Not Knowing What to Do With or Without Machine Intelligence: Evidence from a Natural Experiment Involving Retail Investors

Taha Havakhor, Mohammad Saifur Rahman, Tianjian Zhang, Chenqi Zhu

Preliminary Draft[1]. Please Do Not Cite Without Permission

Abstract

The rapid progress in Big Data technologies is commoditizing their applications and ushering an era of artificial intelligence (AI) where non-traditional users too can take advantage of such advancements. In particular, the ecosystem surrounding application programming interfaces (APIs), which increasingly involves freely available and accessible machine learning tools, is creating and supporting new consumers of data and machine intelligence. Arguably, one of the most vibrant and growing new users of big data and predictions are the retail financial market investors. We are, however, in early stages of understanding to what extent these decision makers rely on machine intelligence as well as the impact of this new input to decision making on the general market outcomes. In this paper, we exploit a natural experiment – the abrupt shutdown of Yahoo! Finance API – to offer initial insights into the impact of big data and machine intelligence on the financial markets. Our difference-in-difference design reveals that retail trades drop by approximately 8.0% in a two-month window centered around the shutdown in firms with below-median institutional ownership, relative to firms with above-median institutional ownership, suggesting that even retail investors are significantly reliant on machine intelligence in making trading decisions. Put differently, a sizable portion of retail investors feel helpless in the absence of machine intelligence. Similarly, the market liquidity deteriorated significantly in the same period, which additionally highlights the pervasiveness of machine intelligence and its role in maintaining the market stability. Additional analysis suggests that the investors who disengaged from the market following the API shutdown were involved in less profitable trades compared to the ones who continued to participate in the market. As such, among the consumers of machine intelligence, those with a lower complementary input (such as financial acumen) may be the ones that depended critically on freely available API-enabled predictions.

Keywords: Application programming interface, retail investor, machine intelligence, big data, natural experiment

JEL Codes: G12, G14, O33, H41

Contact information: Taha Havakhor, Fox School of Business, Temple University, taha.havakhor@temple.edu; Mohammad Saifur Rahman, Krannert School of Management, Purdue University, mrahman@purdue.edu; Tianjian Zhang, Spears School of Business, Oklahoma State University, tj.zhang@okstate.edu; Chenqi Zhu, Stern School of Business, New York University, czhu@stern.nyu.edu.

[1] The bulk of the work was done by a student.

1. Introduction

With the declining cost of Artificial Intelligence (AI) – particularly machine learning, machine intelligence fueled by Big Data is becoming a new central input to various decision scenarios (Agarwal and Dhar 2014; Jain et al. 2018; Kleinberg et al. 2017). This trend is made possible by technologies such as application programming interfaces (API), which enable communications among different platforms. In particular, APIs have dramatically reduced the cost to access and connect Big Data platforms with machine learning tools, thus expanding the user-base of AI-enabled predictions beyond wealthy institutions. Aside from the increase in consumptions by traditional users (e.g. bankers) and the addition of new usages (e.g. genetics) (Agrawal et al. 2018; Meyer et al. 2014; Mullainathan and Spiess 2017), the widespread introduction of APIs and the traction of open software development have the potential to make machine intelligence a vital ingredient of decision making for ordinary decision makers, especially less sophisticated ones in face of demanding prediction needs. As such, we focus on retail investors – i.e., individuals who invest in the stock market for their own personal account rather than for an institution or organization – and their trading activity in financial markets. Specifically, this study intends to provide early insights into machine intelligence consumptions by ordinary retail investors and focuses on unfolding: a) the extent of machine intelligence usage by these new users – retail investors – in a traditional area, b) within these new users, who are most affected by the cheaply available machine intelligence in their decision making, and c) the impact of such intelligence on the market outcomes.

Institutional investors have been historically believed to be better equipped with the talent and technological inputs to turn raw information about stocks, quotes, and financial summaries into valuation signals to guide their investment decisions. On the contrary, retail investors either lacked the talent or the technology or both to make predictions for their investments. The early literature finds that retail investors are subject to various behavioral biases and information processing limitations (Barber and Odean 2007), and suffer more losses than institutional investors (Barber et al. 2008). Recent studies find that retail investors have become more sophisticated over time (Kelley and Tetlock 2013; Kelley and Tetlock 2016), which may be partly due to the advent of new technology and data that assist their decision-makings (Farrell et al. 2018). With the introduction of financial API platforms (such as Google's and Yahoo's) and the fast propagation of open financial prediction applications on those platforms, retail investors could plausibly make more informed decisions at a much lower cost. As such, this allows us to examine how machine intelligence impacts the financial market through retail investors.

To achieve this goal, we capitalize on the natural setting created by the abrupt closure of Yahoo! Finance API. On May 16, 2017, Yahoo! Finance API was suddenly shut down without any warnings or announcements, causing major disruptions for those using it. This exogenous shock provides us with a unique experimental setting to examine the impact of API-enabled, low-cost, predictions on retail investors' trading behavior (relative to institutional investors) and the broader impacts on the liquidity of stock market. Obtaining data related to the actual consumption of prediction by retail investors is incredibly difficult. However, given financial APIs such as Yahoo's are the main source of feeding into financial prediction applications and have a low value on their own, when not combined with a predictive application, the shutdown of Yahoo! Finance API is a good proxy event equivalent to treatment condition of withholding access to API-enabled predictions from the retail investors. Specifically, we employ a difference-in-difference (DID) design by leveraging the well-documented clientele effect that retail investors prefer a different set of public firms relative to institutional investors (Barber and Odean 2000; Ivković et al. 2008).

Therefore, the impact of the Yahoo! Finance API on the extent of prediction consumption by new users can be understood by comparing the extent of investment in individual-preferred public (treatment group) with the extent of investment in institution-preferred firms (control group) before and after the Yahoo! Finance API shutdown.[2] Moreover, we focus on financial acumen as an important complementary skill in processing stock market predictions. The financial acumen of retail investors (as new consumers of API-enabled predictions) is then understood by comparing the profitability of retail investments before and after the shock.

From the DID estimations, we find that retail trades drop by approximately 8.0% in a two-month window centered around the shutdown of Yahoo! Finance API in firms with below-median institutional ownership, relative to firms with above-median institutional ownership. This significant drop highlights the important role of API-enabled predictions for retail investors in today's capital markets. When we extend the window to four or six months, the decrease in retail trades becomes much smaller and statistically insignificant, suggesting that investors gradually switched to other similar technology (e.g., other APIs) or learned to cope with decision making without machine intelligence, albeit a less likely possibility. The immediate drop in retail trades in a short window after the API shutdown provides us with confidence in attributing the drop to the shutdown. Moreover, institutional trading volume does not change significantly in either the relatively short or long windows, mitigating the concern of confounding events that change overall market conditions and lead to fewer trades by both retail and institutional investors.

Also, since retail trades have been shown to provide market liquidity, both theoretically and empirically (Grossman and Stiglitz 1980; Kelley and Tetlock 2013), we examined the impact of Yahoo! Finance API shutdown on the liquidity in markets supported by retail traders versus those

[2] Our estimates are likely to be the lower bound as retail investors also trade on institution-preferred public firms (i.e., firms with above-median institutional holdings), although less frequently.

depending mostly on institutional traders. Consistent with the lack of retail trades lowering the overall market liquidity, we find that the Amihud illiquidity measure (price impact per share traded) and bid-ask spread increases by 5%-12% in the month after the shutdown of Yahoo! Finance API for firms with low institutional holdings, relative to firms with high holdings. The deterioration in overall market liquidity underscores the importance of API-enabled intelligence in the fast-paced stock market.

Our results are obtained after controlling for market conditions (signed and unsigned stock returns, news coverage (Da et al. 2011)) as well as common firm characteristics. The findings are not driven by unobserved firm heterogeneity as the results are essentially unchanged with firm fixed effects. We also control for date fixed effects in each regression, hence macroeconomic shocks should not explain our results. Nevertheless, one may be concerned that firms with low institutional holdings are inherently different from those with high holdings and they have different sensitivities to other changes that coincide with the shutdown of Yahoo! Finance API (e.g., retail investors are more prevailing in firms with low institutional ownership and they trade less frequently in June than May as summer vacations distract them). To address this concern, we conduct placebo tests by repeating the same analyses around May 16, 2016, one year before the actual shutdown of Yahoo! Finance API. If our findings are driven by the distinct sensitivities to other changes such as retail trading seasonality, we should also observe a drop of retail trades within two months after May 16, 2016. However, we do not find any significant changes in retail trades or market liquidity in the placebo sample, thus alleviating this concern and providing further support to our conclusion.

To unfold more about the financial acumen of retail investors that consume API-enabled prediction, we compared the relative profitability of retail trades before and after the shock. Since our previous analysis suggests a decrease in retail trade due to the API shutdown, a comparison between the profitability of retail trades before and after the shock can provide a proxy estimate of the financial acumen of those who stopped trading immediately after the shutdown. Our analysis shows an increase in the profitability of retail trades (conditional on the trades are still executed) immediately after the shutdown. This suggests that those retail investors stopping trades after the shutdown had a lower financial acumen relative to other retail investors. This finding implies that retail investors with a lower complementary input (such as financial acumen) may be the ones that depend critically on API-enabled intelligence. Before discussing the details of our study design and its findings, we provide further information about its setting in the following section.

