in the information age." Working Paper.

Cassella, Stefano, and Huseyin Gulen. 2018. "Extrapolation bias and the predictability of stock returns by price-scaled variables." *The Review of Financial Studies*, 31(11): 4345–4397.

Chaboud, Alain P, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. 2014. "Rise of the machines: Algorithmic trading in the foreign exchange market." *The Journal of Finance*, 69(5): 2045–2084.

Chen, Luyang, Markus Pelger, and Jason Zhu. 2022. "Deep learning in asset pricing." *Management Science*.

Chen, Mark A, Qinxi Wu, and Baozhong Yang. 2019. "How valuable is FinTech innovation?" *The Review of Financial Studies*, 32(5): 2062–2106.

Choi, Hyun-Soo, and Roger Loh. 2021. "Physical frictions and digital banking adoption." *Available at SSRN 3333636*.

Cocco, Joao F, Francisco J Gomes, and Pascal J Maenhout. 2005. "Consumption and portfolio choice over the life cycle." *The Review of Financial Studies*, 18(2): 491–533.

Cohn, Alain, Jan Engelmann, Ernst Fehr, and Michel André Maréchal. 2015. "Evidence for countercyclical risk aversion: An experiment with financial professionals." *American Economic Review*, 105(2): 860–85.

Colliard, Jean-Edouard, Thierry Foucault, and Stefano Lovo. 2022. "Algorithmic Pricing and Liquidity in Securities Markets." *HEC Paris Research Paper*.

Conrad, Jennifer, Sunil Wahal, and Jin Xiang. 2015. "High-frequency quoting, trading, and the efficiency of prices." *Journal of Financial Economics*, 116(2): 271–291.

Cowgill, Bo, and Catherine E Tucker. 2022. "Economics, fairness and algorithmic bias." *Journal of Economic Perspectives*.

D'Acunto, Francesco, Alberto G Rossi, and Michael Weber. 2019. "Crowdsourcing financial information to change spending behavior."

Daniel, Kent, and Tobias J Moskowitz. 2016. "Momentum crashes." *Journal of Financial economics*, 122(2): 221–247.

Dietvorst, Berkeley J, Joseph P Simmons, and Cade Massey. 2018. "Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them." *Management Science*, 64(3): 1155–1170.

Dressel, Julia, and Hany Farid. 2018. "The accuracy, fairness, and limits of predicting recidivism." *Science advances*, 4(1): eaao5580.

D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G Rossi. 2019. "The promises and pitfalls of robo-advising." *The Review of Financial Studies*, 32(5): 1983–2020.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2022. "Predictably unequal? The effects of machine learning on credit markets." *The Journal of Finance*, 77(1): 5–47.

Gargano, Antonio, and Alberto G Rossi. 2022. "Goal Setting and Saving in the FinTech Era." *Journal of Finance*.

Gomes, Francisco, and Oksana Smirnova. 2021. "Stock Market Participation and Portfolio Shares Over the Life-Cycle." *Available at SSRN 3808350*.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of returns and expected returns." *The Review of Financial Studies*, 27(3): 714–746.

Grennan, Jillian, and Roni Michaely. 2020. "Artificial intelligence and the future of work: Evidence from analysts." Working paper, Duke University.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2018. "Time varying risk aversion." *Journal of Financial Economics*, 128(3): 403–421.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. "Empirical asset pricing via machine learning." *The Review of Financial Studies*, 33(5): 2223–2273.

Hagströmer, Björn, and Lars Nordén. 2013. "The diversity of high-frequency traders." *Journal of Financial Markets*, 16(4): 741–770.

Hartzmark, Samuel M. 2015. "The worst, the best, ignoring all the rest: The rank effect and trading behavior." *The Review of Financial Studies*, 28(4): 1024–1059.

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld. 2011. "Does algorithmic trading improve liquidity?" *The Journal of Finance*, 66(1): 1–33.

Hirshleifer, David. 2015. "Behavioral finance." *Annual Review of Financial Economics*, 7: 133–159.

Jansen, Mark, Hieu Nguyen, and Amin Shams. 2020. *Human vs. machine: Underwriting decisions in finance.* Ohio State University, Fisher College of Business, Charles A. Dice Center for Research in Financial Economics.

Odean, Terrance. 1998. "Are investors reluctant to realize their losses?" *The Journal of finance*, 53(5): 1775–1798.

O'hara, Maureen. 2015. "High frequency market microstructure." *Journal of Financial Economics*, 116(2): 257–270.

Pagliaro, Cynthia, Tarun Ramadorai Alberto Rossi Stephen Utkus, and Walther Ansgar. 2022. "Algorithm aversion: Theory and evidence from robo-advice." *Working Paper*.

Rambachan, Ashesh, Jon Kleinberg, Jens Ludwig, and Sendhil Mullainathan. 2020. "An economic perspective on algorithmic fairness." Vol. 110, 91–95.

Rossi, Alberto G, and Stephen P Utkus. 2024. "The Diversification and Welfare Effects of Robo-advising." *Journal of Financial Economics*.

Van Kervel, Vincent, and Albert J Menkveld. 2019. "High-frequency trading around large institutional orders." *The Journal of Finance*, 74(3): 1091–1137.

Weller, Brian M. 2018. "Does algorithmic trading reduce information acquisition?" *The Review of Financial Studies*, 31(6): 2184–2226.

Figure 1: **Trends in Algorithmic Trading**

Panel A of this figure plots the time series of value of algorithmic trading in India; Panel B plots the shares of algorithmic trading value among all trades at NSE; Panel C plots the shares of retail algorithmic trading value among all trades made by retail investors at NSE.

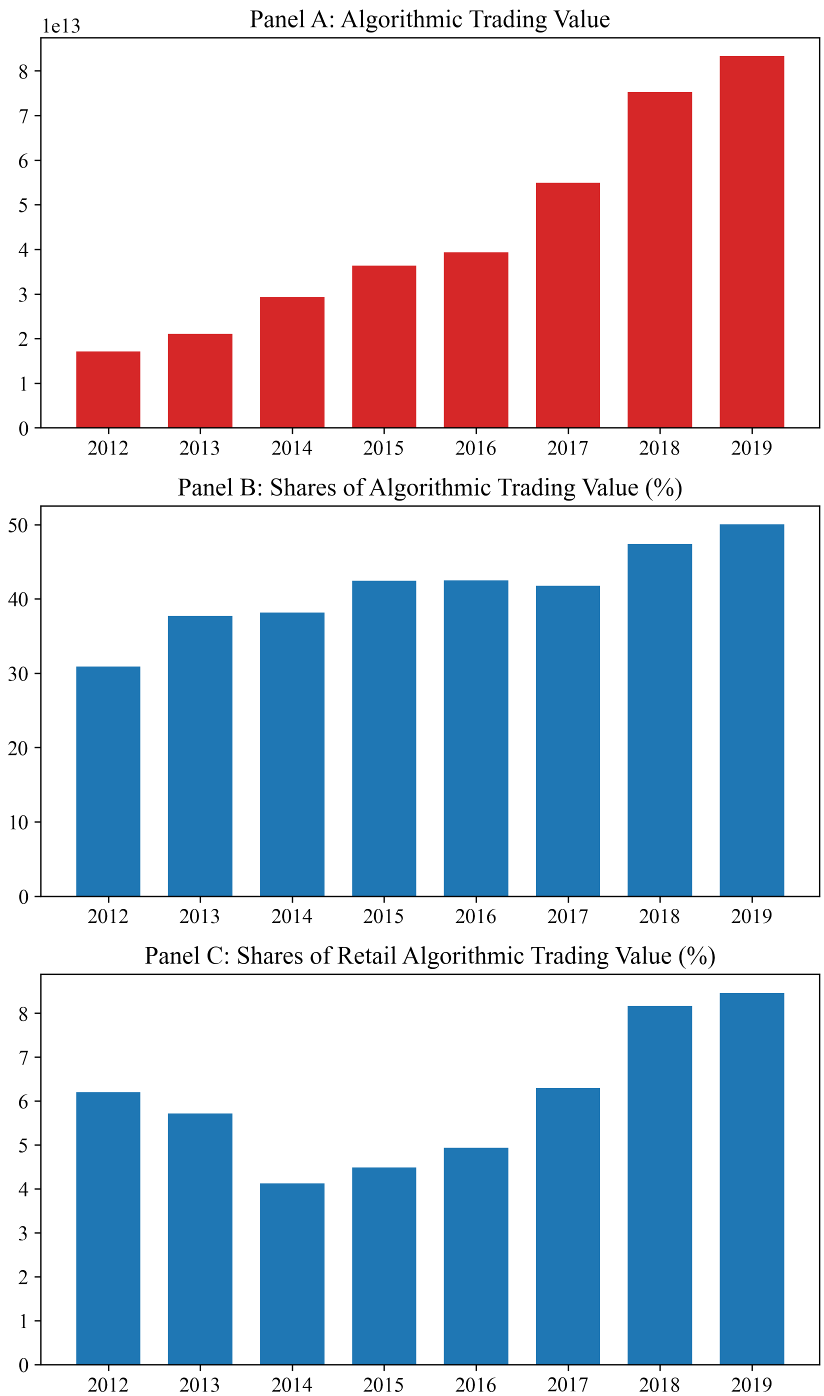


Figure 2: **Screen Display of the Algorithmic Trading Platform: An Illustrative Example**

This figure shows a screenshot of Streak, a popular algo trading platform partnered with Zerodha, as an illustrative example. Figure (a) - Figure (d) display the steps investors take to implement a simple trading strategy based on moving averages. In Figure (a), the investor starts by adding a few stocks from the banking sector, with a one-hour candle interval, and decides to trade with 100 shares; then in Figure (b) the investor takes an entry position when the closing stock price crosses the moving average from below and exit at a stop loss of 5% or a target profit of 5% in Figure (c). Lastly, in Figure (d), the investor names the strategy "MA" and runs a back test on relevant historical data. https://zerodha.com/z-connect/streak/introducing-streak-algo-trade-without-coding

