INSPER - INSTITUTE OF EDUCATION AND RESEARCH
DOCTORAL PROGRAM IN BUSINESS ECONOMICS

GUILHERME LEITE PAIVA

Essays on Individual Investor Behavior

São Paulo

2022

GUILHERME LEITE PAIVA

Essays on Individual Investor Behavior

Dissertation submitted in fulfillment of the requirements for the degree of Doctor in Economics at Insper Institute of Education and Research.
Concentration: Business Economics.
Field: Finance.

Supervisor: Marco Bonomo
Co-supervisor: Ruy M. Ribeiro

São Paulo

2022

Acknowledgements

To be written.

Abstract

PAIVA, G. L. (2022) **Essays on Individual Investor Behavior**. Doctoral Dissertation - Insper, São Paulo 2021.

This dissertation is composed of three self-contained papers on individual investor behavior. The first paper studies the relation between the salience of a stock and retail investors' trading decisions. The study documents that living in a small city that has a local store of a brick-and-mortar firm more than doubles the chances of individuals picking the stock of that firm to day-trade. The second paper provides a life like monetary measure of gains and losses for all investors participating in the Brazilian stock and derivatives market. The paper investigates how the stock selection ability of individuals correlates with their wealth and what are the implications for the redistribution of wealth through the stock market. The study finds that investors with lower wealth do worst, leading to an increase in wealth concentration. The third and final paper examine how social media might affect the trading activity of investors. The study found that trading activity reduces when investor's interaction with social media platform is impaired. Investors shunning away from the markets when they cannot access social media platforms supports the hypothesis that social media are vehicles of information about the market, with the potential to reduce informational frictions and improve market liquidity.

Keywords: retail investor; investor activity; limited-attention; salience; day-trade; stock market; derivatives market; wealth redistribution; return heterogeneity; contrarian behavior; social media; information diffusion.

Resumo

PAIVA, G. L. (2022) **Ensaios sobre o comportamento do investidor pessoa física**. Tese (Doutorado) - Insper, São Paulo 2021.

Esta dissertação é composta por três artigos independentes sobre o comportamento do investidor individual. O primeiro artigo estuda a relação entre a saliência de uma ação e as decisões de investimento dos investidores pessoa física. O estudo documenta que morar em uma cidade pequena que possui uma loja física de uma empresa listada na bolsa de valores mais do que dobra as chances de os indivíduos escolherem ações dessa empresa para *day-trade*. O segundo artigo fornece uma medida monetária fidedigna de lucros e prejuízos para todos os investidores que participam do mercado brasileiro de ações e de derivativos. O artigo investiga como o desempenho dos indivíduos se correlaciona com a sua riqueza e quais são as implicações desse desempenho para a redistribuição da riqueza por meio do mercado acionário. O estudo constata que os investidores com menor riqueza se saem pior, levando a um aumento na concentração de riqueza. O terceiro e último artigo examina como a mídia social pode afetar a atividade dos investidores no mercado acionário. O estudo descobriu que a atividade de negociação reduz quando a interação do investidor com as mídias sociais é prejudicada. Investidores negociando menos quando não podem acessar as mídias sociais corroboram a hipótese de que as mídias sociais são veículos de informações sobre o mercado, com potencial para reduzir fricções de informação e melhorar a liquidez do mercado.

Palavras-chave: investidor pessoa física; atividade do investidor; atenção limitada; saliência; *day-trade*; mercado acionário; mercado de derivativos; redistribuição de riqueza; heterogeneidade no retorno; comportamento contrário; mídia social; difusão de informações.

List of Figures

FIGURES - Chapter 2

FIGURES - Chapter 3

FIGURES - Chapter 4

APPENDIX A – Chapter 3

APPENDIX B – Chapter 4

List of Tables

TABLES - Chapter 2

TABLES - Chapter 3

TABLES - Chapter 4

APPENDIX A – Chapter 3

APPENDIX A – Chapter 4

Contents

1. General Introduction.

This doctoral dissertation is composed of three papers on individual investor behavior a research area of the household finance field. Although each article deals with different problems and contexts, all three papers share the same main data, a highly detailed administrative investor level data from Brazil. Household finance, and in particular the study of the retail investor, has an important pedagogical role. For example, by understanding the repeated mistakes that the individual investor make, we can address this directly in financial education. This doctoral dissertation helps us to advance in the understanding of the Brazilian stock market and the behavior of its participants. Brazil has seen a significant increase in the number of people investing in stocks, making the study of individual investors increasingly relevant.

The first paper of this thesis is a collaboration with Fernando Chague and Bruno Giovannetti, and studies the relation between the salience of a stock and retail investors' trading decisions. The second paper is a collaboration with Marco Bonomo and Ruy Ribeiro and provides a life like monetary measure of profits and losses for all investors in the Brazilian stock and derivatives market. The last paper is a collaboration with Justin Mohr, and studies how social media might affect the behavior of the investors in the stock market.