2. Yahoo! Finance API

Yahoo! Finance is the most popular website for financial information, attracting over 30 million unique daily users[3] (Lawrence et al. 2016). Most users of Yahoo! Finance API are retail investors, as institutional investors have easy access to professional data sources (Da et al. 2011; Lawrence et al. 2016). In addition to visits to the Yahoo! Finance website, many retail investors, especially relatively more sophisticated ones, use its data directly or indirectly through its free API. This API allows users to access historical market data (e.g., stock price, trading volume), intraday real-time market data, as well as some basic financial variables. Many retail investors use API to identify potential investment targets, develop investment strategies, and streamline the portfolio

[3] While the data on the exact market share of Yahoo! Finance API is unavailable, we used Google and Github search volumes to estimate its popularity. Based on various forum posts, we identified four main alternatives of Yahoo! Finance API, namely Google Finance API, AlphaVantage, Intrinio, and Tiingo. The Google trend index measures the relative search frequencies of one or more key words in a given period of time. The highest search volume in the period is assigned the score 100. Figure 1 shows that Yahoo! Finance API was consistently the top financial API until its shutdown in mid-2017. While longitudinal search data is not available on Github, the same search terms suggest Yahoo! Finance API is the most popular financial API on Github as of May 24, 2019 (Figure 2).

rebalance. Some web application developers also use API to redistribute the data to a broader investor base.

Yahoo! Finance API was abruptly shut down on May 16, 2017. While Yahoo! did not offer any notice before the shutdown or explanations afterwards, it was speculated that the main reason for the shutdown was due to financial concerns[4]. While Yahoo! Finance API was widely used, there are a few alternatives that provide similar functionalities. After the shutdown, posts on online communities suggested various alternatives such as Alpha Vangtage and Intrinio. It is probable that some Yahoo Finance API users are aware of these alternatives even before the shutdown, but instant switching to different APIs is unlikely given the learning curves and the fact that some users initially might be waiting for the Yahoo API to be back online. While we do not know when and how many of the Yahoo API users switched to alternative APIs[5], we can get an indirect picture through our analyses of trading volume and market liquidity following the abrupt shutdown.

3. Sample and Data

Our sample firms are Compustat-CRSP firms that exist both before and after the shutdown of Yahoo! Finance API (May 16, 2017). Our main sample period ranges from April 16, 2017 to June 15, 2017 (inclusive), a two-month window centered around the shutdown. We also use longer windows (four- or six-month window around the shutdown) for supplemental analyses or a two-month window in 2016 (one year before the main sample period) for a placebo analysis.

We obtain firm characteristics, stock performance, analyst following, and institutional ownership from standard data sources (Compustat, CRSP, IBES, and Thomas Reuters). To avoid looking-ahead biases, all variables are measured at the latest fiscal year ending before January 1st,

[4] http://blog.intrinio.com/yahoo-finance-api-replacement/

[5] Instant switcher biases against us from finding significant impact of API-enabled machine intelligence on retail trading in the stock market.

2017 hence should be available when trading decisions are made. We identify retail and institutional trades from the Trade and Quote (TAQ) database following Boehmer et al. (2017) and Bushee et al. (2019). The idea behind the classification is that retail trades are often executed off-exchange and offered a small price discount relative to the national best bids and offers (Boehmer et al. 2017). Specifically, we classify retail sale (buy) trades as those with TAQ exchange code "D" (indicating off-exchange trades) and prices 0.1-0.4 cents above (below) a round penny. To be more conservative, trades with prices at a round penny or near the half-penny (0.4-0.6 cents, inclusive) are not classified. Non-retail trades larger than $50,000 are classified as institutional trades (Bushee et al. 2019), as larger trades are more likely to be submitted by institutional investors. Our classification misses out some retail trades and institutional trades as not every retail trade is off-exchange or receives price discount and institutional investors nowadays often break down their trades into smaller ones. Boehmer et al. (2017) compare this classification to a proprietary datasets of retail trades and show the retail trades identified by this classification are representative. Moreover, the measurement errors should not systematically affect our results as this classification results in similar errors both before and after the API shutdown. To make the trading volume more comparable across firms, we scale the shares traded by retail or institutional investors by total shares outstanding and remove its normal level (the corresponding median scaled trading volume for the same day of the week over the past ten weeks) to construct the abnormal trading volume (*Ab_Retail_Vol* and *Ab_Institutional_Vol*).[6]

We construct two widely-used daily measures of market liquidity. The first one is Amihud's illiquidity measure (*AIM*), which quantifies the price impact per shares traded (Amihud 2002). The second one is the bid-ask spread (*Spread*), the difference between the ask price and bid price,

[6] The trading volume aggregates shares traded in buy and sale transactions. We do not differentiate buy from sale transactions in the main analyses as they are both affected by the API availability.

scaled by the middle point of the bid and ask. A higher value of AIM and Spread indicates lower liquidity (higher illiquidity). Both variables are constructed based on daily data from CRSP. Please refer to Table A1 for details on variable definition and data sources.

Panel A of Table 1 reports the summary statistics for the main sample, which includes 192 thousand daily observations from April 16, 2017, to June 15, 2017, for 4,481 unique firms. To minimize the influence of outliers, all variables are winsorized at 1% and 99% except dummy variables and variables that have been taken logarithm. On a typical day, the retail (institutional) trades we identified from TAQ account for 0.08% (0.08%) of shares outstanding. After removing the normal level of retail (institutional) trading volume, the average abnormal retail (institutional) trades account for 0.02% (0.03%) of shares outstanding. The average firm is modestly large and levered, regularly covered by media and financial analysts. The average ROA is negative, but the median ROA is slightly positive.

Panel B splits the sample firms based on the percentage of institutional holdings as reported in 13-f filings on December 31, 2016. *Low_IH*=1 (treatment group) indicates firms with below-median institutional holdings and *Low_IH*=0 (control group) indicates below-median. Pre and Post indicate one month before and after the shutdown of Yahoo! Finance API, respectively. The univariate analysis shows that both the raw and abnormal retail trades fall significantly in the post period for firms with low institutional holdings. Specifically, *Retail_Vol* declines from 0.112 to 0.103, an 8.1% decrease in a month, statistically significant at 1%. We do not observe a significant decrease in retail trades for firms with high institutional holdings during the same period. Therefore, the univariate difference-in-differences (DID, the post-pre difference for firms with low institutional holdings relative to the same difference for firms with high institutional holdings) is also significantly negative. Institutional trades reduce slightly in the post period and the magnitude is similar for both group of firms, hence the DID estimate is insignificant. We also observe significant deterioration in market liquidity (increases in *AIM* and *Spread*). The daily measures of control variables do not change much around the Yahoo! Finance API shutdown, except the absolute value of stock return increases slightly but the increase is only significant at 10% level.

4. Empirical Analyses

4.1 Regression Specification

In this section, we formally test the impact of the shutdown of Yahoo! Finance API on trading and market liquidity using the following regression specification.

$$Outcome_{it} = \alpha + \beta \cdot Post_t \times Low\_IH_i + \gamma \cdot W_{it} + Date\ FE + Firm\ FE + \epsilon_{it}.$$

where i represents the firm and t the date. The outcome variable is abnormal retail or institutional trades (*Ab_Retail_Vol* or *Ab_Institutional_Vol*) or market liquidity (*AIM* or *Spread*). The key variable of interest is the interaction term *Post* × *Low_IH*. Both are dummy variables, indicating the period after the shutdown of Yahoo! Finance API and the firms with below-median institutional holdings, respectively. They are not included in the regressions independently because their direct impacts are absorbed by date and firm fixed effects, respectively. β is a DID estimate, with a positive (negative) value indicating that the outcome variable increases (decreases) after the shutdown in firms with low institutional holdings, relatively to those with higher holdings. W_{it} represents a set of firm-day level control variables including stock return, absolute value of stock return, and news coverage, following Da et al. (2011). Date fixed effects control for any changes in macroeconomic conditions that affect firms with both high and low institutional holdings. Firm fixed effects control for time-invariant firm heterogeneities. In robustness checks, we drop firm fixed effects and add a set of common firm characteristics measured as of the latest fiscal year ending before January 1, 2017.

4.2 Retail and Institutional Trades

To illustrate the changes in retail trades around the shutdown of Yahoo! Finance API, we plot average daily retail trading volume scaled by shares outstanding (*Retail_Vol*) in Figure 3. The solid line represents firms with below-median institutional holdings and the dash one above-median institutional holdings. The solid line is always above the dashed line, which is not surprising given that retail trading is generally more active in firms with low institutional holdings. Despite the level differences, the two lines follow similar trends in the month before the shutdown, validating the parallel trend assumption for DID analyses. Within three days after the shutdown, the solid line drops substantially, while the dashed line does not change noticeably except a minor spike. The sharp decrease in retail trades suggests that the drop in retail trades is very likely due to the Yahoo API shutdown.