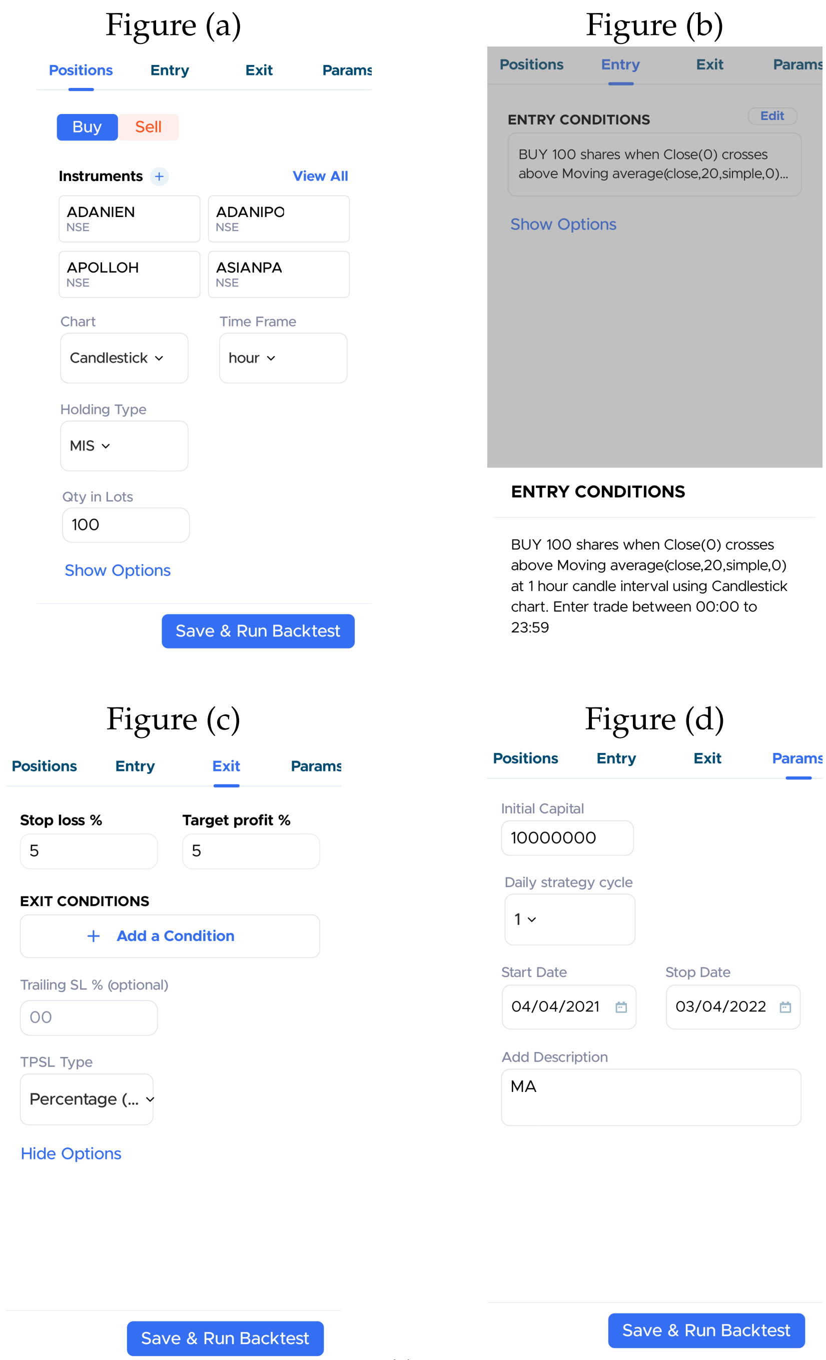


Figure 3: **Geography of Retail Algo Traders**

The plot figure plots the geographical distribution of retail algorithmic investors across districts who trade at NSE from 2012 to 2019. The left and right panel presents the statistics based on the total algo trading value and number of algo investors in our data.

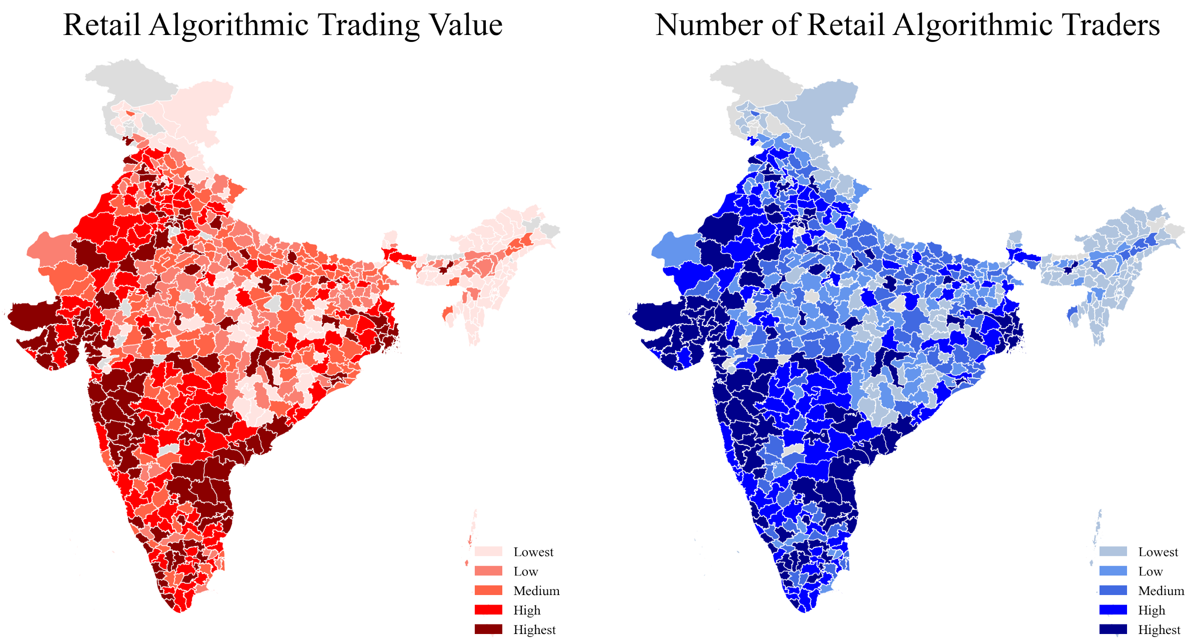


Figure 4: **Trends in Composition of Retail Algorithmic Traders**

This figure presents the time series of algorithmic trading share by six different investor groups. We classify investors, based on their age as of the beginning of sample period (2012), into (1) young (below 35); (2) mid-aged (aging from 35 to 60); (3) old (above 60) investors. Then we refer to the gender counterparts of these three categories and consider six investor groups in total. Panel A and B plot the share based on (1) value of algo trade and (2) number of algo investors respectively.

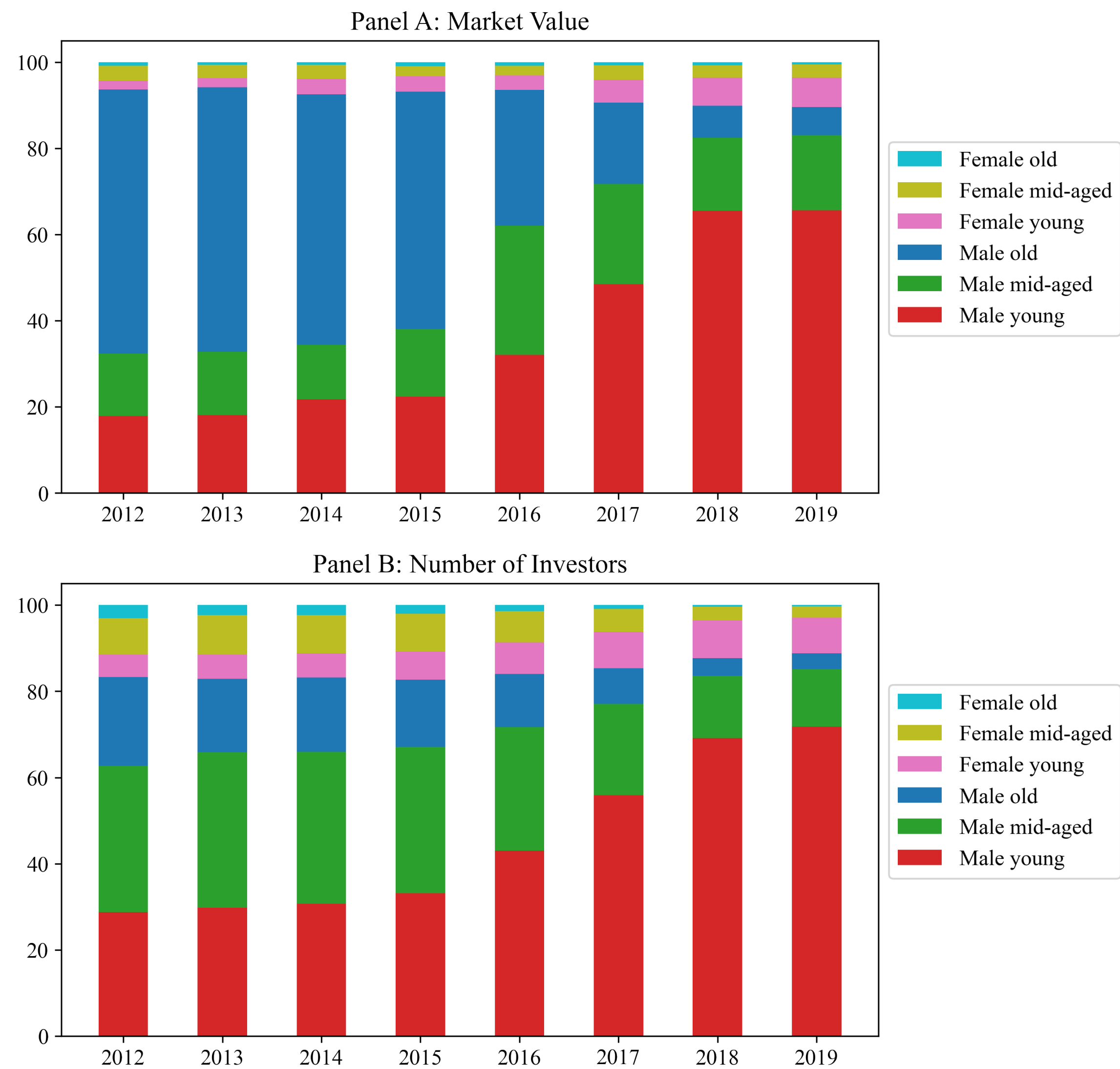


Figure 5: **Composition of Retail Algorithmic Traders: Across Stocks**

This figure presents algorithmic trading share for each investor group among (1) stocks with different market capitalization in Panel A and (2) value and growth stocks in Panel B. Stocks are assigned to groups (small and large stocks) based on their market cap at the end of each year and the breakpoints are 50th percentiles of cross-section distribution. Similarly, value (vs growth) stock is defined based on their price-to-book value at the end of each year and the breakpoints are 50th percentiles of cross-section distribution. The algo trading share is calculated based on the market value of algo trade.