This study was financed in part the by the São Paulo Research Foundation (FAPESP), grant 2020/09648-7, by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES) - Finance Code 88881.623310/2021-01 (PSDE program), and by the Brazilian Financial and Capital Markets Association (ANBIMA), throughout the *XVII Prêmio* ANBIMA *de Mercado de Capitais* – 2021.The opinions, hypotheses and conclusions or recommendations expressed in this material are the responsibility of the author and do not necessarily reflect the views of FAPESP, ANBIMA or CAPES.

2. Out of sight, out of mind: Local stores and retail day-trading

Abstract

Salience often covary with information. Hence, empirically showing that the salience of a stock, in itself, affects retail investors trading decisions is challenging. We document that living in a small city that has a local store of a brick-and-mortar firm more than doubles the chances of individuals picking the stock of that firm to *day-trade*. This suggests a direct relation between salience and retail investors' trading decisions: a local store in a small city i) increases the visual salience of the firm for the city residents but ii) does not provide any useful information for day-trading, which depends exclusively on high-frequency indicators. We explore the granularity of our dataset to control for indirect channels that can make retail day-trading correlate with local stores.

Keywords: retail investors decisions; limited-attention; salience; day-trade

1. Introduction.

Retail investors have the difficult task of choosing in which stocks to invest. With limited resources to evaluate the entire universe of stocks, they tend to focus on the ones that are salient[1] to them (Barber and Odean, 2008, and Barber, Lin, and Odean, 2019). Empirically showing that the salience of a stock, *in itself*, affects retail investors trading decisions is, however, challenging. Ideally, one would need a large experiment in which the salience of a group of stocks is exogenously changed. The issue with observational data is that usual proxies of salience (e.g., abnormal volume, abnormal returns, press coverage) often covary with the arrival of new information, which can also affect trading decisions.

In this paper, we investigate the relation between salience and retail trading by exploring a measure of salience that is uninformative for the retail trading activity that we choose to focus on. We use the existence of a local store in a small city of a brick-and-mortar exchange-listed firm[2] as our measure of (visual) salience of the firm to the retail traders who live in the city. We then focus on *day-trading*[3] to ensure that our salience measure is uninformative to the investor. Day-trading is a short-lived trading strategy that lasts minutes, hours at most. As such, day-traders rely exclusively on high-frequency indicators[4] to implement their strategies. Accordingly, a local store in a small city is unable to provide any information that can be useful specifically for day-trading.

To investigate the relation between local stores in small cities and retail day-trading, we explore two rich Brazilian datasets. The first contains the addresses of all stores from 2012 to

---

[1] According to Bordalo, Gennaioli, and Shleifer (2022), a stimulus is salient when it attracts the decision maker's attention "bottom up," automatically and involuntarily, and this can occur because of contrast with surroundings, surprise, or prominence.

[2] A street-side business that offers products and/or services to its customers face-to-face in an office or store.

[3] Day-trading is a trading strategy that involves buying and selling the same financial asset on the same day in the same quantity. According to a 2017 article in Forbes, "day-trading is the new sexy that gets an inordinate amount of hype" (https://www.forbes.com/sites/nealegodfrey/2017/07/16/day-trading-smart-or-stupid/#411e5d8a1007).

[4] For instance, intraday price variation, order sizes, and signed measures of trading flow (Bernstein, 1995).

2017 of all brick-and-mortar firms that are listed in the stock market between 2012 and 2017. This represents 60 firms with total market capitalization of US$ 322 billion (as of January 2015). The second dataset contains all day-trades in the Brazilian equity market of all individuals from 2012 to 2017 (a total of 8,746,980 day-trades performed by 190,655 individuals) and, crucially, the cities where these individuals live.

There were 5,570 cities in Brazil in 2017. Our baseline cutoff to define a small city is having a population of less than 100 thousand individuals (5,270 cities).[5] The main reason we focus on local stores and day-trades in *small* cities is because we do not have in our dataset the complete addresses of individuals we only observe the cities where they live. Hence, we cannot divide larger cities into neighborhoods and relate the day-trading activity in each neighborhood with the existence of a local store close by. However, we also believe that by focusing on small cities we have a cleaner empirical exercise: people usually circulate a lot in large cities (e.g., by living far from where they work) and, hence, they can see many stores which are not close to where they live.