We report the regression results of abnormal retail trades in Panel A of Table 2. Columns 1 and 2 use a two-month window centered around the shutdown of Yahoo! Finance API ([April 16, 2017, June 15, 2017]). In Column 1, we include a set of common firm characteristics and industry fixed effects. In Column 2, firm fixed effects are included and firm characteristics are dropped as they do not change during our sample period hence are absorbed by firm fixed effects. Consistent with the graphic evidence, both columns report a negative coefficient on the interaction term (*Post*×*Low_IH*), significant at 1% level. This result suggests that abnormal retail trades drop significantly after the shutdown for firms with low institutional holdings. Once we extend the sample period to four months in Columns 3-4 and to six months in Columns 5-6, the coefficient on *Post*×*Low_IH* gradually becomes smaller and statistically insignificant, suggesting that retail investors gradually find alternative data sources to substitute Yahoo! Finance API.

Panel B of Table 2 reports the evidence on abnormal institutional trades. The coefficients on *Post×Low_IH* are insignificant across the three different sample periods. The only exception is Column 5 where the coefficient is significant at 10% level with industry fixed effects. Once firm fixed effects are included, the coefficient becomes insignificant again (Column 6). Overall, the insignificant change in institutional trades around the shutdown of Yahoo! Finance API reassures that our finding is not caused by confounding events that affect trade decisions in general.

For control variables, we find that both retail trading and institutional trading increase on days with larger stock movements and news coverage. Interestingly, retail investors are less active in loss firms or firms with intensive R&D, possibly because these firms are too complicated for retail investors to understand. In contrast, institutional trading volume increases with R&D intensity as they are in a better position to understand the risk and potential of R&Ds. Moreover, both retail and institutional investors are more active in firms with intensive advertising expenditures, which is consistent with advertising campaigns increase firms' visibility and attract investor attention.

4.3 Market Liquidity

In this section, we study the change in market liquidity around the shutdown of Yahoo! Finance API. On the one hand, theoretical and empirical evidence shows that the active participation of retail investors in the stock market improves the market liquidity (e.g., Grossman and Stiglitz 1980; Kelley and Tetlock 2013). More retail investors make it easier for any given investor to find a counterparty to trade with. More importantly, in a market with relatively more retail investors, who are less sophisticated on average, any given investor will be less concerned with adverse selection (trading with counterparties with information advantage). Consequently, investors are more willing to trade with each other, resulting in a lower price impact per share traded and lower bid-ask spread (Greene and Smart 1999; Han et al. 2016). Based on the above intuition, we expect that market liquidity deteriorates after the shutdown of Yahoo! Finance API, as the shutdown leads to fewer retail trades. On the other hand, if the exit of retail investors is compensated by institutional investors, the stock market liquidity should not change much. Taken together, it is an empirical question as to how market liquidity changes around the shutdown of Yahoo! Finance API.

Table 3 shows that the stock market becomes more illiquid after the shutdown. Economically, according to the estimates reported in Column 2(4), AIM (Spread) increase by 11.7% (5.0%) after the shutdown for firms with low institutional holdings, relative to firms with high institutional holdings. This economically considerable deterioration in market liquidity underscores the importance of AI-enabled decision-making in the full functioning of the overall market.

4.4 Placebo Tests

One may be concerned that our findings are driven by the inherent differences between firms with low versus high institutional holdings. The level difference in these two groups should not affect our DID estimates (firm fixed effects control for the level differences), as long as the level does not change after the API shutdown for reasons unrelated to but coincide with the shutdown. Similarly, our DID estimates are also not affected by other events that happened after the API shutdown (date fixed effects control for the average impact of these events), as long as other events do not affect the two groups differently. The remaining concern is that the two groups of firms have different sensitivities toward other events. One such example is that retail investors are more prevailing in firms with low institutional ownership and they trade less frequently in June than May as summer vacations distract them. To test such possibility, we repeat the same analyses on the same set of sample firm in a placebo period, a two-month window centered around May 16, 2016, one year before the actual shutdown. If our findings are driven by confounding events that affect the two groups differently, we should find similar results in the placebo sample.

The results using the placebo sample are reported in Table 4. The regression specifications are exactly the same as the ones used in Table 2 and 3. Across the 8 regression specifications, none of the coefficients on the interaction term is significant, regardless of the inclusion of industry or firm fixed effects. The non-results provide us with more confidence in attributing our findings to the shutdown of API rather than other confounding events that somehow affect the two groups differently. The Appendix section further discusses additional robustness checks.

4.5 Characteristics of API Consumers

So far, the analyses have addressed how Yahoo! finance API shutdown impacts trading volumes and market liquidity. To better understand the financial acumen of the API consumers, we examine the association between the extent of retail trades (we separate buys and sells due to the inherent nature of their trading) and the subsequent cumulative abnormal return of the traded stocks. A difference in the profitability of investment targets before and after the shock, combined with our already-established knowledge about the decrease in retail investments after the shock, provides insights about the extent of financial acumen of the API consumers. We use the Fama-Macbeth regression for the pre- and post-period, as follows.

$$CAR = \alpha + \beta_1 \cdot Ab\_Retail\_Buy + \beta_2 \cdot Ab\_Retail\_Sell + \eta \cdot Z + \xi,$$

where *Ab_Retail_Buy (sell)* stands for the buying (selling) volume by retail investors (scaled by outstanding shares) with a median adjustment (see Table A1 for the detailed definitions). Z represents a set of firm-day level control variables (see Table 5) following Kelley and Tetlock (2013). Table 5 reports the regression results on the relative profitability of retail buys and sells in the two-month window around the shutdown. The dependent variable is the buy-and-hold cumulative abnormal returns (CAR), that is, cumulated buy-and-hold individual stock return minus the corresponding market return over different horizons (i.e., buy and hold the stock for 1 week, 3 weeks, etc.). CAR[1W] is calculated daily for individual stock's one week CAR starting from the next day. CAR[2W, 4W] is CAR from the beginning of week 2 to the end of week 4. CAR[5W, 8W] is similarly defined. The results across the six regressions using the entire sample (columns 1-6) suggest the significant results are concentrated in the first month (CAR[1W] and CAR[2W, 4W]) and the significance fades as the horizon extends to the second month. Subsample analyses for control and treatment groups are therefore conducted for CAR[1W, 4W].

We performed a "difference-in-difference" comparison of the coefficients for the four subsamples (treatment_post, treatment_pre, control_post, control_pre) using Welch's t-tests, and found that retail buys become relatively more profitable (or less loss-making) in the treated firms after the API shutdown. We did not find significant changes for retail sells, which may be due to the fact that selling stocks could be driven by liquidity reasons, independent from the availability of API-enabled prediction products. This evidence suggests the absence of API-enabled prediction products filters out less sophisticated retail investors, resulting in a higher profitability of the average retail trade conditional on the trade taking place. We report similar analyses with four- and six-month windows centered around the API shutdown in Table 6 and 7, respectively. With recovery of trading volumes after one month (Table 2), differences in profitability between the control and treatment groups gradually converged and eventually vanished (Table 6).

5. Conclusions

Put together, the study provides preliminary insights about the introduction of new users (i.e., retail investors) to a traditional area of prediction usage (i.e., stock market investment). Consistent with prior studies that documented beneficial applications of forecasting via machine intelligence (Kleinberg et al. 2017; Miklós-Thal and Tucker 2019), our findings suggest the absence of API-enabled prediction products negatively impacts the financial market. Moreover, the evidence after the immediate shutdown of the API shows an increased profitability to retail trades, indicating that the new consumers who critically depend on the API-enabled economy of prediction products are likely to have a deficit in an input to decision making which is complementary to prediction, i.e., judgement (Agrawal et al. 2019). This study, to the best of our knowledge, is among the first attempts to document the ways in which the consumption of machine intelligence and prediction is changed in recent years and highlights the new trend in including new users in traditional areas where prediction is applied. Despite its findings, the study is a preliminary effort in this area. Specifically, the question about the *nature* of prediction consumption by new users is an essential inquiry to be pursued. Future studies shall address how retail investors' (new users') prediction consumption is different from those made by institutional investors (traditional consumers).

In addition, the study offers insights to the stock market context. The performance of stock markets is often influenced by reducing information frictions through new technologies (Clemons and Weber 1997; Dewan and Mendelson 1998; Zhang and Zhang 2015). Retail investors, in particular – who may not have an army of financial consultants at their service, may face more information frictions compared to their institutional counterparts. Retail investors are an important part of the stock market. 14% of US households directly invest in individual stocks (Bricker et al. 2017), and they as a whole own $16.8 trillion in stocks, more than any other single investor groups including mutual funds and hedge funds (Fed 2018). Moreover, retail investors are of great importance to regulators[7] . As such, reducing information frictions, through increasing the access to reliable prediction products, can eradicate major risks that retail investors otherwise endure and result in an inflow of micro investments to markets.

[7] https://www.sec.gov/news/speech/remarks-economic-club-new-york; Current SEC Chairman Jay Clayton stated, "Our analysis starts and ends with the long-term interests of the Main Street investor [retail investor]."

References

Agarwal, R., and Dhar, V. 2014. "Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research," *Information Systems Research* (25:3), pp. 443-448.

Agrawal, A., Gans, J., and Goldfarb, A. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Press.