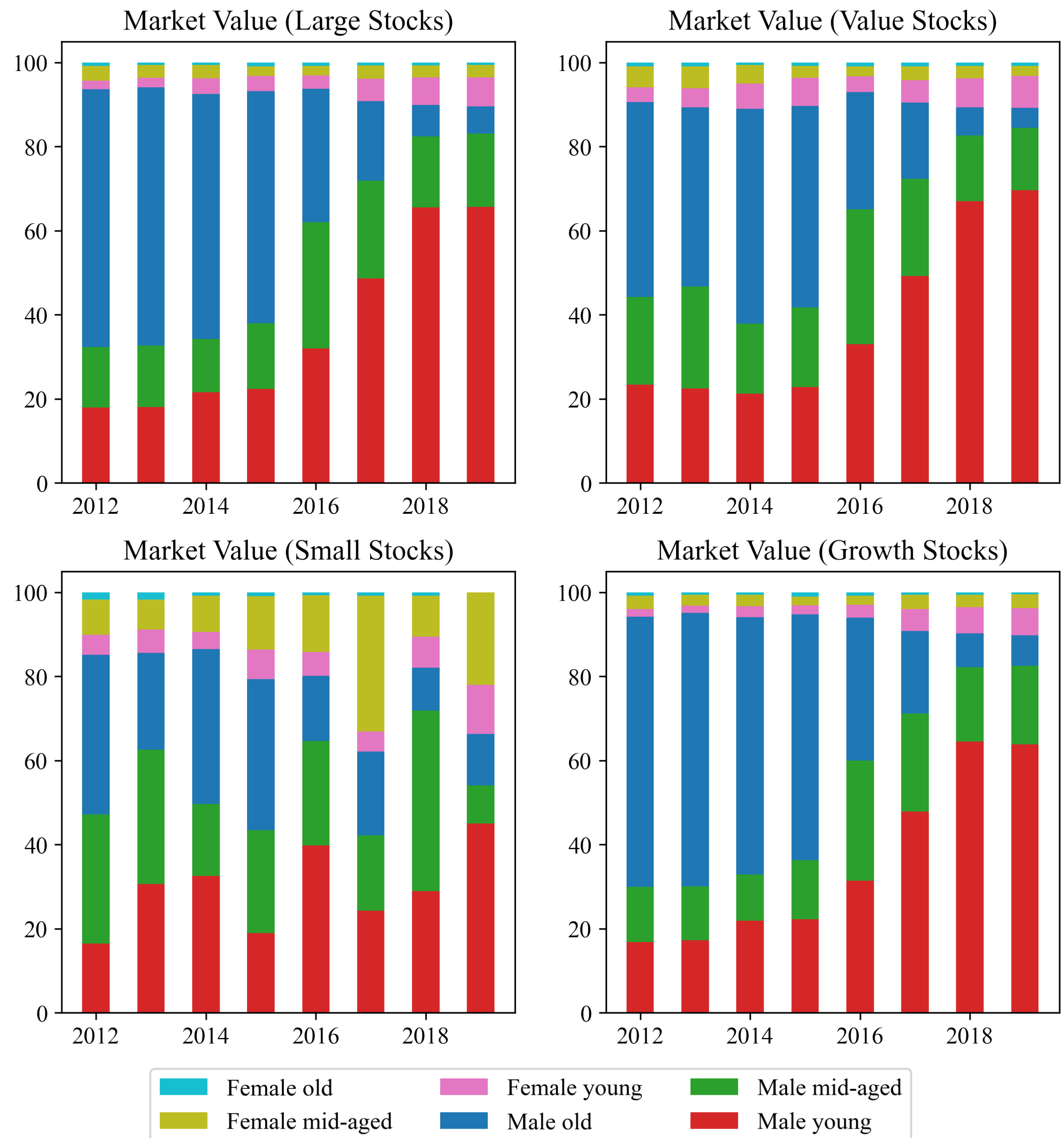


Figure 6: **Composition of Retail Algo Traders: Market Condition**

This figure presents algorithmic trading share for each investor group during booming and bust markets. We follow Daniel and Moskowitz (2016) and define, for each month, boom market if the excess cumulative market return in the past 24 month is positive and bust market otherwise. The algo trading shares based on the market value of algo trade and number of algo investors are shown respectively.

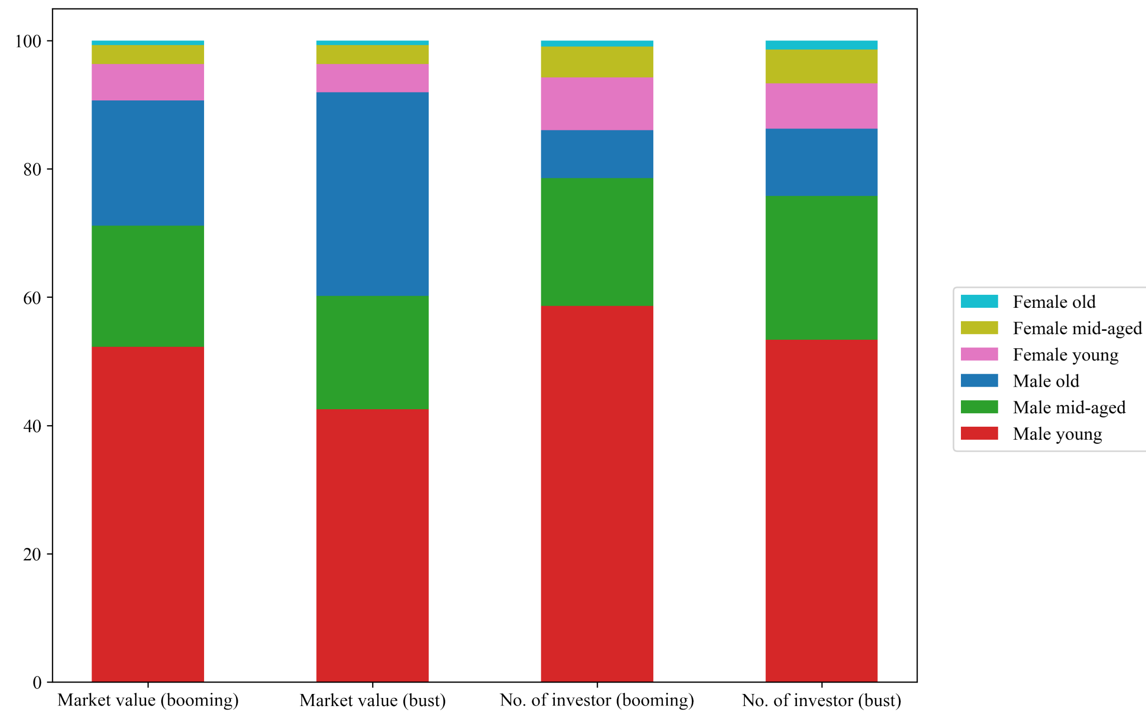


Figure 7: **Performance of Retail Algo vs Human Trading**

This figure shows the investment performance for algo vs non-algo by retail investors around the algo adoption time. Performance measures include (1) Return: the value-weighted average market-adjusted return (5- or 10-day) of all trades in month t; (2) Risk (volatility): for each trade submitted in month t, we first calculate the standard deviation of daily return over the next 5 or 10 trading days and compute the value-weighted average across all trades in a given month; (3) Sharp ratio: for each trade in month t, we obtain the trade-level Sharpe ratio as the ratio of 5/10-day market-adjusted return to standard deviation of daily return over the next 5/10 trading days and then take the value-weighted average across all trades in a given month.