Our main findings are the following. First, in an analysis within firm-month (across cities), we find that the likelihood of a firm being day-traded by individuals in a given month is 2.0 percentage points higher in a small city that has a local store of that firm compared to a small city that has no local store of that firm. Second, in an analysis within city-month (across firms), we find that the likelihood of a day-trade in a small city in a given month is 1.5 percentage point higher for a firm that has a local store in the city than for a firm that has no local store in the city. Third, in an analysis within city-firm (across months), we find that the likelihood of a given firm being day-traded in some small city is 0.9 percentage point higher in the months in which there is a local store of that firm in that city compared to the months when there is no local store. These effects are economically important: the unconditional probability of occurring some retail day-trade in a month for our average brick-and-mortar stock in our average small city is below 1% during the whole period.

To estimate these effects, we account for the fact that the locations of the stores across Brazilian small cities and over time are not randomly defined by the brick-and-mortar firms. For instance, a reasonable concern is that a firm is more likely to open a local store where (and when) its unobserved regional popularity is higher. This would generate a positive relation between local stores and retail day-trading if individuals are more likely to day-trade stocks

---

[5] We use different cutoffs for robustness analyses.

that are popular to them. However, the granularity of our dataset allows us to control for indirect channels like this. The dataset we build is at the city-firm-month level and, given these three dimensions, we can explore a number of different fixed-effects in the regressions.

First, we compare the day-trading activity by individuals in a given pair firm-month across two small cities (one with a local store of the firm, the other without) that are located in the same micro-region[6] of Brazil, keeping regional popularity constant. Second, we compare the day-trading by individuals in a given pair city-month across two different firms (one with a local store in the city, the other without), controlling for firm-microregion-month fixed-effects, which capture all omitted variables that vary in the firm-microregion-month dimension for instance, firms' regional popularity. Third, we compare the day-trading by individuals within a given pair city-firm across two different months (one with a local store of the firm in the city, the other without), also controlling for firm-microregion-month fixed-effects to account for possible dynamics in the regional popularity of the firm. Additionally, we also include city-month and firm-month fixed-effects in the regressions to control for all unobservables that vary in these dimensions.

To make things more concrete, consider the following three examples, one for each of the three specifications described in the previous paragraph. Lojas Americanas is a Brazilian retail company founded in 1940. In 2017, the firm had 1,144 local stores in Brazil. Suppose a small city A that has a Lojas Americanas store in month t and another small city B, located in the same micro-region of A, which does not have a Lojas Americanas store in month t. We test whether, in month t, there is a greater chance of a day-trade of Lojas Americanas by individuals from city A compared to by individuals from city B. To ensure that the comparison occurs only across small cities that are in the same micro-region, the regression includes firm-microregion-month fixed-effects. Furthermore, the regression also includes city-month fixed-effects to control for all possible social-economic differences across both cities that may affect both day-trading and the existence of the local store of Lojas Americanas (e.g., per capita income, population, and unobservables).

Second, considering the same small city A, suppose that while it has a Lojas Americanas store in month t, it does not have a store from another large Brazilian retail company, say, Magazine Luiza (846 stores in 2017). To compare day-trading by individuals

---

[6] Brazilian micro-regions are defined by IBGE, the Brazilian Institute of Geography and Statistics. There are 558 different micro-regions in Brazil, which are narrower in the more populated areas.

from city A in month t across both firms, we include firm-microregion-month and city-sector fixed-effects in the regression. The firm-microregion-month fixed-effects control for all possible differences across both firms that can vary regionally and over time and can affect both day-trading and the existence of a local store in the city, the regional popularity of the firms, for instance. The city-sector fixed-effects ensure that we are comparing only firms with similar business characteristics.

Third, consider that we are comparing across two different months the day-trading in Lojas Americanas by individuals who live in city A. In one month, there is a Lojas Americanas store in city A; in the other, there is no store. To control for changes in the popularity of Lojas Americanas in the region of city A that could affect both retail day-trading and the store's existence, we employ our stock-microregion-month fixed-effects. Moreover, to control for city A dynamics and Lojas Americanas dynamics that could also affect both retail day-trading and the store's existence, we employ both city-month and firm-month fixed-effects. Hence, in this specification Lojas Americanas store will need to vary over all three dimensions for the estimation of its effect on day-trading, a triple difference model.

Importantly, we show that when the local store has a diminished impact on the salience of the firm, the name that appears on the local storefront is different from the firm name, the effects of the local store on day-trading is either significantly smaller or null. We also present evidence that the effect of the salience of a local store coexists with the effect of other high frequency (daily) proxies of salience, with attention-grabbing events of a firm having stronger effect for investors who live in cities with local stores of that same firm. We additionally perform several robustness exercises to show that the results remain qualitatively the same under alternative definitions and samples. We use alternative definitions for day-trade and small cities, run our regression and controls at the daily level, and also look at the relation between stores and day-trading in medium and large cities.