Agrawal, A., Gans, J. S., and Goldfarb, A. 2019. "Exploring the Impact of Artificial Intelligence: Prediction Versus Judgment," *Information Economics and Policy* (Forthcoming).

Amihud, Y. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets* (5:1), pp. 31-56.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., and Odean, T. 2008. "Just How Much Do Individual Investors Lose by Trading?," *The Review of Financial Studies* (22:2), pp. 609-632.

Barber, B. M., and Odean, T. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *The Journal of Finance* (55:2), pp. 773-806.

Barber, B. M., and Odean, T. 2007. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *The Review of Financial Studies* (21:2), pp. 785-818.

Boehmer, E., Jones, C. M., and Zhang, X. 2017. "Tracking Retail Investor Activity." SSRN 2822105.

Bricker, J., Dettling, L. J., Henriques, A., Hsu, J. W., Jacobs, L., Moore, K. B., Pack, S., Sabelhaus, J., Thompson, J., and Windle, R. A. 2017. "Changes in Us Family Finances from 2013 to 2016: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin* (103).

Bushee, B. J., Cedergren, M. C., and Michels, J. 2019. "Does the Media Help or Hurt Retail Investors During the IPO Quiet Period?." SSRN 3160949.

Clemons, E. K., and Weber, B. W. 1997. "Information Technology and Screen-Based Securities Trading: Pricing the Stock and Pricing the Trade," *Management Science* (43:12), pp. 1693-1708.

Da, Z., Engelberg, J., and Gao, P. 2011. "In Search of Attention," *The Journal of Finance* (66:5), pp. 1461-1499.

Dewan, S., and Mendelson, H. 1998. "Information Technology and Time-Based Competition in Financial Markets," *Management Science* (44:5), pp. 595-609.

Farrell, M., Green, T. C., Jame, R., and Markov, S. 2018. "The Democratization of Investment Research: Implications for Retail Investor Profitability and Firm Liquidity." SSRN 3222841.

Fed, R. 2018. "Financial Accounts of the United States: Flow of Funds, Balance Sheets, and Integrated Macroeconomic Accounts, Third Quarter 2018."

Greene, J., and Smart, S. 1999. "Liquidity Provision and Noise Trading: Evidence from the “Investment Dartboard” Column," *The Journal of Finance* (54:5), pp. 1885-1899.

Grossman, S. J., and Stiglitz, J. E. 1980. "On the Impossibility of Informationally Efficient Markets," *The American Economic Review* (70:3), pp. 393-408.

Han, B., Tang, Y., and Yang, L. 2016. "Public Information and Uninformed Trading: Implications for Market Liquidity and Price Efficiency," *Journal of Economic Theory* (163), pp. 604-643.

Ivković, Z., Sialm, C., and Weisbenner, S. 2008. "Portfolio Concentration and the Performance of Individual Investors," *Journal of Financial and Quantitative Analysis* (43:3), pp. 613-655.

Jain, H., Padmanabhan, B., Pavlou, P. A., and Santanam, R. T. 2018. "Call for Papers—Special Issue of Information Systems Research—Humans, Algorithms, and Augmented Intelligence: The Future of Work, Organizations, and Society," *Information Systems Research* (29:1), pp. 250-251.

Kelley, E. K., and Tetlock, P. C. 2013. "How Wise Are Crowds? Insights from Retail Orders and Stock Returns," *The Journal of Finance* (68:3), pp. 1229-1265.

Kelley, E. K., and Tetlock, P. C. 2016. "Retail Short Selling and Stock Prices," *The Review of Financial Studies* (30:3), pp. 801-834.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. 2017. "Human Decisions and Machine Predictions," *The Quarterly Journal of Economics* (133:1), pp. 237-293.

Lawrence, A., Ryans, J., Sun, E., and Laptev, N. 2016. "Yahoo Finance Search and Earnings Announcements." SSRN 2804353.

Meyer, G., Adomavicius, G., Johnson, P. E., Elidrisi, M., Rush, W. A., Sperl-Hillen, J. M., and O'Connor, P. J. 2014. "A Machine Learning Approach to Improving Dynamic Decision Making," *Information Systems Research* (25:2), pp. 239-263.

Miklós-Thal, J., and Tucker, C. 2019. "Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination between Sellers?," *Management Science* (Forthcoming).

Mullainathan, S., and Spiess, J. 2017. "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives* (31:2), pp. 87-106.

Newey, W. K., and West, K. D. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelationconsistent Covariance Matrix," *Econometrica* (55:3), pp. 703-708.

Zhang, X. M., and Zhang, L. 2015. "How Does the Internet Affect the Financial Market? An Equilibrium Model of Internet-Facilitated Feedback Trading," *MIS Quarterly* (39:1).

Figure 1. Worldwide Google Search Volume Index on Finance APIs

This figure depicts the worldwide weekly Google search volume index from 2014 to 2019 for the five popular Finance APIs. The highest search volume in the period is assigned the score 100. The search volume for Yahoo Finance API is represented by the purple line.

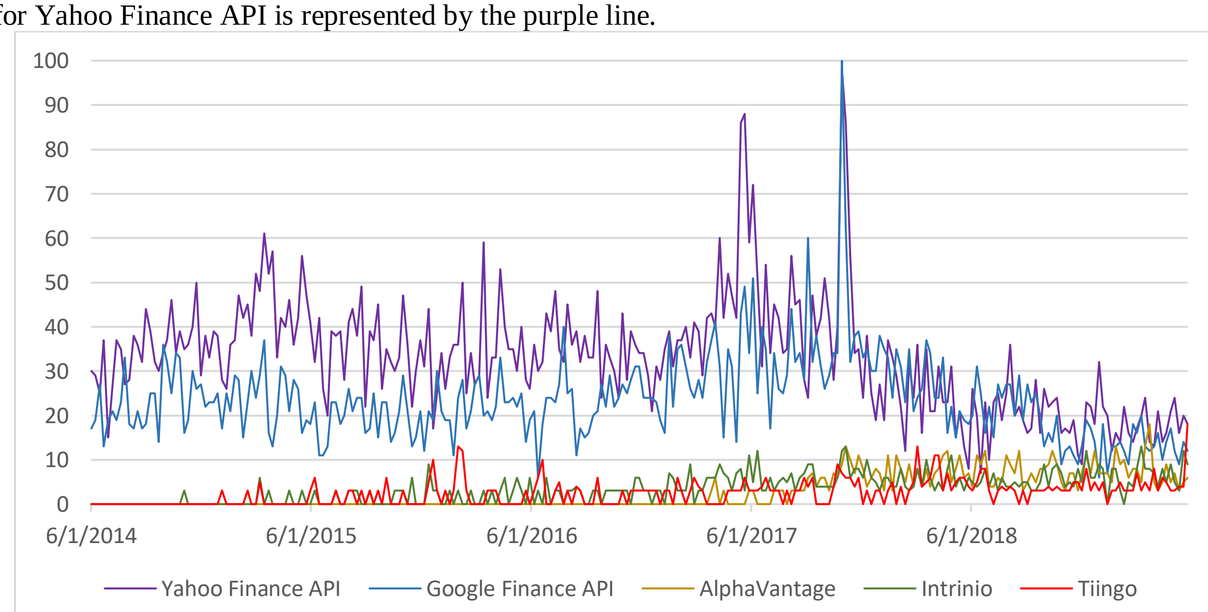

Figure 2: Repository Finance API Results on Github
This figure depicts the number of repository search results from Github.com for the five popular Finance APIs on May 24, 2019.

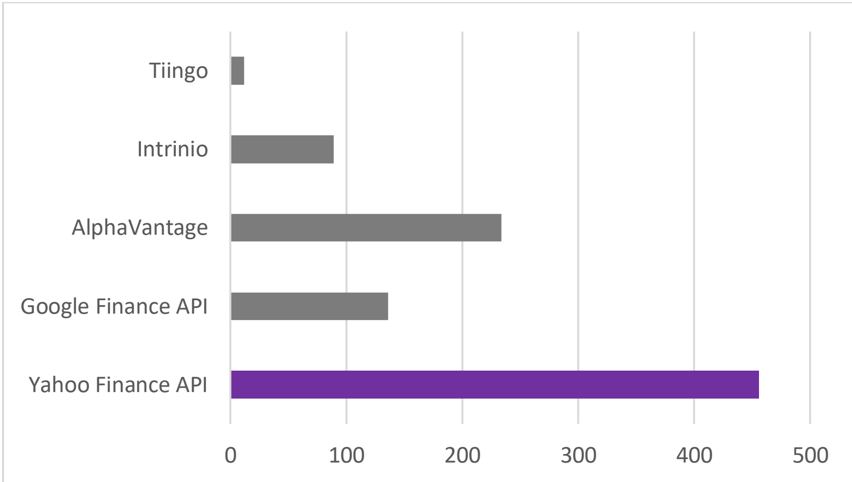

Figure 3. Retail trades around the shutdown of Yahoo! Finance API
This figure plots daily retail trading volume around the shutdown of Yahoo! Finance API for firms with below-median (solid line) or above-median institutional holdings (dashed line). Retail trades are identified from TAQ, following Boehmer et al. (2017). The y-axis is retail trading volume scaled by total shares outstanding, multiplied by 100. The vertical dashed line indicates the shut-down of Yahoo! Finance API.