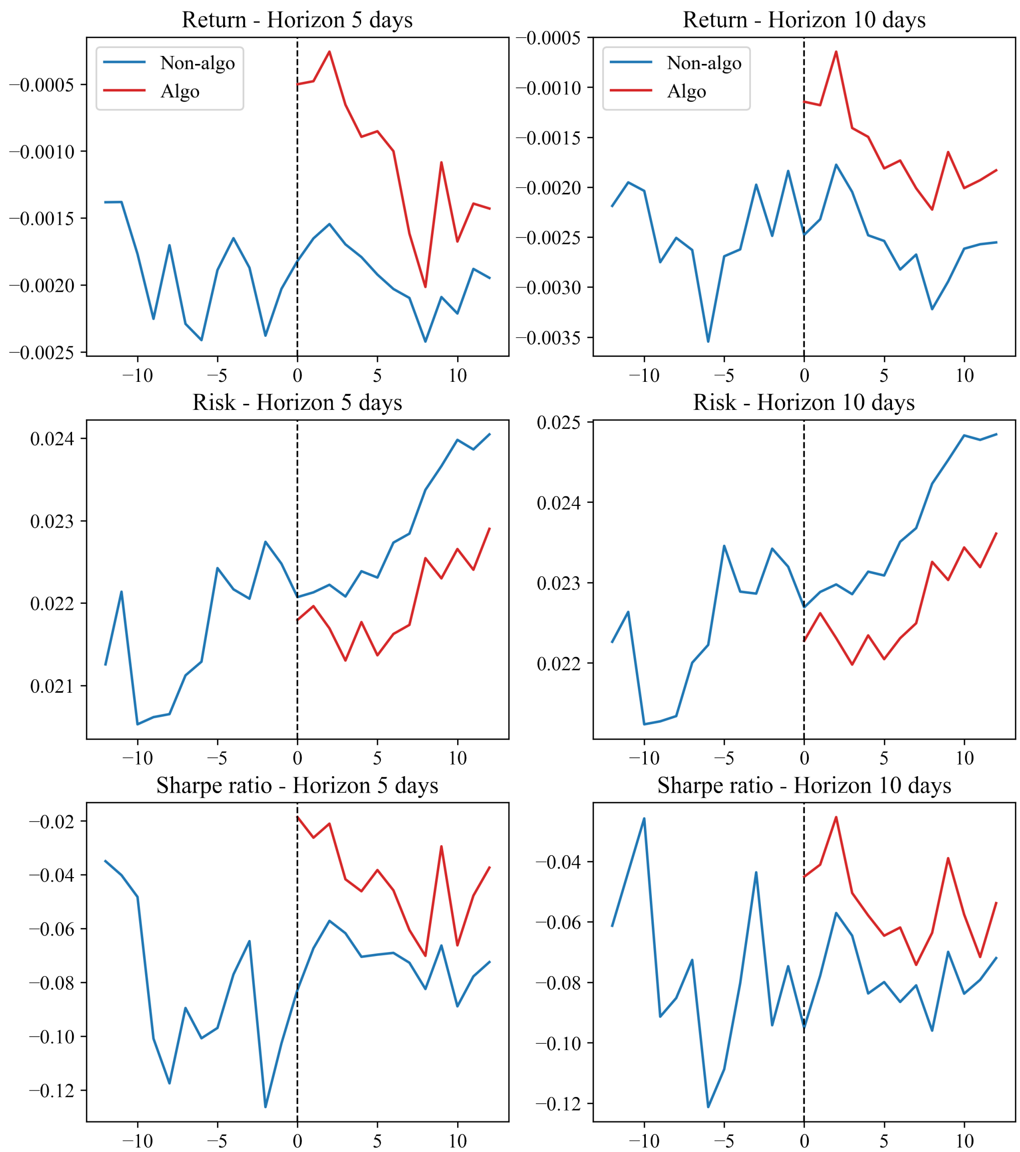


Figure 8: **Responsiveness of Retail Algo vs Human Trading**

The figure illustrates the change in reaction speed to new market information(i.e)earnings announcements) for both algorithmic and non-algorithmic trades around the time of algorithmic adoption. Each bar in the figure denotes the average speed across retail investors before adoption, and for non-algorithmic and algorithmic trades after adoption. The speed measures in the left and right figures are computed using the number of shares traded and the dollar volume of shares, respectively.

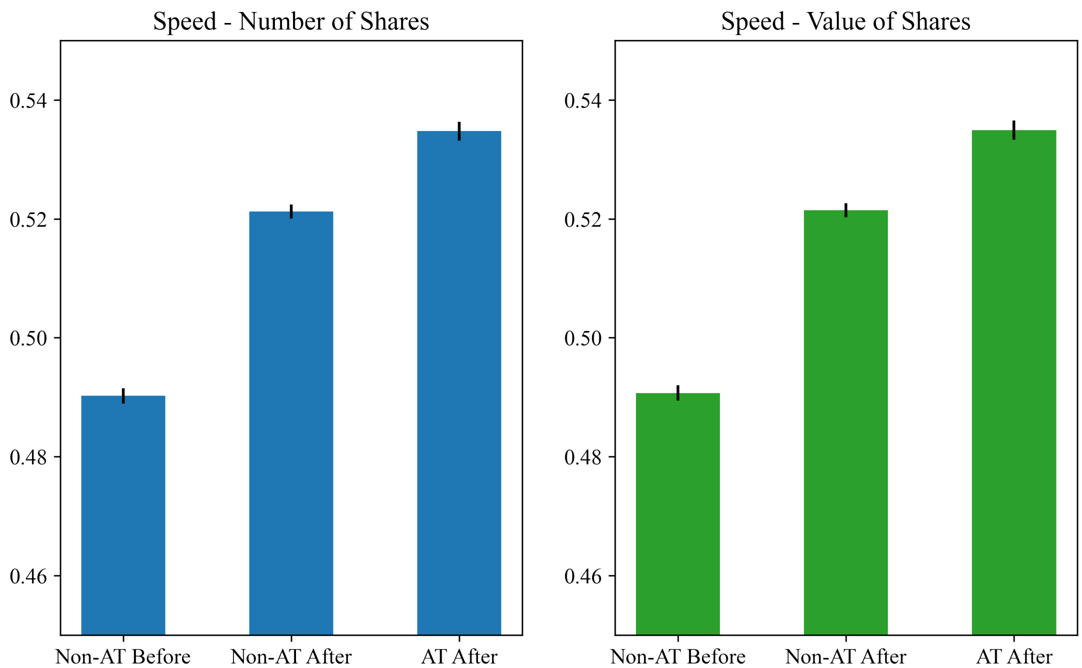


Figure 9: **Cross-Section of Change in Trading Responsiveness**

The figure illustrates the cross-section of change in reaction speed to new market information(i.e earnings announcements) for both algorithmic and non-algorithmic trades around the time of algorithmic adoption, conditioning on the level of trading responsiveness ex ante. We categorize all investors into four groups, from low to high levels of trading responsiveness, based on their average value ex ante. Each bar reports the percentage of investors who improved (i.e., increased) their trading responsiveness after algorithmic adoption. This is done by comparing the metric between (1) non-algorithmic trades before and after adoption, as well as (2) non-algorithmic trading before and algorithmic trades after adoption. The speed measures in the top and bottom figures are computed using the number of shares traded and the dollar volume of shares, respectively.

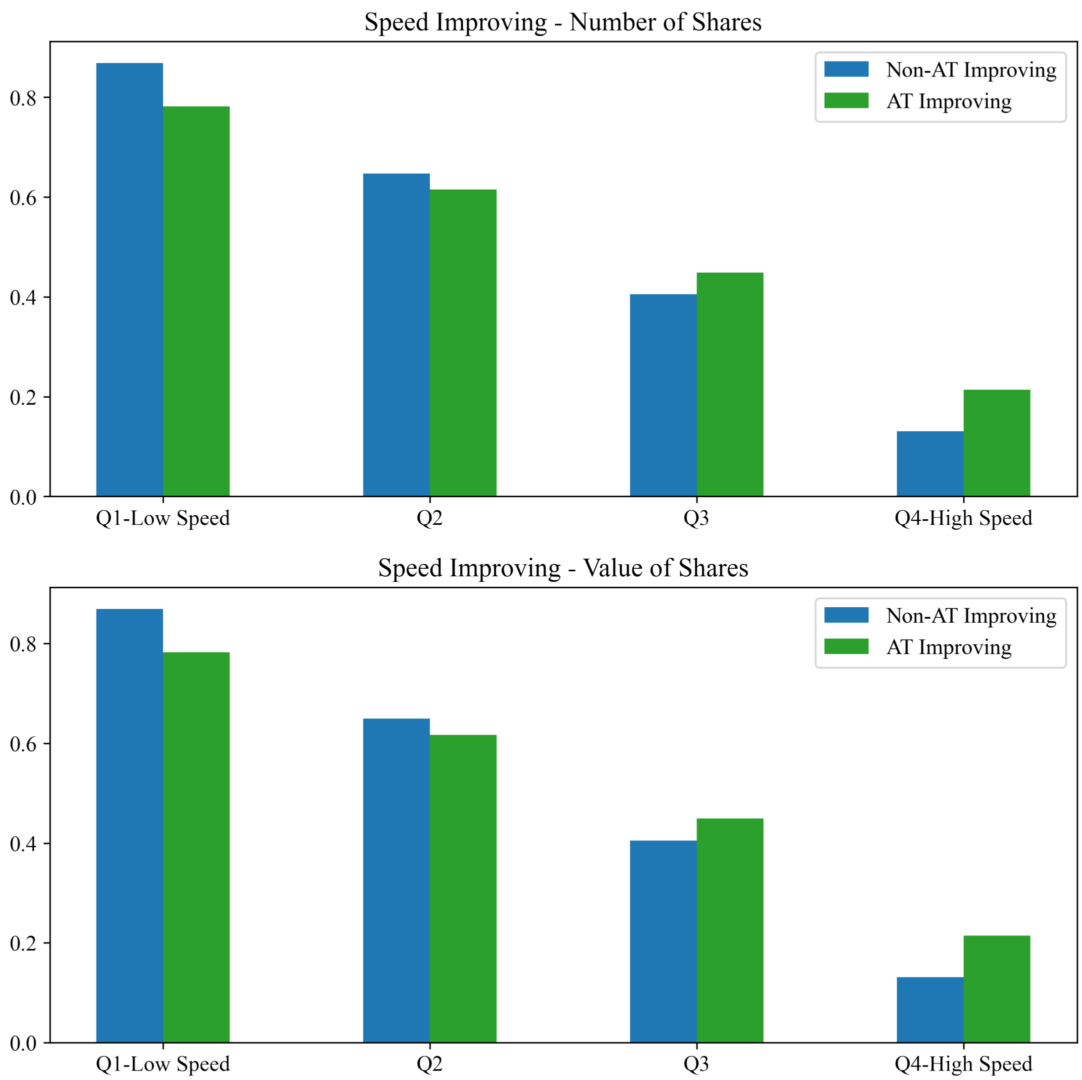
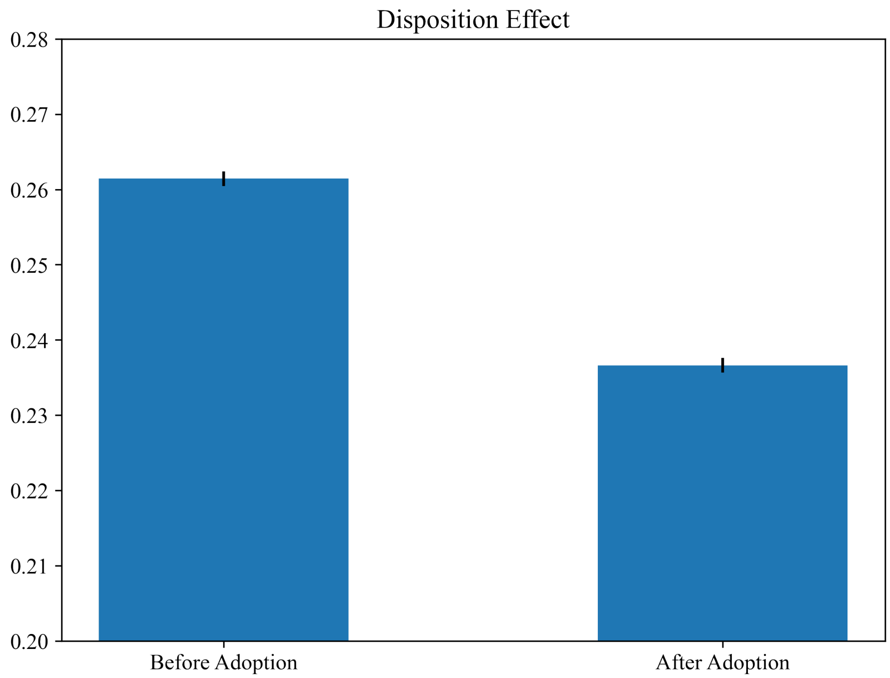


Figure 10: **Retail Algo vs Human Trading: Disposition Effect**

Panel A of the figure illustrates the change in disposition effect around the time of algorithmic adoption. Each bar in the figure denotes the average disposition effect across retail investors before and after adoption. Panel B plots the cross-section of change in disposition effect before and after adoption, conditioning on the extent of bias ex ante. We categorize all investors into four groups, from low to high levels of bias, based on their average value ex ante. Each bar reports the percentage of investors who who experienced an improvement in the disposition effect (i.e., decreased) after adoption.