We enhance our empirical analysis by studying the case of a specific firm in our sample, Lojas Americanas. The firm opened its first store in 1929 in Rio de Janeiro and went public in 1940, having its shares traded on the Brazilian stock market since then. During our sample period, the firm put in practice an expansion plan named "85 years in 5." The number of small cities with a local store went from 52 in 2012 to 231 in 2017. We explore this expansion to perform a difference-in-differences exercise. We look at the evolution of retail day-trading in the small cities where a store was opened and in the small cities where a store was not open.

First, we show that, before the event, the probability of retail day-trading, after we control for the population, income, number of stock market investors, and location of each city, is not statistically different between the two groups of cities. We then show that, in the years after the event, the probability of a day-trade in the cities where the local store is opened begins to increase, while the probability of a day-trade in the cities where the local store is not opened remains constant.

Barber and Odean (2008) suggest that the limited attention of retail investors can cause a buying pressure on a stock when it becomes salient: when deciding which stocks to buy among all existing stocks, retail investors allocate their limited attention to the stocks that are more salient. Since retail investors in general only sell what they have in their portfolios (a few stocks), selling decisions are less affected by salience. This is a very relevant hypothesis. If true, under limits to arbitrage, salience could then cause stock overpricing, at least temporarily. Barber and Loeffler (1993), Liang (1999), Da, Engelberg, and Gao (2011) and Engelberg, Sasseville, and Williams (2012) provide empirical evidence relating overpricing to salience.

Exactly as what happens with buying decisions, when individuals must decide which stocks they are going to day-trade, they also must choose among all stocks available in the market. Accordingly, the limited attention of day-traders should also be binding, just as highlighted by for buying decisions, and salience should also play an important role for day-trading: stocks that are salient should be more day-traded. By relating the existence of local stores in small cities with retail day-trading and considering that there is no relevant information for day-traders in these local stores (we empirically show that this is indeed the case, as expected), we deepen our understanding on the effects of salience on retail investors.

On a more general level, our empirical evidence corroborates a growing theoretical literature that emphasizes the role of salience in economic choice (Bordalo, Gennaioli, and Shleifer, 2012, Bordalo, Gennaioli, and Shleifer, 2013, and Bordalo, Gennaioli, and Shleifer, 2020). According to the survey of Bordalo, Gennaioli, and Shleifer (2022), when decision makers choose, their attention is allocated to the salient attributes of the choice options. Attributes of an option are differentially salient based on i) the contrast with the attributes of the other options, ii) the surprise compared to their usual values, and iii) the prominence with which they are displayed or retrieved. In our case, individuals are choosing among stocks to be day-traded and not among goods to be purchased, as in their models, but our salience measure for the stocks, the existence of a local store in a small city, is clearly related to prominence.

Accordingly, our evidence is consistent with the general theoretical prediction that salience should affect individuals' choices.

More directly, we contribute to the empirical literature that studies the effects of salience on retail investors. Barber and Odean (2008) show that retail investors are net buyers of stocks that present high levels for variables related to salience (trading volume, absolute returns, news). Hartzmark (2015) shows that individuals are more likely to sell stocks that are extremely ranked in their portfolios (both in terms of cumulative return since purchase and alphabetically). Kaniel and Parham (2017) show that flows to mutual funds increase when they are mentioned in Wall Street Journal "Category Kings" ranking list, compared to those funds which just missed making the list. Wang (2017) shows that ranking a stock in a more salient place can affect small investors and market variables such as volatility and volume. Choi, Haisley, Kurkoski, and Massey (2017) documents that the salience of savings rates affect 401k contributions. Frydman and Wang (2020) show that the salience of a stock's purchase price affects the disposition effect. Ozik, Sadka, and Shen (2021) show that retail trading in Robinhood platform exhibits a sharp increase among stocks with high COVID19-related media coverage. Barber, Huang, Odean, and Schwarz (2022) show that the purchase behavior of Robinhood users is highly correlated, what suggests that they engage in salience-induced trading.

We also contribute to the literature on the local bias of retail investors. It is well-documented that retail investors tilt their trading activity towards local stocks, usually defined as stocks with headquarters close to the investor - see, for instance, Huberman (2001), Ivkovic and Weisbenner (2005), and Seasholes and Zhu (2010). The reasons why local stocks are appealing to retail investors, however, are still not fully clear. This is something important to be understood since, by investing locally, individuals become subject to shocks that may affect both their earnings and their investments.

The most common explanation for the local bias is that investors may have some informational advantage by living close to the firm's headquarters. This is, however, controversial. On the one hand, Ivkovic and Weisbenner (2005), Massa and Simonov (2006), and Bodnaruk (2009) present evidence consistent with the existence of some informational advantage. On the other hand, Grinblatt and Keloharju (2001), Huberman (2001), Keloharju, Knüpfer, and Linnainmaa (2012), Goetzmann and Kumar (2008), Seasholes and Zhu (2010), and Døskeland and Hvide (2011) present evidence of the contrary, i.e., that individuals' investors do not outperform in local stocks. According to the evidence provided by our paper, salience could be an important reason under the local bias. A close headquarter could simply increase the salience of the stock.