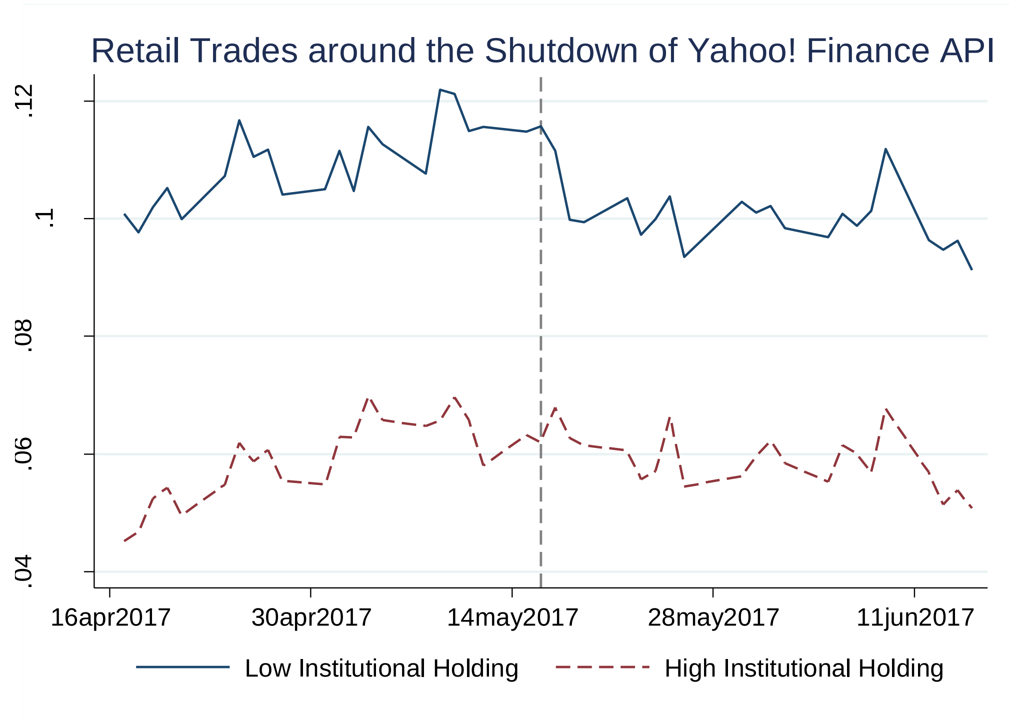

Table 1. Summary statistics

Panel A of this table reports the summary statistics of the key variables used in the main sample of this study (two-month window centered around the shutdown of Yahoo! Finance API). Each observation is a firm-trading day for daily measures. Firm characteristics are measured as of the most recent fiscal year before the sample starting date. Panel B presents the univariate comparisons for daily measures for firms with below- or above- median institutional holdings (*Low_IH*=1 or *Low_IH*=0) around the shutdown of Yahoo! Finance API. *Pre* and *Post* indicate the sub-periods before and after the shutdown, respectively. See Appendix A for detailed variable definitions. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Summary statistics

	N	Mean	S.D.	P25	P50	P75
Daily Measures						
Retail_Vol	192299	0.0839	0.2025	0.0102	0.0251	0.0652
Institutional_Vol	192299	0.0828	0.1720	0.0000	0.0198	0.0841
Ab_Retail_Vol	192218	0.0194	0.1319	-0.0085	0.0000	0.0135
Ab_Institutional_Vol	192218	0.0319	0.1354	-0.0092	0.0000	0.0248
AIM	191239	0.0768	0.2557	0.0002	0.0015	0.0175
Spread	192252	0.5230	1.0204	0.0375	0.1172	0.4662
Ret	192253	0.0003	0.0234	-0.0096	0.0000	0.0106
|Ret|	192253	0.0162	0.0188	0.0042	0.0101	0.0208
News	175915	0.2242	0.5336	0.0000	0.0000	0.0000
Firm Characteristics						
Size	4481	6.8406	2.0972	5.3315	6.8342	8.2690
ROA	4481	-0.0711	0.2906	-0.0448	0.0123	0.0516
Loss	4481	0.3522	0.4777	0.0000	0.0000	1.0000
R&D	4481	0.0659	0.1577	0.0000	0.0000	0.0458
Advertising	4481	0.0086	0.0246	0.0000	0.0000	0.0027
Leverage	4481	0.2620	0.2425	0.0477	0.2214	0.4089
Analysts	4481	0.9728	0.9385	0.0000	0.6931	1.6094

Panel B. Univariate Comparison

	Low_IH=1			Low_IH=0			
Variables	Pre	Post	MeanDiff	Pre	Post	MeanDiff	DID
Retail_Vol	0.112	0.103	-0.009***	0.060	0.060	0.000	-0.009***
Institutional_Vol	0.053	0.049	-0.004***	0.118	0.112	-0.006***	0.002
Ab_Retail_Vol	0.026	0.021	-0.006***	0.015	0.015	0.000	-0.005**
Ab_Institutional_Vol	0.026	0.022	-0.004***	0.044	0.037	-0.007***	0.003
AIM	0.143	0.152	0.010***	0.005	0.006	0.001**	0.009***
Spread	0.900	0.927	0.027***	0.126	0.128	0.003	0.025***
Ret	0.000	-0.001	-0.001***	0.002	0.000	-0.001***	0.000
|Ret|	0.019	0.019	0.000	0.014	0.013	0.000	0.000*
News	0.194	0.152	-0.042***	0.298	0.249	-0.050***	0.008

Table 2. Retail versus Institutional trades around the shutdown of Yahoo! Finance API

This table reports the regressions results of retail or institutional trading volume around the shutdown of Yahoo! Finance API (May 16, 2017). The sample period is two-, four- or six-month window (indicated in the table header) centered around May 16, 2017. Retail and institutional trades are identified from TAQ, following Boehmer et al. (2017) and Bushee et al. (2019). The dependent variable is abnormal retail trading volume (*Ab_Retail_Vol*) in Panel A and abnormal institutional trading volume (*Ab_Institutional_Vol*) in Panel B. The key variable of interest is the interaction between *Post* (indicating the period after the shutdown of Yahoo! Finance API) and *Low_IH* (indicating firms with below-median percentage of institutional holdings). See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Retail Trades

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**2-Month Window**		**4-Month Window**		**6-Month Window**	
Post×Low_IH	-0.008***	-0.007***	-0.004	-0.003	-0.002	-0.002
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Ret	0.164***	0.192***	0.198***	0.221***	0.206***	0.220***
	(0.031)	(0.028)	(0.023)	(0.022)	(0.021)	(0.019)
|Ret|	2.803***	2.590***	2.807***	2.710***	2.833***	2.762***
	(0.081)	(0.064)	(0.075)	(0.065)	(0.069)	(0.061)
News	0.021***	0.025***	0.021***	0.026***	0.020***	0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Low_IH	0.003		0.001		0.000	
	(0.002)		(0.002)		(0.001)	
Size	-0.000		0.000		-0.000	
	(0.001)		(0.001)		(0.001)	
ROA	-0.010		-0.022***		-0.020***	
	(0.010)		(0.008)		(0.006)	
Loss	-0.007***		-0.009***		-0.009***	
	(0.003)		(0.002)		(0.002)	
R&D	-0.030*		-0.023*		-0.010	
	(0.016)		(0.013)		(0.010)	
Advertising	0.189***		0.089**		0.079***	
	(0.057)		(0.036)		(0.030)	
Leverage	0.002		0.001		0.005*	
	(0.004)		(0.003)		(0.003)	
Analysts	0.001		0.001		0.001	
	(0.001)		(0.001)		(0.001)	
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	175,611	175,611	342,624	342,624	509,195	509,195
R-squared	0.171	0.331	0.171	0.254	0.177	0.237

Table 2. Panel B: Institutional Trades

VARIABLES	(1) 2-Month Window	(2)	(3) 4-Month Window	(4)	(5) 6-Month Window	(6)
Post×Low_IH	0.003	0.003	-0.002	-0.001	-0.002*	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Ret	0.088***	0.053***	0.109***	0.083***	0.103***	0.079***
	(0.022)	(0.020)	(0.017)	(0.015)	(0.015)	(0.014)
|Ret|	1.579***	1.618***	1.480***	1.535***	1.586***	1.628***
	(0.052)	(0.046)	(0.045)	(0.041)	(0.044)	(0.039)
News	0.029***	0.037***	0.027***	0.037***	0.028***	0.037***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Low_IH	-0.016***		-0.015***		-0.017***	
	(0.002)		(0.002)		(0.001)	
Size	0.002***		0.001		0.000	
	(0.001)		(0.001)		(0.000)	
ROA	0.026***		0.022***		0.021***	
	(0.004)		(0.003)		(0.003)	
Loss	0.002		-0.001		-0.001	
	(0.002)		(0.002)		(0.001)	
R&D	0.013*		0.016***		0.012**	
	(0.007)		(0.006)		(0.006)	
Advertising	0.202***		0.092***		0.093***	
	(0.048)		(0.032)		(0.028)	
Leverage	0.004		0.003		0.006***	
	(0.003)		(0.002)		(0.002)	
Analysts	0.000		-0.000		0.000	
	(0.001)		(0.001)		(0.001)	
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	175,611	175,611	342,624	342,624	509,195	509,195
R-squared	0.076	0.193	0.081	0.144	0.084	0.136