Panel A Change in Disposition Effect



Panel B Cross-Section Based on the Extent of Ex Ante Bias

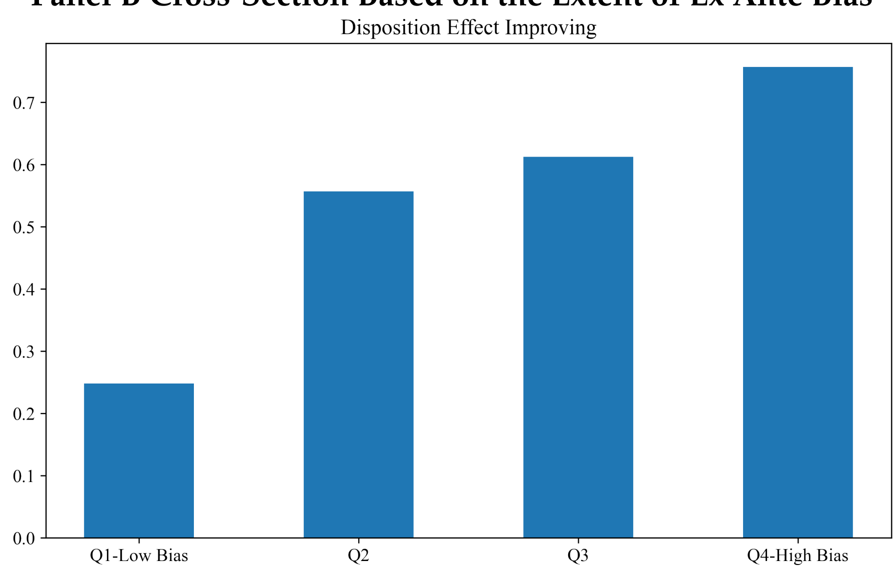
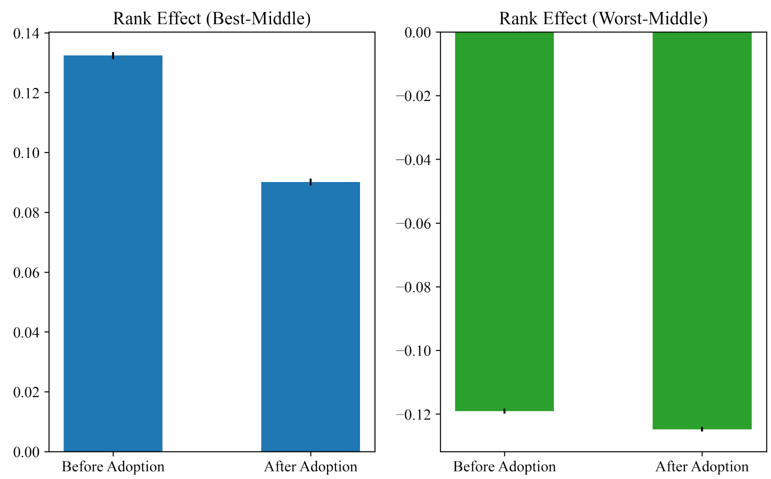


Figure 11: **Retail Algo vs Human Trading: Rank Effect**

Panel A of the figure illustrates the change in rank effect around the time of algorithmic adoption. Each bar in the figure denotes the average disposition effect across retail investors before and after adoption. Panel B plots the cross-section of change in rank effect before and after adoption, conditioning on the extent of bias ex ante. We categorize all investors into four groups, from low to high levels of bias, based on their average value ex ante. Each bar reports the percentage of investors who who experienced an improvement in the rank effect (i.e., decreased) after adoption.

Panel A Change in Rank Effect



Panel B Cross-Section Based on the Extent of Ex Ante Bias

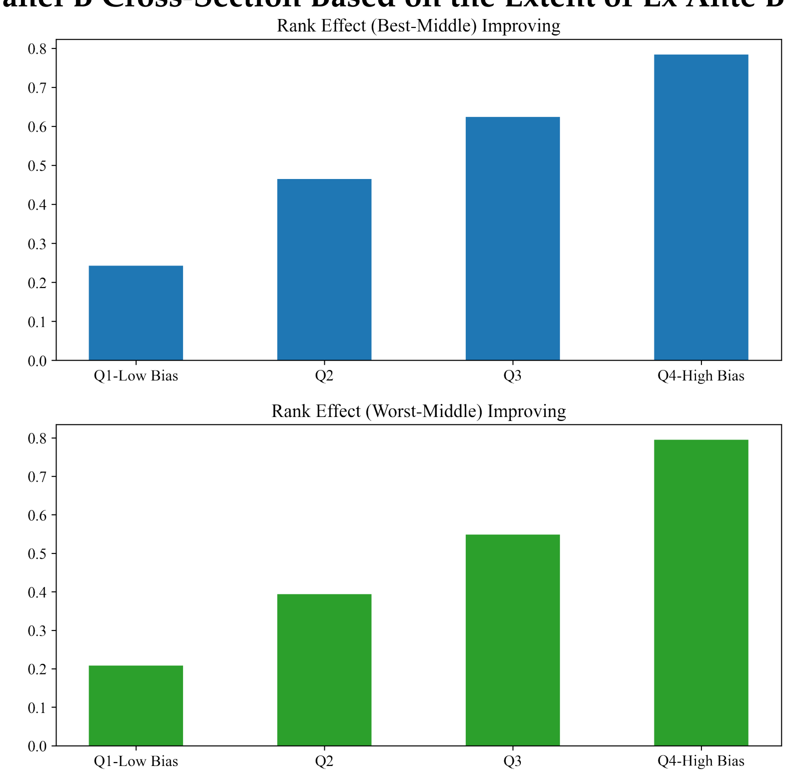


Figure 12: **Algo Trader Survival Rate**

Panel A plots the Kaplan-Meier survival function for the retail algo trader. Entry trade is defined as the first time when the retail trader first adopts algo trading. Quitting is defined as the first trade after which we observe no algo trading in the next 12 consecutive months.

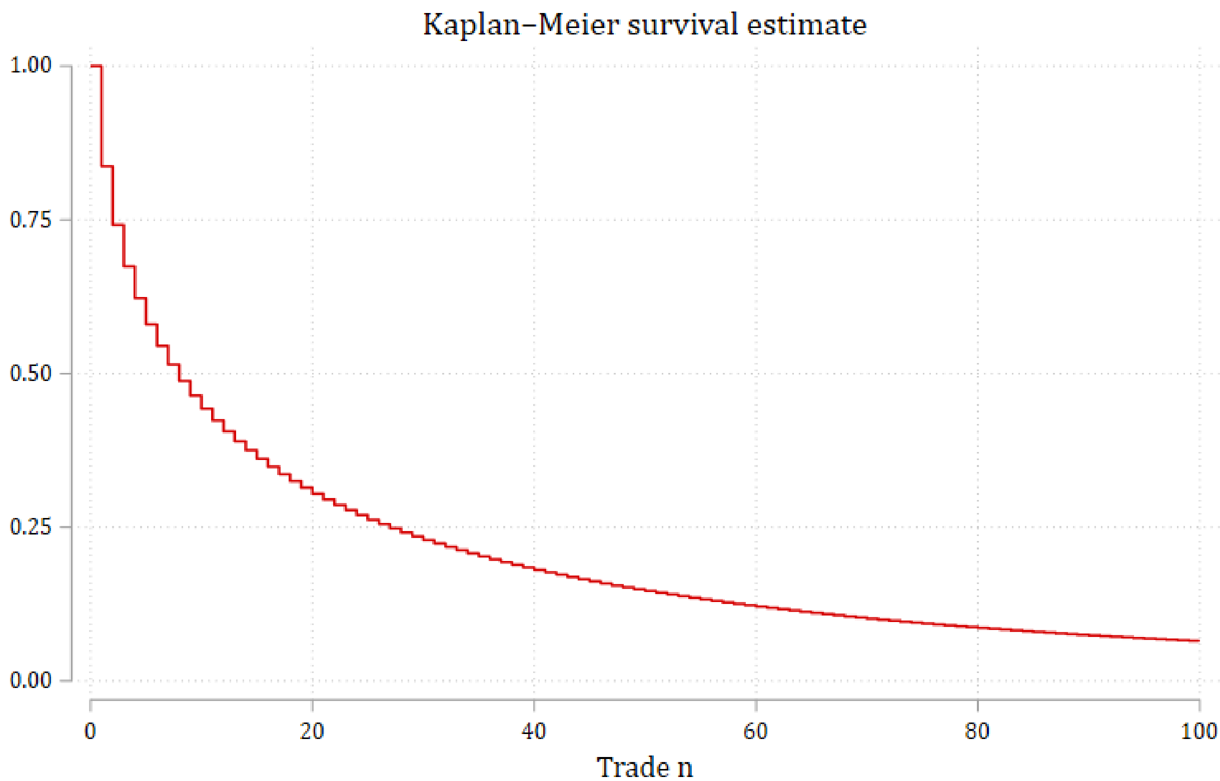


Figure 13: **Retail Algo Trade Size**

Panel A of this figure plots the distribution of initial algo trade size for retail investors. Each bin represents 10,000 India rupees. Panel B presents how the average trade size evolves over the trading sequence. Trade sizes ($n > 1$) are standardized by the initial trade and computed among algo investors who do not exit at trade n.