Finally, our paper also relates to the literature that studies retail day-trading. Linnainmaa (2003), Jordan and Diltz (2003), Choe and Eom (2009), Ryu (2012), Kuo and Lin (2013), Barber, Lee, Liu, and Odean (2014), Barber et al. (2019), and Chague, De-Losso, and Giovannetti (2019) show that day-trading is common among individuals, who in general lose.

The remainder of the paper is organized as follows. Section 2 presents our datasets and relevant descriptive statistics. Section 3 shows the main empirical results. Section 4 shows the robustness exercises. Finally, Section 5 concludes.

2. Data: day-trading, small cities, micro-regions, and local stores.

We rely on data from two sources. First, the retail trading data come from the *Comissão de Valores Mobiliários* (CVM), the Brazilian equivalent to the SEC. The dataset is at the investor-stock-day level and contains the volume purchased, the volume sold, the quantity of shares purchased, and the quantity of shares sold by all retail investors in the Brazilian stock market from 2012 to 2017. We also observe the city of residence of each individual, which is crucial for our analyses.

Our focus in this paper is on a particular type of trading strategy: day-trading. We define an investor-stock-day observation as a day-trade if the quantity of shares purchased is equal to the quantity of shares sold.[7] There are 8,846,980 day-trades performed by 190,655 individuals in Brazil from 2012 to 2017 (across all stocks). Table 1 presents some descriptive statistics for these 190,655 individuals who performed at least one day-trade from 2012 to 2017. On average they are 40 years old (median of 37), display 198 trades (purchases, sells or day-trades) at the stock-day level (median of 70), 46 day-trades at the stock-day level (median of 4), an average purchasing volume of R$ 20,523 at the stock-day level (median of R$ 7,456), and an average purchasing volume of R$ 37,655 on a day-trading stock-day (median of R$ 12,034). Volumes

---

[7] In the robustness section, we say there is a day-trade if the investor both purchases and sells shares of the same stock on the same day, not necessarily in the same quantities.

on day-trading stock-day observations tend to be higher since individuals can use leverage when they are day-trading.

Our second data source is RAIS (Relação Anual de Informações Sociais) dataset.[8] It comes from the Brazilian Ministry of Labor and contains detailed information about all formally employed workers in Brazil. At the worker-month level, the dataset contains the worker's job description, wage, employer identification, job address, among many other information. We use this dataset to obtain the location of all stores of the 60 brick-and-mortar companies listed in the Brazilian stock market between 2012 and 2017. We also determine the dates when a new store opens in the city by looking at the date when the first worker of the store is hired; we do the same to infer store closures. In our baseline analyses, we focus on stores in small cities, which we define as the ones with less than 100 thousand people in 2017 (5,270 cities). In the robustness section, we change this definition and consider medium and large cities. The distribution of the number of local stores of these 60 firms in each triple (small city, firm, month) is naturally concentrated in 0 (96.13%), assuming the following other values: 1 (3.49%), 2 (0.34%), and 3 or more (0.04%).[9]

Table 1 - Day-traders descriptive statistics.

	statistics across 190,655 individuals					
	average	pct 5	pct25	pct50	pct 75	pct 95
number of trades (stock-day)	197.55	4	26	70	175	683
number of day-trades (stock-day)	46.40	1	1	4	19	148
average volume purchased (stock-day) in R$	20,523	1,008	3,236	7,456	17,775	69,367
average volume purchased in day-trades (stock-day) in R$	37,655	1,061	4,396	12,034	32,427	138,185
age in 2015	40.3	23	31	37	49	66

Note: this table presents descriptive statistics for the 190,655 individuals who made at least one day-trade between 2012 and 2017. A day-trade is defined as a stock-day observation in which the individual purchased and sold the same quantities of the stock. For each individual we compute the number of stock-day observations with some trading activity, the number of stock-day observation with a day-trading activity, the average volume among the stock-day purchases not related to day-trade, the average volume purchased in a stock-day observation with a day-trade, and his or her age in 2015.