Table 3. Market liquidity

This table reports the regressions results of market liquidity in a two-month window centered around the shutdown of Yahoo! Finance API. The dependent variable is daily Amihud's illiquidity measure (*AIM*) in Columns 1-2 and daily relative bid-ask spread (*Spread*) in Columns 3-4. The key variable of interest is the interaction between *Post* (indicating the period after the shutdown of Yahoo! Finance API) and *Low_IH* (indicating firms with below-median percentage of institutional holdings). See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	*AIM*		*Spread*	
Post×Low_IH	0.009***	0.009***	0.027***	0.026***
	(0.002)	(0.002)	(0.009)	(0.009)
Ret	-0.273***	-0.229***	-0.276**	0.127
	(0.035)	(0.025)	(0.117)	(0.089)
|Ret|	2.235***	2.121***	3.559***	1.500***
	(0.122)	(0.079)	(0.342)	(0.145)
News	-0.000	-0.016***	0.042***	-0.012***
	(0.002)	(0.001)	(0.009)	(0.002)
Low_IH	0.048***		0.293***	
	(0.004)		(0.020)	
Size	-0.037***		-0.208***	
	(0.002)		(0.009)	
ROA	-0.060**		-0.296***	
	(0.027)		(0.092)	
Loss	-0.000		0.032	
	(0.008)		(0.034)	
R&D	-0.140***		-0.416***	
	(0.041)		(0.143)	
Advertising	-0.115		-0.369	
	(0.117)		(0.459)	
Leverage	-0.030***		-0.090*	
	(0.011)		(0.046)	
Analysts	0.000		-0.006	
	(0.003)		(0.012)	
Date FE	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes
Observations	175,112	175,112	175,660	175,660
R-squared	0.210	0.625	0.335	0.752

Table 4. Falsification tests

This table reports the results of falsification tests in a two-month window centered around May 16, 2016, one year before the shutdown of Yahoo! Finance API. The dependent variable is abnormal retail trading volume (*Ab_Retail_Vol*) in Columns 1-2, abnormal institutional trading volume (*Ab_Institutional_Vol*) in Columns 3-4, daily Amihud's illiquidity measure (*AIM*) in Columns 5-6, and daily relative bid-ask spread (*Spread*) in Columns 7-8. The key variable of interest is the interaction between Post (indicating the period after the shutdown of Yahoo! Finance API) and *Low_IH* (indicating firms with below-median percentage of institutional holdings). See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	*Ab_Retail_Vol*		*Ab_Institutional_Vol*		*AIM*		*Spread*	
Post×Low_IH	0.001	0.001	0.002	0.002	-0.005	-0.005	0.003	0.006
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.011)	(0.011)
Ret	0.061***	0.066***	0.070***	0.056***	-0.275***	-0.227***	-0.351***	-0.077
	(0.017)	(0.016)	(0.018)	(0.017)	(0.042)	(0.030)	(0.124)	(0.103)
|Ret|	1.801***	1.744***	1.193***	1.221***	2.900***	2.562***	4.064***	1.546***
	(0.049)	(0.040)	(0.040)	(0.035)	(0.179)	(0.104)	(0.434)	(0.176)
News	0.019***	0.022***	0.024***	0.032***	0.013***	-0.021***	0.091***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.010)	(0.003)
Low_IH	-0.001		-0.012***		0.072***		0.333***	
	(0.001)		(0.002)		(0.006)		(0.022)	
Size	-0.001***		0.001*		-0.066***		-0.265***	
	(0.000)		(0.000)		(0.003)		(0.011)	
ROA	-0.020***		0.015***		-0.081**		-0.496***	
	(0.007)		(0.004)		(0.039)		(0.124)	
Loss	-0.013***		-0.001		-0.008		-0.047	
	(0.002)		(0.002)		(0.012)		(0.039)	
R&D	0.015		0.016*		-0.260***		-0.611***	
	(0.015)		(0.009)		(0.066)		(0.219)	
Advertising	0.070**		0.081***		0.151		-0.091	
	(0.032)		(0.030)		(0.197)		(0.612)	
Leverage	0.006*		0.013***		-0.014		-0.048	
	(0.003)		(0.003)		(0.018)		(0.065)	
Analysts	-0.001*		-0.002*		0.004		0.003	
	(0.001)		(0.001)		(0.004)		(0.015)	
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	162,087	162,087	162,087	162,087	161,333	161,333	162,095	162,095
R-squared	0.191	0.333	0.066	0.157	0.234	0.653	0.316	0.740

Table 5. Cumulative abnormal returns (2-month window)

This table reports the daily Fama-Macbeth regressions of future returns on abnormal retail buy and sell trading volume. The dependent variable is the buy-and-hold abnormal returns (buy-and-hold individual stock return minus the corresponding market return) for the next week starting from the next day in Columns 1-2, from week 2 to week 4 in Columns 3-4, from week 5 to week 8 in Columns 5-6, and from next day to the end of week 4 in Columns 7-10. The key variables of interest are abnormal retail buys (sells), which is the retail buy (sell) trading volume scaled by shares outstanding, minus its median value over the last 10 weeks. The sample includes firm-day observations during 2-month window centered around the shutdown of Yahoo! Finance API (from April 16, 2017 to June 15, 2017). In Columns 7-8 (9-10), only firms with above-median (below-median) institutional ownership are included. Newey and West (1987) standard errors with lags of two are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	CAR[1W] & All		CAR[2W,4W] & All		CAR[5W,8W] & All		CAR[1W,4W] & Low_IH=0		CAR[1W,4W] & Low_IH=1	
VARIABLES	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Ab_Retail_Buy	0.071	1.616*	-0.557	2.123	-1.219	0.354	6.695*	-1.525	-1.101	6.251***
	(0.632)	(0.810)	(1.406)	(1.503)	(1.630)	(1.315)	(3.586)	(2.720)	(1.745)	(1.350)
Ab_Retail_Sell	-2.468***	-3.195***	-1.019	-3.373***	-2.588	0.942	-2.724	-4.332	-5.295***	-7.683***
	(0.606)	(0.578)	(1.319)	(1.170)	(2.406)	(1.387)	(3.264)	(3.844)	(1.168)	(1.263)
|Ret[0]|	-8.718**	-10.960***	-15.462***	-1.381	-5.943	-13.362*	-14.965***	-1.795	-29.489***	-11.001*
	(3.473)	(2.710)	(4.161)	(5.549)	(8.325)	(7.133)	(3.940)	(10.697)	(4.455)	(5.763)
News	0.063	0.065	0.171	0.023	0.211***	-0.001	0.266**	0.322***	0.156	-0.004
	(0.058)	(0.094)	(0.101)	(0.051)	(0.064)	(0.068)	(0.120)	(0.090)	(0.126)	(0.093)
Size	0.069**	0.060	0.258***	-0.049	0.027	0.388***	0.221**	-0.455***	0.267***	0.138**
	(0.032)	(0.041)	(0.038)	(0.076)	(0.075)	(0.137)	(0.105)	(0.087)	(0.027)	(0.054)
BTM	-0.131	-0.011	-0.060	-0.018	0.046	0.513***	-1.510***	-0.003	0.223***	-0.019
	(0.112)	(0.144)	(0.099)	(0.078)	(0.142)	(0.180)	(0.253)	(0.291)	(0.073)	(0.110)
Ret[0]	3.766	-6.856**	4.781*	-3.853	11.709*	2.042	12.809**	-2.900	6.978*	-14.299***
	(2.747)	(2.911)	(2.650)	(4.687)	(6.502)	(2.702)	(5.445)	(8.603)	(3.818)	(4.665)
CAR[-1W]	0.024	-0.013	0.068**	-0.024*	0.046	0.029*	0.059*	-0.057***	0.107***	-0.034*
	(0.015)	(0.013)	(0.025)	(0.013)	(0.052)	(0.015)	(0.031)	(0.019)	(0.023)	(0.017)
CAR[-2W,-4W]	0.010	0.011	0.089***	-0.009	-0.015	0.048***	0.109***	-0.031	0.088***	0.012
	(0.008)	(0.016)	(0.009)	(0.009)	(0.012)	(0.009)	(0.016)	(0.026)	(0.017)	(0.014)
						Ab_Retail_Buy (Col 10-Col 9) – (Col 8-Col 7):				**15.572*** **
						Ab_Retail_Sell (Col 10-Col 9) – (Col 8-Col 7):				**-0.780**
Observations	85,599	89,487	85,564	89,275	85,425	89,024	44,268	46,375	41,296	42,900
R-squared	0.017	0.028	0.025	0.014	0.018	0.024	0.032	0.037	0.023	0.014

Table 6. Cumulative abnormal returns (4-month and 6-month windows)