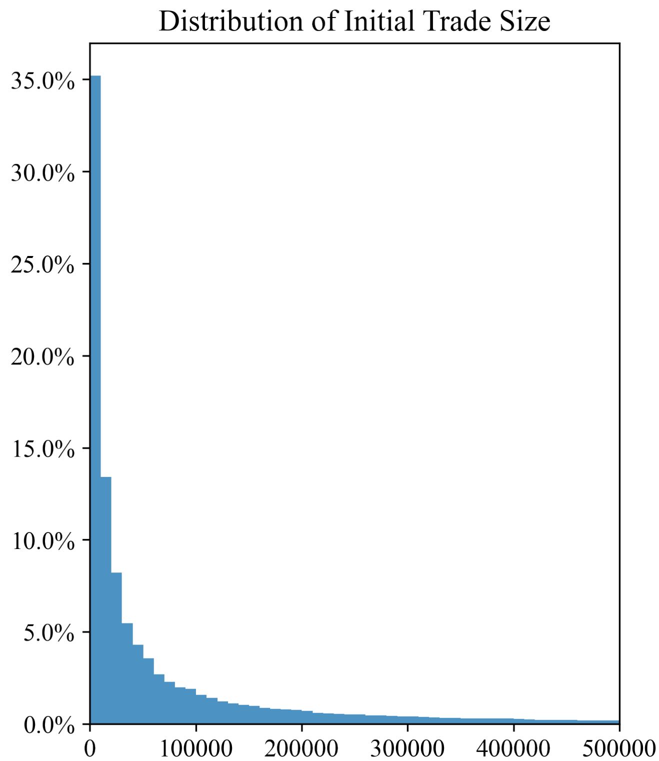


Table 1: **Summary Statistics**

	Mean	SD	p25	Median	p75
Panel A Entire sample					
Number of years with trades (entire sample)	2.86	2.08	1	2	4
Number of stocks traded (per year)	3.36	2.55	1	2.5	4.83
Number of months with trades (per year)	10.23	15.41	1.67	4	11.5
Number of trades (per year)	44.41	104.3	2	7.33	32.5
Average value of shares traded (000, per year)	3,453	11,808	17	115	1,008
Age in 2012	33.29	15.02	22	30	43
Female	0.24	0.43	0	0	0
Panel B Accounts with algo trades					
Number of years with trades (entire sample)	4.02	2.19	2	3	6
Number of months with trades (per year)	5.46	2.66	3.25	5.14	7.5
Number of stocks traded (per year)	26.16	23.78	8	18.25	36.67
Number of trades (per year)	149.97	183.28	23.67	75.5	197
Average value of shares traded (000, per year)	15,272	24,048	533	4,036	17,545
Age in 2012	31.72	13.43	22	29	39
Female	0.17	0.38	0	0	0

This table presents the summary statistics of our sample. Panel A reports means, standard deviations, and percentile distributions of trading activities and demographics for all retail. Panel B provides similar statistics of algo investors.

Table 2: **Investment Performance of Algo Trades vs Human trades**

Variables	Sample	Non-Algo Trades [-12, -1] (1)	[1,12] (2)	Algo Trades [1,12] (3)	Difference	(3)–(1)	(3)–(2)
Panel A 5-day							
Sharpe ratio	Mean	-0.1238	-0.0676	-0.0366	Mean	0.0872***	0.0310***
	SE	1.8759	1.6526	2.0689	t-statistic	20.6237	8.259
Return	Mean	-0.0024	-0.0019	-0.0009	Mean	0.0015***	0.0010***
	SE	0.0314	0.0285	0.0404	t-statistic	18.5999	14.2577
Volatility	Mean	0.0221	0.0228	0.022	Mean	-0.0001***	-0.0008***
	SE	0.0121	0.0112	0.0136	t-statistic	-3.6774	-32.533
Panel B 10-day							
Sharpe ratio	Mean	-0.1179	-0.0776	-0.0462	Mean	0.0716***	0.0314***
	SE	1.7877	1.6439	2.0766	t-statistic	17.1855	8.3588
Return	Mean	-0.003	-0.0026	-0.0015	Mean	0.0015***	0.0011***
	SE	0.0414	0.0376	0.0511	t-statistic	14.7221	12.0977
Volatility	Mean	0.0228	0.0236	0.0226	Mean	-0.0002***	-0.0010***
	SE	0.0109	0.0105	0.0119	t-statistic	-7.019	-42.7701

This table contrasts the investment performance for algo and non-algo trades by retail investors around the adoption time (i.e., $t = 0$). [–12, –1] indicates the window from twelve months to one month prior to the adoption. Performance measures include (1) Return: the value-weighted average of market-adjusted return (5- or 10-day) of all trades submitted in month t; (2) Volatility: for each trade submitted in month t, we first calculate the standard deviation of daily return over the next 5 or 10 trading days and compute the value-weighted average across all trades in a given month; (3) Sharp ratio: for each trade in month t, we obtain the trade-level Sharpe ratio as the ratio of 5- or 10-day market-adjusted return to standard deviation of daily return over the next 5 or 10 trading days and then take the value-weighted average across all trades in a given month. See appendix for detailed variable definitions. The Columns (1)–(3) present the sample statistics for algo and non-algo trades, while the last two columns display the performance differences and t-statistics. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 3: **Who Benefits More from Adopting Algorithmic Trading**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sharpe Ratio Gap(5-day, %)				Sharpe Ratio Gap(10-day, %)			
D(mid-aged)	6.059***				5.103***			
	(0.709)				(0.711)			
D(old)	10.789***				9.898***			
	(1.475)				(1.611)			
Female	2.545***				3.212***			
	(0.860)				(1.154)			
Log_PortfolioValue	-2.662***				-2.520***			
	(0.131)				(0.144)			
Stock No.		-0.152***				-0.168***		
		(0.030)				(0.039)		
Portfolio Concentration		19.974***				19.377***		
		(1.498)				(1.709)		
Sailence_Bias			1.399**				1.894**	
			(0.590)				(0.742)	
Extropolation_Bias				1.908***				1.500**
				(0.719)				(0.739)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	366,103	366,103	366,103	366,103	366,103	366,103	366,103	366,103
Adjusted R-squared	0.003	0.002	0.001	0.001	0.002	0.001	0.000	0.000

This table presents estimates of cross-section regression at the investor level to link the algo-trade-induced performance improvement with investor characteristics. The performance improvement is measured as the average gap in adjusted Sharpe ratio between algo and non-algo trades during the one-year period after algo adoption. In Columns (1) and (5), we explore the relation with investor demographics, including age, gender and level of wealth (proxied by the average value of portfolio during the 12-month period prior to algo adoption). In Columns (2) and (6), we focus on investors' portfolio characteristics, including number of stocks held and portfolio Herfindahl index, during the twelve months before adopting algo trading. Columns (3)–(4) and (7)–(8) present the results when we consider two potential behavioral bias revealed from trading activities ex ante. Availability and extrapolation bias refer to the investor's propensity to purchase attention-grabbing (momentum) stocks. See appendix for detailed variable definitions. District fixed effects are included in all specifications. Standard errors are clustered by district and are reported in parentheses under the coefficient estimate. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: **Algorithmic Trading and Investor Behavior: Trading Responsiveness and Behavioral Biases**

Panel A. Speed (Number of Shares, %)		
	AT After minus Non-AT Before	Non-AT After minus Non-AT Before
Change after adoption	4.450***	3.099***
(p-value)	(0.00)	(0.00)
Panel B. Behavioral Bias (%)		
	Disposition Effect	Rank Effect (Best-Middle)
	After minus Before	After minus Before
Change after adoption	-2.479***	-4.230***
(p-value)	(0.00)	(0.00)

This table tests whether the adoption of algorithmic trading induces significant changes in (1) trading responses to new market information; (2) behavioral biases. Panel A reports the results for trading responsiveness. Each column denotes the average difference between (1) non-algorithmic trading before and after, and (2) non-algorithmic trading before and algorithmic trading after. Panel B reports the results for two measures of behavioral biases. Change in the disposition effect is the difference between the proportion of gains realized (PGR) and the proportion of losses realized (PLR) for each investor before and after the adoption. Change in the rank effect is the average difference between the number of best-performing stocks sold and the number of mid-performing stocks sold before and after the adoption. Each cell reports difference coefficients and the associated p-values.

Table 5: **Cox proportional hazard model of Algo-trade Exit**

	(1)	(2)	(3)	(4)
Panel A Performance of next five trading days				
Return(Trade n)	-1.046***	-0.926***		
	(0.084)	(0.081)		
Trade Average Return(Trade 1 to $n-1$)			-1.097***	-0.956***
			(0.096)	(0.092)
Log Size(Trade n)	-0.065***	-0.077***	-0.065***	-0.077***
	(0.001)	(0.001)	(0.001)	(0.001)
Year-month FE	No	Yes	No	Yes
Observations	565,858	565,858	565,858	565,858
Panel B Performance of next ten trading days				
Return(Trade n)	-0.960***	-0.742***		
	(0.066)	(0.062)		
Trade Average Return(Trade 1 to $n-1$)			-1.050***	-0.781***
			(0.075)	(0.071)
Log Size(Trade n)	-0.065***	-0.077***	-0.065***	-0.077***
	(0.001)	(0.001)	(0.001)	(0.001)
Year-month FE	No	Yes	No	Yes
Observations	565,858	565,858	565,858	565,858

This table presents the estimate of Cox proportional hazard rate model to test the impact of past performance on the investor's decision to quit algo trading. The time-varying covariates include (1) investor's performance of nth trade; (2) the average performance of past trade from 1 to $n-1$; (3) the log-size of n-th trade. Both (1) and (2) are defined based on either 5-day or 10-day return following n-th trade in Panel A and B, respectively. Year-month fixed effects are included in Columns (2) and (4). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 6: **Past Performance and Retail Algorithmic Trading Size**

	(1)	(2)	(3)	(4)
5-day Return (Trade $n-1$)	0.837***	1.020***		
	(0.094)	(0.097)		
10-day Return (Trade $n-1$)			0.826***	0.829***
			(0.071)	(0.074)
Log Size (Trade $n-1$)	0.735***	0.420***	0.735***	0.420***
	(0.001)	(0.001)	(0.001)	(0.001)
Year-month FE	Yes	Yes	Yes	Yes
Trade FE	Yes	Yes	Yes	Yes
Investor FE	No	Yes	No	Yes
Observations	11,588,408	11,534,512	11,588,408	11,534,512
Adjusted R-squared	0.602	0.669	0.602	0.669