The top map in Figure 1 presents the location of all 5,270 small cities in Brazil. The blue dots indicate the cities that have at least one store from a listed firm; the red dots indicate

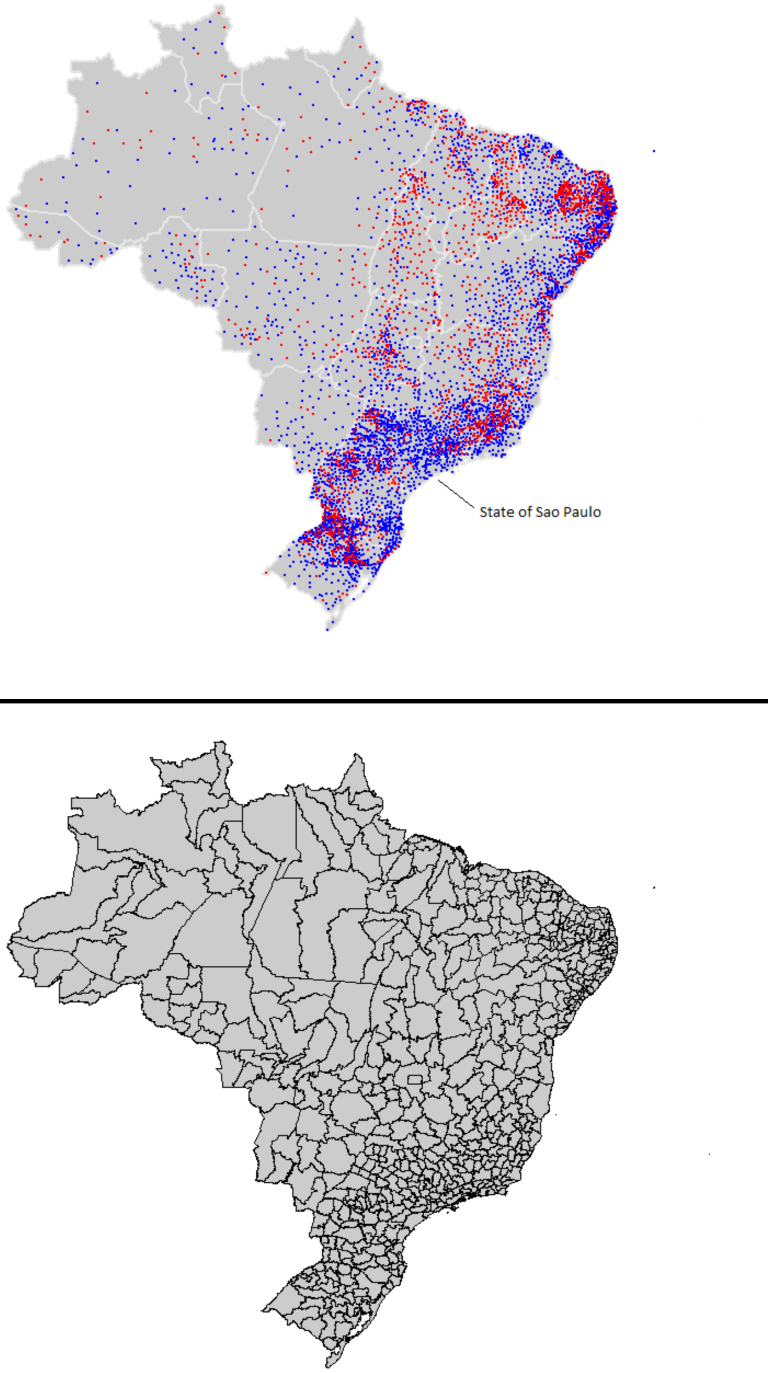
---

[8] This rich dataset has been successfully used in the labor economics literature (to mention a few MenezesFilho, Muendler, and Ramey, 2008, Meghir, Narita, and Robin, 2015, Ulyssea, 2018).

[9] In Brazil, the large bulk of local stores in small cities are family owned or run by small businesses that are not listed in the stock market.

the cities without stores from listed firms. The white borders indicate the 26 Brazilian states plus the Federal District of Brasília, the Brazilian capital. The bottom map in Figure 1 presents the 558 micro-regions of Brazil, which we use throughout the paper to control for any unobserved regional characteristics. These 558 micro-regions are defined by IBGE (the Brazilian Institute of Geography and Statistics) and are smaller in more populated areas. The average number of small cities in a micro-region is 9.5, the median is 8, the minimum is 1, the 25th percentile is 5, the 75th percentile is 13, and the maximum is 39.