This table reports the daily Fama-Macbeth regressions of future returns on abnormal retail buy and sell trading volume, respectively. The dependent variable is the buy-and-hold abnormal returns (buy-and-hold individual stock return minus the corresponding market return) from week 1 to week 4 The key variables of interest are abnormal retail buys (sells), which is the retail buy (sell) trading volume scaled by shares outstanding, minus its median value over the last 10 weeks. The sample includes firm-day observations during 4-month (Column 1-4) and 6-month window (Column 5-8) centered around the shutdown of Yahoo! Finance API (May 16, 2017). In Columns 1, 2, 5, 6 (3, 4, 7, 8), only firms with above-median (below-median) institutional ownership are included. Newey and West (1987) standard errors with lags of two are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	4-Month Window				**6-Month Window**			
	CAR[1W,4W] & Low_IH=0		CAR[1W,4W] & Low_IH=1		CAR[1W,4W] & Low_IH=0		CAR[1W,4W] & Low_IH=1	
VARIABLES	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Ab_Retail_Buy	3.306	-1.655	0.419	2.194	3.370*	-0.945	0.624	0.190
	(2.319)	(2.585)	(1.488)	(1.591)	(1.833)	(1.948)	(1.312)	(1.375)
Ab_Retail_Sell	0.292	-2.738	-5.533***	-7.256***	1.016	-1.594	-4.750***	-6.963***
	(2.161)	(3.174)	(0.898)	(1.087)	(1.715)	(2.172)	(0.957)	(1.000)
|Ret[0]|	-36.660***	2.812	-41.713***	-18.573***	-32.929***	-2.100	-38.625***	-12.973**
	(7.409)	(8.839)	(4.097)	(5.955)	(5.500)	(8.762)	(3.266)	(6.448)
News	0.261***	0.287***	-0.109	-0.149*	0.140	0.184***	-0.156	-0.099
	(0.095)	(0.058)	(0.132)	(0.088)	(0.086)	(0.064)	(0.100)	(0.065)
Size	-0.056	-0.080	0.277***	0.361***	-0.044	0.025	0.239***	0.349***
	(0.106)	(0.146)	(0.031)	(0.091)	(0.078)	(0.122)	(0.034)	(0.076)
BTM	-1.344***	0.103	0.277***	0.373**	-1.256***	-0.106	0.457***	0.323**
	(0.261)	(0.174)	(0.099)	(0.145)	(0.237)	(0.161)	(0.125)	(0.133)
Ret[0]	1.698	-5.049	2.858	-9.831***	2.477	-1.979	1.063	-7.470***
	(4.872)	(5.635)	(2.883)	(3.160)	(3.465)	(4.431)	(2.254)	(2.472)
CAR[-1W]	0.018	-0.078***	0.081***	-0.022	0.018	-0.050***	0.079***	-0.012
	(0.024)	(0.019)	(0.017)	(0.016)	(0.020)	(0.017)	(0.014)	(0.014)
CAR[-2W,-4W]	0.058***	-0.040***	0.079***	-0.011	0.055***	-0.024**	0.070***	-0.018**
	(0.018)	(0.014)	(0.013)	(0.012)	(0.017)	(0.011)	(0.011)	(0.008)
Ab_Retail_Buy		**(Col 4-Col3) – (Col 2-Col 1):**		**6.736**	**(Col 8-Col 7) – (Col 6-Col 5):**			**3.881**
Ab_Retail_Sell		**(Col 4-Col3) – (Col 2-Col 1):**		**1.307**	**(Col 8-Col 7) – (Col 6-Col 5):**			**0.397**
Observations	88,536	88,509	82,624	81,514	128,588	134,770	120,036	123,636
R-squared	0.030	0.036	0.028	0.026	0.029	0.034	0.025	0.024

Appendices

Variable definitions are provided in **Table A1**. As a robustness test, we assess whether the results are sensitive to the control and treatment group classification. We replicate the main analyses while excluding sample firms whose institutional holdings fall in the middle 20 percent. Effectively, we changed the treatment and control groups from those with below and above median institutional holdings to 0-40% and 60%-100%, respectively. The results in **Table A2** are consistent with that of the main analyses (Table 2, 3).

Moreover, since our earlier analysis (Table 5) indicates that the profitability of retail trades relatively increases after the shutdown of the API, presumably due to the shutdown's immediate deterring impact on less savvy traders, we seek to further test if those results are due to the similarity in investment targets that both institutional traders and savvy retail traders (those who continue making trades in the immediate aftermath of the shutdown) pursue. Therefore, we seek to eliminate targets of investment that are dissimilar across the institutional and retail traders, by only retaining the targets that find a match across the retail or institutional targets. Using a coarsened exact matching (CEM) procedure based on firm size, ROA, loss, financial leverage, R&D and advertising expenditures, the number of analysts following the company (see Table A1), we see that the DID coefficient becomes insignificant both in trading volume and market liquidity estimations (see **Table A3**). This non-result indicates that the shutdown does not have an impact on the volume and size of trades made by retail investors who behave similar to institutional traders.

We further differentiate retail and institutional trading volumes in the main analyses (Table 2) into buy and sell volumes for retail and institutional trades. Abnormal retail buys (sells) is the retail buy (sell) trading volume scaled by shares outstanding, and minus its median value over the last 10 weeks. Institutional buy and sell volumes are similarly defined. Using the same predictors as in

Table 2, we found that both retail buys and sells drop by a similar magnitude after the shutdown of Yahoo! Finance API (**Table A4**). The impact on institutional buy and sell volumes are insignificant.

Table A1. Variable definitions

Variables	**Definitions**	**Data Source**
Retail_Vol	Shares of trades initiated by retail investors, scaled by total shares outstanding and multiplied by 100. Retail trades are identified based on TAQ exchange code (D) and a small price improvement (0-0.4 cents, exclusive, above (below) a round cent for sale (buy) transactions), following Boehmer et al. (2017).	TAQ & CRSP
Institutional_Vol	Shares of trades initiated by institutional investors, scaled by total shares outstanding and multiplied by 100. Institutional trades are non-retail trades with trade size above $50,000, following Bushee et al. (2019).	TAQ & CRSP
Ab_Retail_Vol	*Retail_Vol* minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
Ab_Institutional_Vol	*Institutional_Vol* minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
Ab_Retail_Buy	Shares of trades bought by retail investors (scaled by total shares outstanding and multiplied by 100) minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
Ab_Retail_Sell	Shares of trades sold by retail investors (scaled by total shares outstanding and multiplied by 100) minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
Ab_ Institutional_Buy	Shares of trades bought by institutional investors (scaled by total shares outstanding and multiplied by 100) minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
Ab_ Institutional_Sell	Shares of trades sold by institutional investors (scaled by total shares outstanding and multiplied by 100) minus its median for the same day of the week over the past 10 week.	TAQ & CRSP
AIM	Amihud(2002) illiquidity measure, the natural logarithm of the ratio of absolute stock return to dollar volume [$1{,}000{,}000 \times \mid ret \mid \div (prc \times vol)$]	CRSP
Spread	Daily bid-ask spread based on CRSP data, $100 \times (ask - bid)/[(ask + bid)/2]$.	CRSP
Ret	Delist adjusted stock returns.	CRSP
|Ret|	Absolute value of *Ret*.	CRSP
News	Natural logarithm of one plus the number of news articles on the Dow Jones Edition of RavenPack with relevance store above 20 (the company name can be identified somewhere in the story).	RavenPack
Size	Natural logarithm of market capitalization at the fiscal year end.	Compustat
ROA	Return on assets (*ib/at*).	Compustat
Loss	Dummy variable, one if *ROA*<0	Compustat
R&D	R&D intensity (*xrd/at*).	Compustat
Advertising	Advertising intensity (*xad/at*).	Compustat
Leverage	Financial leverage ((*dltt+dlc)/at*).	Compustat
Analysts	Natural logarithm of one plus the number of financial analysts following the company.	IBES
IH	Institutional holdings, shares owned by Institutional investors scaled by total shares outstanding.	Thomson Reuters
BTM	Book to market ratio.	Compustat & CRSP

Table A2: Main analyses with different sample construction: exclude the middle 20%

Panel A, B, and C of this table are robustness checks for Table 2 Panel A and B, and Table 3, respectively. The regression specifications in this table are exactly the same as before. The only difference is that firms whose institutional holdings fall in the middle 20 percent were excluded here from the sample. See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Retail trades

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	2-Month Window		4-Month Window		6-Month Window	
Post×Low_IH	-0.010***	-0.010***	-0.004	-0.004	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Ret	0.184***	0.221***	0.226***	0.256***	0.234***	0.253***
	(0.035)	(0.032)	(0.026)	(0.024)	(0.024)	(0.022)
|Ret|	2.874***	2.647***	2.882***	2.777***	2.908***	2.832***
	(0.090)	(0.072)	(0.084)	(0.073)	(0.078)	(0.068)
News	0.024***	0.027***	0.024***	0.028***	0.022***	0.026***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Low_IH	0.003		-0.000		-0.001	
	(0.003)		(0.002)		(0.002)	
Size	-0.001		-0.001		-0.001	
	(0.001)		(0.001)		(0.001)	
ROA	-0.011		-0.023**		-0.021***	
	(0.011)		(0.009)		(0.007)	
Loss	-0.009***		-0.010***		-0.011***	
	(0.003)		(0.002)		(0.002)	
R&D	-0.029*		-0.026*		-0.012	
	(0.017)		(0.014)		(0.011)	
Advertising	0.211***		0.096**		0.087**	
	(0.068)		(0.042)		(0.036)	
Leverage	0.001		0.000		0.004	
	(0.005)		(0.004)		(0.003)	
Analysts	0.002		0.001		0.002*	
	(0.002)		(0.001)		(0.001)	
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	139,362	139,362	271,849	271,849	404,017	404,017
R-squared	0.174	0.334	0.174	0.257	0.180	0.240