This table presents the evidence on the impact of past performance on choices of future algo-trade size. The unit of observation is a single algo trade. The data include the algo-trading records of all retail investors until quitting, which is defined as the last trade after which we observe no algo trading in 12 consecutive months. We regress the log-size of n-th algo trade against the outcome of $n-1$-th algo-trade, measured as 5-day or 10-day return following the trade. For all specifications, we include as controls the log size of algo-trade $n-1$, year-month fixed effects and indicator variables that represent investor's first, second trade, etc. Columns (2) and (4) also consider investor's fixed effects. Standard errors are clustered by investor. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Internet Appendix

Table A1: **Variable Definition**

Variable	Definition
Trading Share$_{g,t}$	Ratio of total algo trading value for investor group g to that of all retail investors at year t. We also consider using the number of algo investors in each investor group to construct similar measures.
Return(trade)$_{i,j,t}$	Ratio of the transaction price of stock j at day t and the stock's closing price (adjusted for splits and dividend) at day $t+h$ ($h = 5, 10, 20, or 30$)
Algo Return(Month)$_{i,t}$	The value-weighted average of market-adjusted return of all algo trades submitted by investor i in month t
Algo Volatility(Month)$_{i,t}$	For each trade submitted in month t, we first calculate the standard deviation of daily return over the next $h(h = 5, 10, 20, 30)$ trading days and compute the value-weighted average across all trades in a given month
Algo SharpeRatio(Month)$_{i,t}$	For each trade in month t, we obtain the trade-level Sharpe ratio as the ratio of h-day market-adjusted return to standard deviation of daily return over the next $h(h = 5, 10, 20, 30)$ trading days and then take the value-weighted average across all trades in a given month
Log(Portfolio Value)$_i$	The average value of portfolio by investor i during the 12-month period prior to her algo adoption
Stock No$_i$	The average number of stocks held by investor i during the 12-month period prior to her algo adoption
Portfolio Concentration$_i$	The average portfolio Herfindahl index held by investor i during the 12-month period prior to her algo adoption
Sailence Bias$_i$	The investor's propensity to purchase attention-grabbing stocks. We first calculate the percentage of investor i's purchase of attention-grabbing stocks among all her trading in month t and compute the average during the 12-month period prior to her algo adoption. We define a stock to be attention-grabbing if it is ranked in the top 10% of performance over the last month.
Extrapolation Bias$_i$	The investor's propensity to stocks after a sequence of positive return. We first calculate the percentage of investor i's purchase of momentum stocks among all her trading in month t and compute the average during the 12-month period prior to her algo adoption. We define a stock to be momentum if its performance is ranked in the top 10% of performance over the last 12 months.
D(young)$_i$	The indicator variable that equals one if the age (as of 2012) of the investor is below 35.

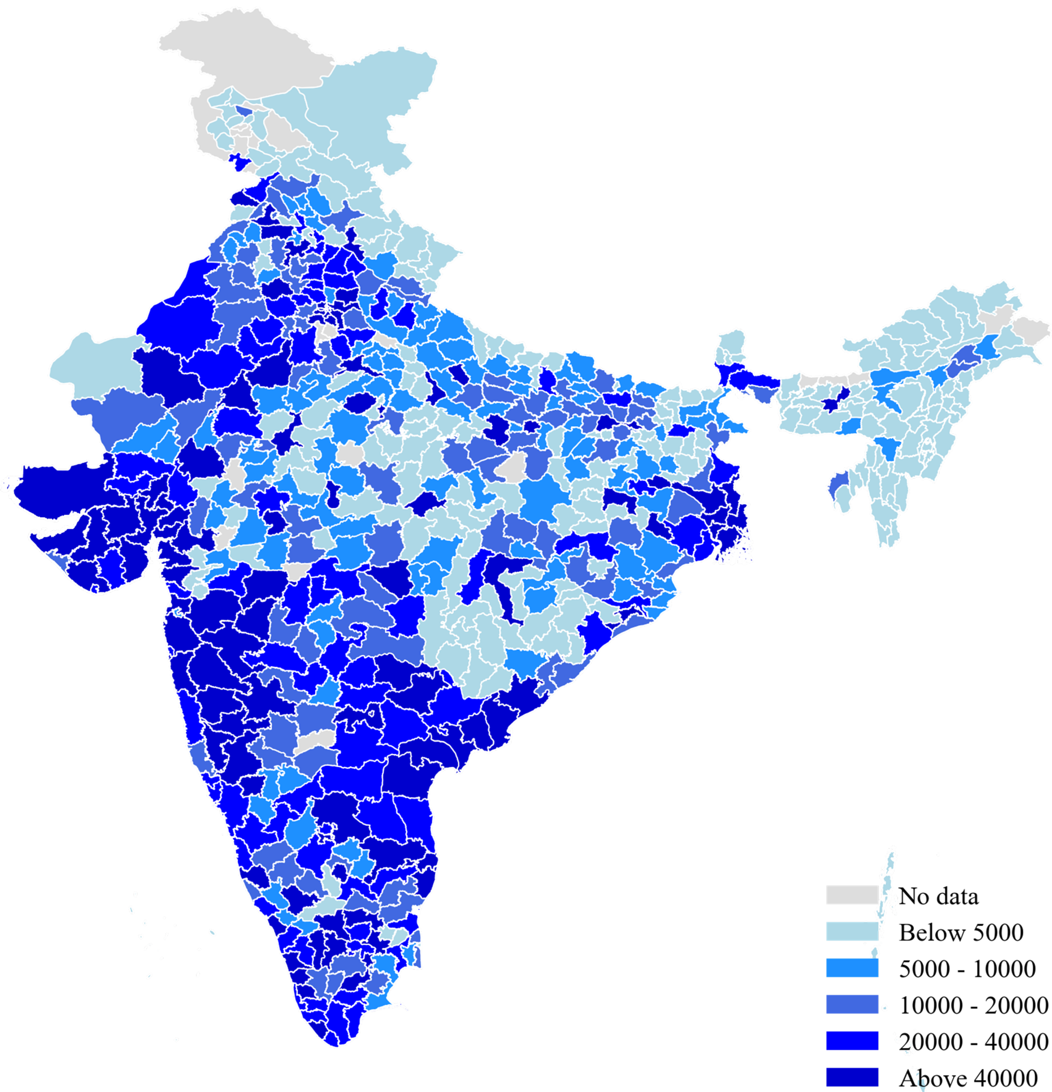
Table A1: **Variable Definition, Continued**

D(mid-aged)$_i$	The indicator variable that equals one if the age (as of 2012) of the investor is between 35 and 60.
D(old)$_i$	The indicator variable that equals one if the age (as of 2012) of the investor is above 60.
Female$_i$	The indicator for female investors.

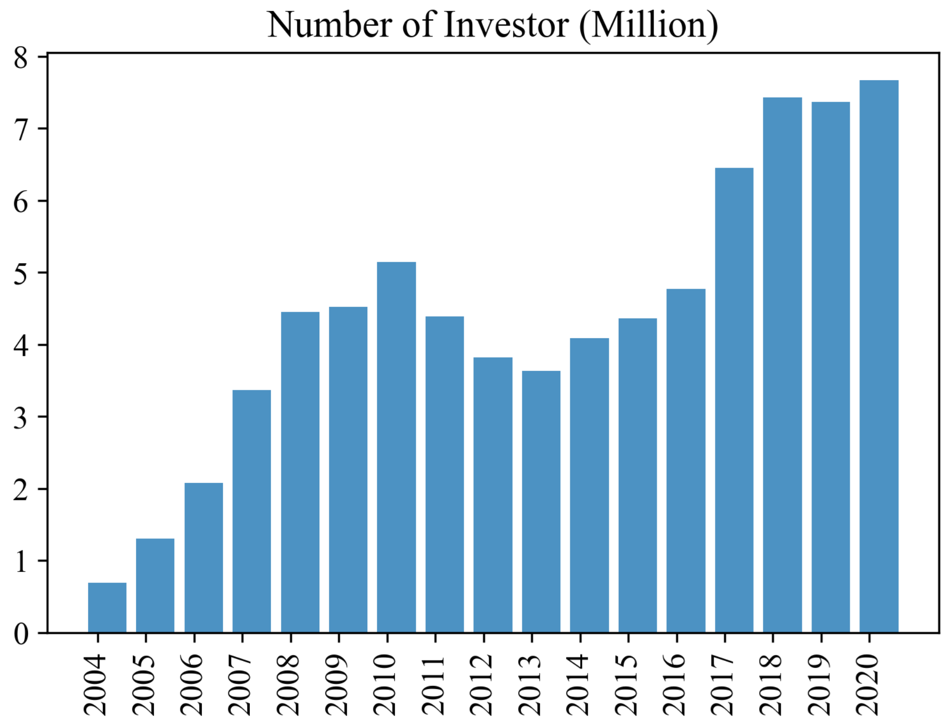
Figure A1: **Statistics of NSE Investors and Trading**

Panel A of this figure plots the number of retail investors across districts who trade at NSE from January 2004 to June 2020. Panel B presents the histogram plot of (1) number of trading investors and (2) total annual turnover for both retail and institutional investors in our data.