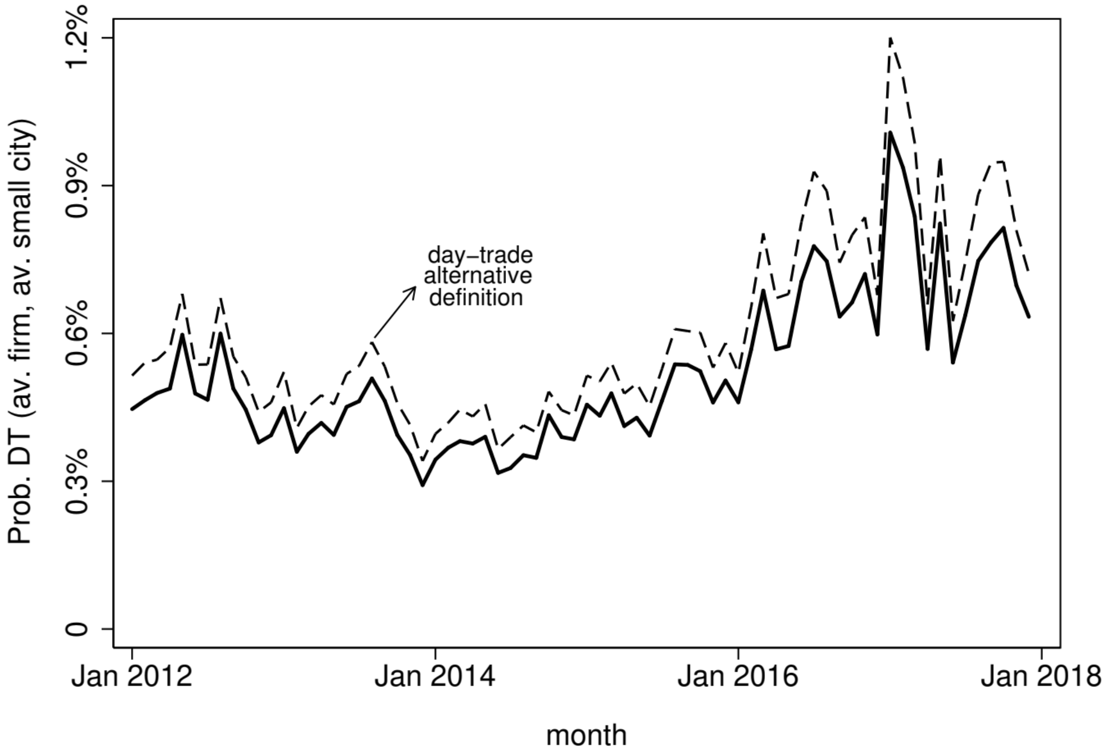
Figure 1 - Small cities and microregions



Note: the top map shows the 5,270 small cities in Brazil (the ones with less than 100 thousand people). The ones in blue have a local store of some of the 60 firms between 2012 and 2017; the ones in red have no local store of any of the 60 firms in the period; the white frontiers represent the Brazilian states. The bottom map presents the micro-regions of Brazil, which were defined by IGBE (the Brazilian Institute of Geography and Statistics) in 1990.

We look at the individuals who live in these 5,270 Brazilian small cities and we investigate their day-trading activity in the 60 brick-and-mortar listed firms. Figure 2 presents the probability of a retail day-trade in each month between 2012 and 2017 for our average firm, among the 60 brick-and-mortar firms, in our average small city (solid line). That is, the figure shows the average in each month across all pair's city-firm of a dummy variable that is one in case we observe some day-trading on that stock by individuals living in that small city and zero otherwise. As we can see, this probability reaches a minimum of around 0.3% by the end of 2013. After this point, the probability increases and gets closer to 1.0% by the end of 2017, when the number of retail day-traders increase in the Brazilian stock market.[10] This is an important baseline number in our paper; we will estimate how the presence of a local store in a small city can affect this unconditional probability.

Figure 2 - Probability of retail day-trading in the average small city and firm



Note: this figure presents the probability of a retail day-trade in each month between 2012 and 2017 for our average stock in our average small city. That is, the figure shows the average in each month across all pair's city-firm of a dummy variable that is one in case we observe some day-trading on that stock by individuals living in that small city and zero otherwise. The solid line is for the baseline definition of day-trade (same positive quantity purchased and sold). The dashed line is for the alternative definition of a day-trade used in the robustness section (positive quantity purchased and positive quantity sold).

The 60 brick-and-mortar companies listed in the Brazilian stock market have a combined market capitalization of US$ 322 billion in January 2015 (the middle of the sample), which corresponds to about 38% of the total market capitalization of the Brazilian stock exchange at the time. Table 2 shows the list of the firms, their sector, market capitalization, total number of stores in Brazil in the year with the highest value, and the number of small

---

[10] Brazil has seen a significant increase in the stock market participation by retail investor in recent years (https://www.b3.com.br/pt_br/noticias/investidores.htm).

cities with at least one local store in each year a missing value indicates that the firm was not listed in the Brazilian stock market in that year. The list of sectors represented is: real estate (11), services (11), retailers (9), banking and financial services (8 firms), apparel retailers (6), education (6), malls (6), and healthcare (3). As the table shows, firms from the financial sector, the large commercial banks, are present in more small cities than any other sector.[11] Also, 11 of the 60 firms have no local store in any small city during the sample period (but did have in larger cities). One particular firm, Lojas Americanas (row 43 in the table), shows a strong expansion in small cities in the period that we will explore in the empirical analysis.