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2. Panel B: Institutional trades

VARIABLES	(1) 2-Month Window	(2)	(3) 4-Month Window	(4)	(5) 6-Month Window	(6)
Post×Low_IH	0.002	0.003	-0.002	-0.002	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Ret	0.062***	0.030	0.088***	0.065***	0.084***	0.064***
	(0.024)	(0.022)	(0.018)	(0.017)	(0.016)	(0.015)
|Ret|	1.524***	1.566***	1.420***	1.477***	1.531***	1.572***
	(0.056)	(0.050)	(0.048)	(0.044)	(0.047)	(0.042)
News	0.032***	0.039***	0.030***	0.038***	0.031***	0.039***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Low_IH	-0.018***		-0.018***		-0.021***	
	(0.003)		(0.002)		(0.002)	
Size	0.002***		0.001		0.000	
	(0.001)		(0.001)		(0.001)	
ROA	0.025***		0.020***		0.020***	
	(0.004)		(0.003)		(0.004)	
Loss	0.002		-0.000		-0.001	
	(0.002)		(0.002)		(0.002)	
R&D	0.018**		0.016***		0.013**	
	(0.008)		(0.006)		(0.006)	
Advertising	0.183***		0.081**		0.083***	
	(0.056)		(0.037)		(0.032)	
Leverage	0.000		0.001		0.003	
	(0.003)		(0.002)		(0.002)	
Analysts	0.000		-0.000		0.000	
	(0.002)		(0.001)		(0.001)	
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	139,362	139,362	271,849	271,849	404,017	404,017
R-squared	0.078	0.194	0.080	0.141	0.084	0.133

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2. Panel C: Market liquidity

	(1)	(2)	(3)	(4)
VARIABLES	AIM		Spread	
Post×Low_IH	0.011***	0.011***	0.035***	0.033***
	(0.003)	(0.003)	(0.012)	(0.012)
|Ret|	-0.289***	-0.250***	-0.202	0.164
	(0.040)	(0.029)	(0.136)	(0.106)
News	2.442***	2.340***	3.844***	1.689***
	(0.140)	(0.091)	(0.389)	(0.170)
Low_IH	-0.007**	-0.019***	0.018*	-0.015***
	(0.003)	(0.001)	(0.010)	(0.003)
Ret	0.048***		0.325***	
	(0.005)		(0.026)	
Size	-0.043***		-0.226***	
	(0.003)		(0.011)	
ROA	-0.041		-0.253**	
	(0.030)		(0.101)	
Loss	0.004		0.045	
	(0.010)		(0.039)	
R&D	-0.135***		-0.424***	
	(0.047)		(0.159)	
Advertising	-0.079		-0.070	
	(0.145)		(0.559)	
Leverage	-0.033**		-0.098*	
	(0.013)		(0.054)	
Analysts	-0.001		-0.012	
	(0.003)		(0.015)	
Date FE	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes
Observations	138,879	138,879	139,411	139,411
R-squared	0.229	0.624	0.356	0.746

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3. Main analyses after coarsened exact matching

Panel A and B of this table are robustness checks for Table 2 and 3, respectively. The 4- and 6-month window results (insignificant) are not reported for brevity. The regression specifications are the same as in the main analyses. The only difference is that the analyses here were conducted after coarsened exact matching based on firm characteristics (*Size* ~ *Analysts*) as of the beginning of the sample period. See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Retail and institutional trades

	(1)	(2)	(3)	(4)
VARIABLES	Retail trading volume (2-month)		Institutional trading volume (2-month)	
Post×Low_IH	-0.001	-0.001	0.004	0.005
	(0.003)	(0.003)	(0.003)	(0.003)
Ret	-0.030	-0.016	0.067	0.049
	(0.051)	(0.051)	(0.043)	(0.040)
|Ret|	1.584***	1.590***	1.247***	1.367***
	(0.110)	(0.108)	(0.096)	(0.091)
News	0.016***	0.015***	0.024***	0.024***
	(0.002)	(0.002)	(0.003)	(0.002)
Low_IH	-0.000		-0.011***	
	(0.003)		(0.003)	
Size	0.002		0.001	
	(0.001)		(0.001)	
ROA	-0.000		0.041**	
	(0.027)		(0.019)	
Loss	-0.001		0.000	
	(0.004)		(0.004)	
R&D	0.038		0.063	
	(0.038)		(0.039)	
Advertising	-0.227**		-0.200	
	(0.113)		(0.133)	
Leverage	0.004		0.007	
	(0.005)		(0.006)	
Analysts	-0.001		-0.001	
	(0.002)		(0.002)	
Date FE	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes
Observations	68,965	68,965	68,965	68,965
R-squared	0.144	0.294	0.057	0.163

Table A3. Panel B: Market liquidity

	(1)	(2)	(3)	(4)
VARIABLES	*AIM*		*Spread*	
Post*Low_IH	-0.001	-0.002	-0.004	-0.004
	(0.004)	(0.004)	(0.013)	(0.013)
Ret	-0.155**	-0.109***	-0.246	-0.063
	(0.068)	(0.040)	(0.200)	(0.130)
|Ret|	1.786***	1.564***	3.979***	1.321***
	(0.228)	(0.149)	(0.869)	(0.231)
News	-0.009***	-0.010***	-0.001	-0.002
	(0.003)	(0.001)	(0.013)	(0.003)
Low_IH	0.034***		0.234***	
	(0.008)		(0.041)	
Size	-0.027***		-0.175***	
	(0.004)		(0.015)	
ROA	-0.094		-0.242	
	(0.061)		(0.353)	
Loss	0.024		0.077	
	(0.015)		(0.077)	
R&D	-0.159		0.385	
	(0.173)		(1.185)	
Advertising	-0.048		0.941	
	(0.279)		(1.643)	
Leverage	-0.014		-0.021	
	(0.017)		(0.095)	
Analysts	0.005		-0.002	
	(0.004)		(0.019)	
Date FE	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes
Observations	68,747	68,747	68,988	68,988
R-squared	0.148	0.600	0.227	0.792

Table A4: Buy versus sell trades

This table reports the impact of Yahoo! Finance API shutdown on buy and sell volumes, respectively. The sample period is the two-month window centered around May 16, 2017. The dependent variables are abnormal retail trading buy and sell (*Ab_Retail_Buy*, *Ab_Retail_Sell*) and abnormal institutional buy and sell (*Ab_Institutional_Buy*, *Ab_Institutional_Sell*). See Table A1 for detailed variable definitions. Robust standard errors clustered by firm are reported in the parentheses. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) Ab_Retail_ Buy	(2) Ab_Retail_ Sell	(3) Ab_ Institutional _Buy	(4) Ab_ Institutional _Sell	(5) Ab_Retail_ Buy	(6) Ab_Retail_ Sell	(7) Ab_ Institutional _Buy	(8) Ab_ Institutional _Sell
Post×Low_IH	-0.004***	-0.004***	0.001	0.001	-0.003***	-0.004***	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ret	0.147***	0.025*	0.078***	0.014	0.163***	0.038***	0.060***	-0.005
	(0.018)	(0.015)	(0.012)	(0.012)	(0.017)	(0.014)	(0.011)	(0.011)
|Ret|	1.432***	1.378***	0.776***	0.785***	1.320***	1.275***	0.784***	0.809***
	(0.042)	(0.040)	(0.026)	(0.028)	(0.033)	(0.033)	(0.024)	(0.024)
News	0.010***	0.010***	0.014***	0.015***	0.012***	0.013***	0.018***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Low_IH	0.002	0.001	-0.010***	-0.009***				
	(0.001)	(0.001)	(0.001)	(0.001)				
Size	-0.000	-0.001	0.001***	0.001***				
	(0.000)	(0.000)	(0.000)	(0.000)				
ROA	-0.008	-0.003	0.013***	0.014***				
	(0.005)	(0.005)	(0.002)	(0.002)				
Loss	-0.004***	-0.003***	0.001	0.001				
	(0.001)	(0.001)	(0.001)	(0.001)				
R&D	-0.016**	-0.016**	0.007*	0.007*				
	(0.008)	(0.008)	(0.004)	(0.004)				
Advertising	0.099***	0.090***	0.104***	0.112***				
	(0.030)	(0.028)	(0.024)	(0.027)				
Leverage	0.002	0.001	0.003**	0.002				
	(0.002)	(0.002)	(0.002)	(0.002)				
Analysts	0.001	0.001	0.000	0.000				
	(0.001)	(0.001)	(0.001)	(0.001)				
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	No	No	No	No
Firm FE	No	No	No	No	Yes	Yes	Yes	Yes
Observations	175,663	175,663	175,663	175,663	175,663	175,663	175,663	175,663
R-squared	0.163	0.157	0.070	0.064	0.316	0.308	0.176	0.172