Panel A Geographical Distribution



Panel B Time-series Plot of NSE Trading Data



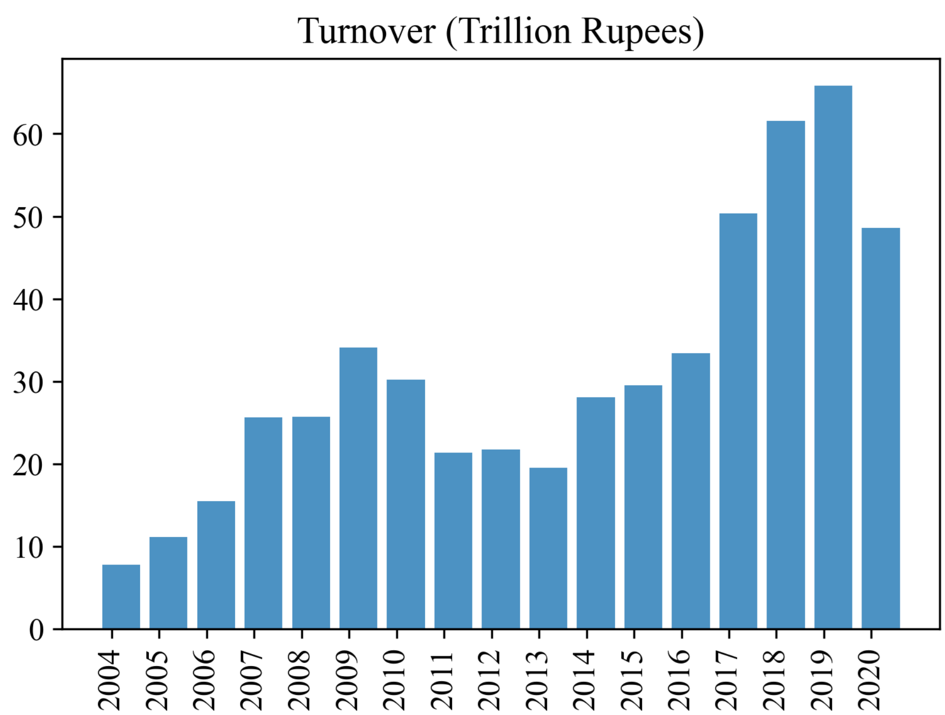


Figure A2: **Composition of Retail Algo Traders: Across Stocks**

This figure presents algorithmic trading share for each investor group among (1) stocks with different market capitalization in Panel A and (2) value and growth stocks in Panel B. Stocks are assigned to groups (small and large stocks) based on their market cap at the end of each year and the breakpoints are 50th percentiles of cross-section distribution. Similarly, value (vs growth) stock is defined based on their price-to-book value at the end of each year and the breakpoints are 50th percentiles of cross-section distribution. The algo trading share is calculated based on the number of algo investors.

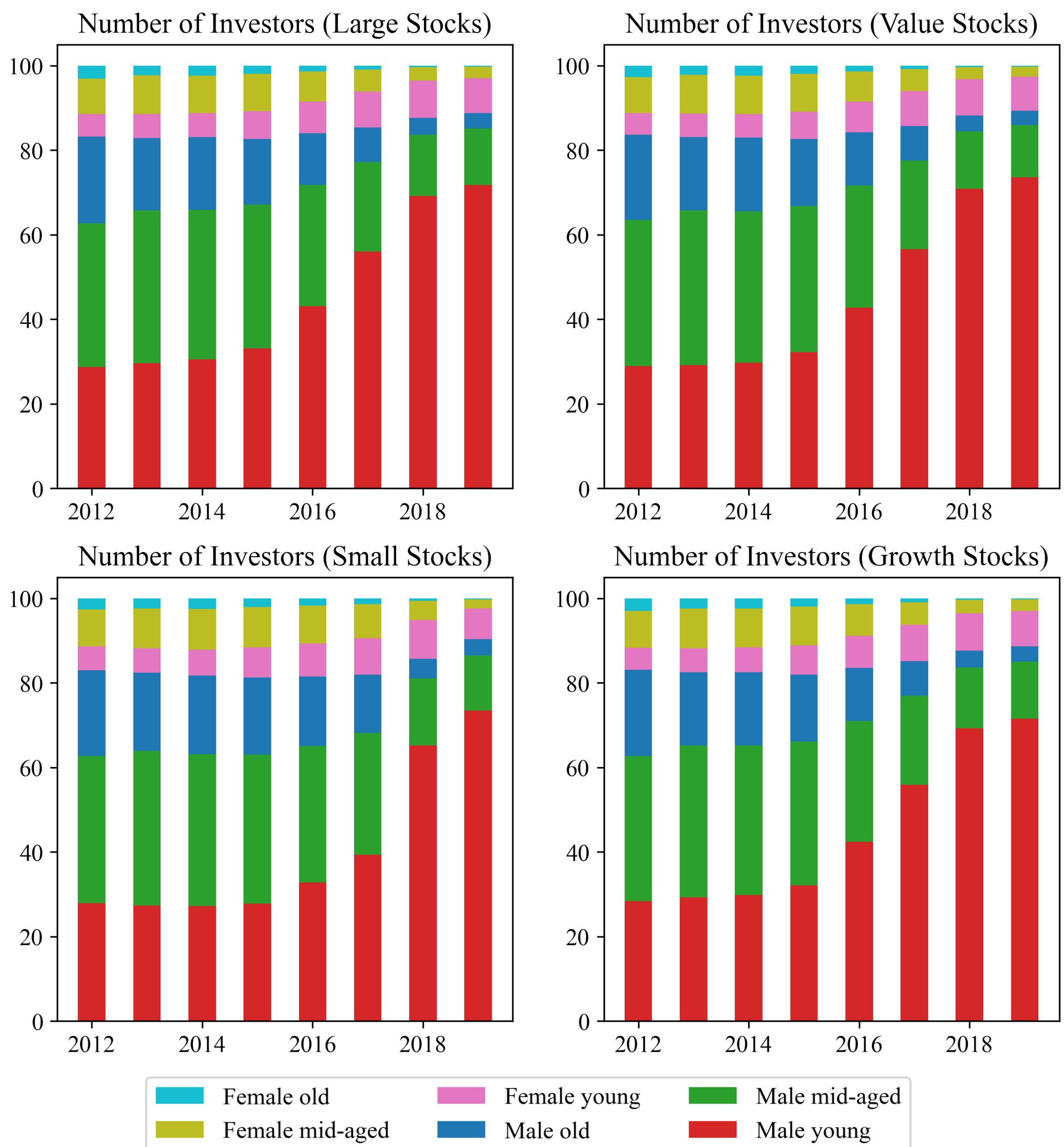


Figure A3: **Performance of Retail Algo vs Human Trading**

This figure contrasts the investment performance for algo vs non-algo (other trading) by retail investors around the algo adoption time, using alternative performance measurement of 20 or 30-day market-adjusted return. Performance measures include (1) Return: the value-weighted average market-adjusted return (20 or 30-day) of all trades in month *t*; (2) Risk (volatility): for each trade submitted in month *t*, we first calculate the standard deviation of daily return over the next 20 or 30 trading days and compute the value-weighted average across all trades in a given month; (3) Sharp ratio: for each trade in month *t*, we obtain the trade-level Sharpe ratio as the ratio of 20/30-day market-adjusted return to standard deviation of daily return over the next 20 or 30 trading days and then take the value-weighted average across all trades in a given month.

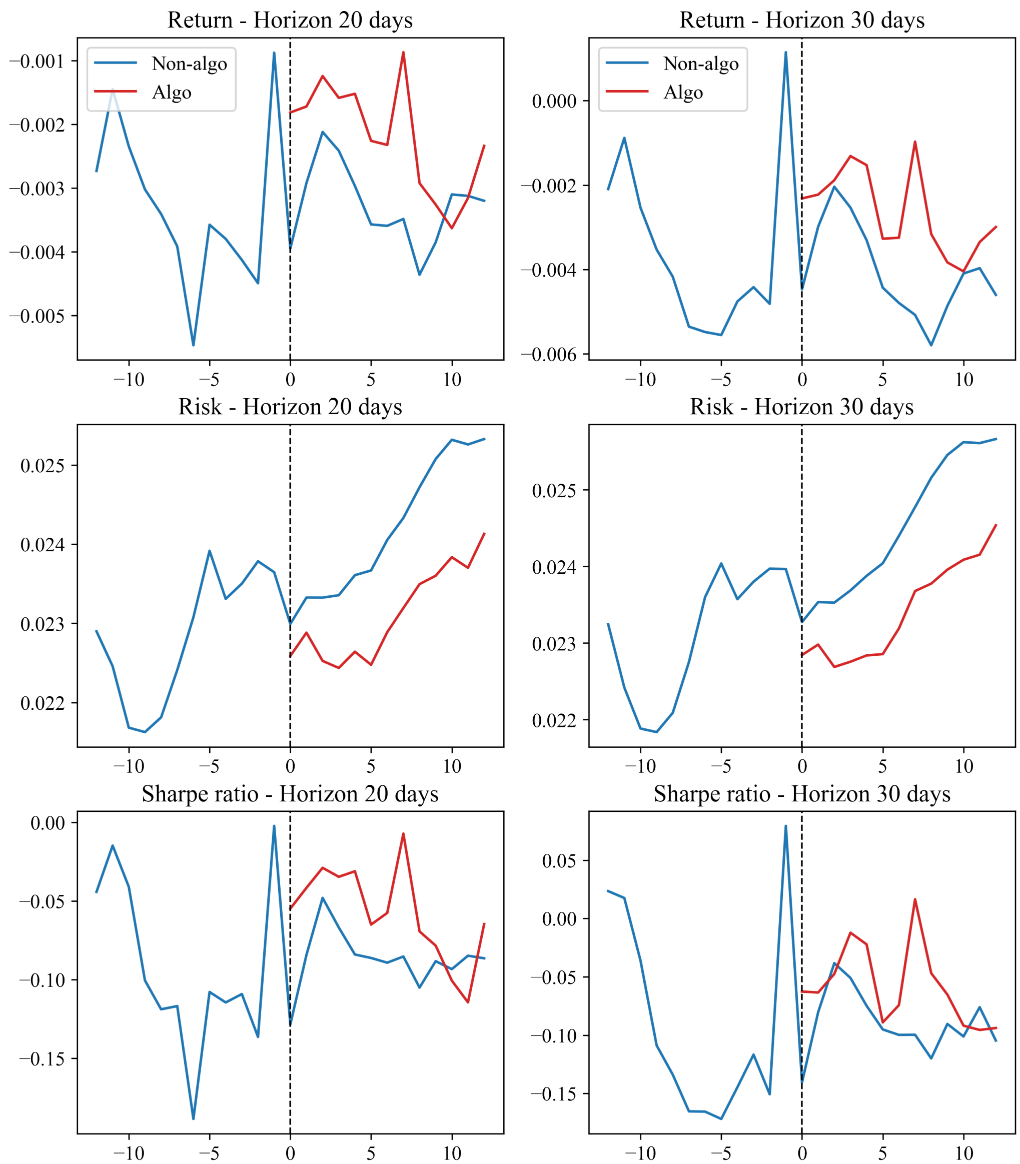


Figure A4: **Hazard Ratio for Algo Trade Exit and Past performance**

The figure presents the hazard ratio for retail algo trading exit conditional on past performance. Quitting is defined as the trade since which we observe no algo trading in the next 12 consecutive months. Both the hazard ratio and 95% confidence interval are reported for different return categories relative to the benchmark return group, (-5bps, 0), where the hazard ratio is one by construction. Return range(in basis points) covers the lower and higher end of the returns earned in the next five trading days.