2.1. Local stores in small cities do not provide useful information for day-trading.

Day-trading is a very short-lived trading strategy that lasts minutes, hours at most. As such, information for day-traders can only come from high-frequency indicators such as intraday price variation, order sizes, and signed measures of trading flow (see, for instance, Bernstein, 1995). Accordingly, a local store in a small city should not provide any useful information for day-trading for long-horizon investment, in turn, there could be, in principle, some valuable information in some local stores (see, for instance, Gerken and Painter, 2022).

---

[11] These are bank branches, not ATMs

Table 2 – Brick-and-mortar firms' descriptive statistics.

#	Firm name	Ticker	Sector	Stores show firm name?	Mkt. cap. in Jan 2015 (US$ million)	Total number of stores in Brazil	Number of small cities with a local store					
							2012	2013	2014	2015	2016	2017
1	Alpargatas	ALPA	apparel	n	1,935	57	11	11	11	10	9	9
2	Arezzo	ARZZ	apparel	y	1,005	66	3	3	4	5	5	5
3	Hering	HGTX	apparel	y	1,564	113	12	13	15	15	14	15
4	Guararapes	GUAR	apparel	n	2,071	321	9	6	6	5	3	3
5	Lojas Marisa	AMAR	apparel	y	973	438	1	1	1	1	1	1
6	Lojas Renner	LREN	apparel	y	4,065	524	1	1	2	3	4	7
7	Anhanguera	AEDU	education	y	1,944	190	16	16	16			
8	Kroton Educacional	KROT	education	n	6,226	190	16	16	16	16	16	16
9	Anima Educação	ANIM	education	n	541	52		3	3	4	5	6
10	Ser Educacional	SEER	education	n	930	75		0	0	0	0	0
11	Somos Educação	SEDU	education	n	1,158	163	8	7	7	7	7	7
12	Estácio Part.	YDUQ	education	y	1,955	191	2	2	2	2	2	3
13	Banco ABC Brasil	ABCB	finance	y	753	4	0	0	0	0	0	0
14	Banrisul	BRSR	finance	y	2,360	541	341	342	342	342	342	341
15	BB Seguridade	BBSE	finance	n	21,482	5713		2899	2915	2921	2916	2909
16	Banco Bradesco	BBDC	finance	y	54,454	6525	2028	2028	2019	2012	1998	1977
17	Banco do Brasil	BBAS	finance	y	28,117	5725	2871	2899	2914	2920	2915	2908
18	Banco BTG Pactual	BPAC	finance	y	7,737	24	1	1	1	1	1	1
19	Itau-Unibanco	ITUB	finance	y	65,719	5046	942	933	932	933	931	900
20	Banco Santander BR	SANB	finance	y	23,694	2541	527	527	526	495	494	493
21	Alliar	AALR	health	n	662	121					9	9
22	Fleury	FLRY	health	y	1,310	226	3	3	3	3	2	2
23	Hermes Pardini	PARD	health	y	1,021	134						7
24	Aliansce	ALSC	malls	n	1,102	39	0	0	0	0	0	0
25	Sonae Sierra Brasil	ALSO	malls	n	632	16	0	0	0	0	0	0
26	BR Malls Part.	BRML	malls	n	3,322	104	0	0	0	0	0	0
27	Iguatemi	IGTA	malls	y	1,654	55	0	0	0	0	0	0
28	Jereissati Part.	MLFT	malls	n	548	55	0	0	0	0	0	0
29	Multiplan	MULT	malls	n	3,853	117	0	0	0	0	0	0

30	Direcional	DIRR	real estate	y	488	81	0	1	1	1	2	2
31	Even	EVEN	real estate	y	498	9	1	1	1	1	1	1
32	Eztec	EZTC	real estate	y	1,196	43	0	1	1	1	1	1
33	Gafisa	GFSA	real estate	y	365	59	1	0	2	3	4	4
34	Cyrela Realt	CYRE	real estate	y	2,070	158	0	0	0	2	2	2
35	Helbor	HBOR	real estate	y	386	7	0	0	0	0	0	0
36	JHSF Part.	JHSF	real estate	y	485	37	1	1	1	1	2	2
37	MRV	MRVE	real estate	y	2,030	212	6	5	4	2	2	2
38	PDG Realt	PDGR	real estate	y	272	146	0	0	0	0	0	0
39	Rossi Residencial	RSID	real estate	y	76	166	1	1	0	0	0	0
40	Tenda	TEND	real estate	y	517	31						0
41	BR Pharma	BPHA	retail	n	283	836	81	91	95	95	91	86
42	Carrefour BR	CRFB	retail	y	11,712	592						12
43	Lojas Americanas	LAME	retail	y	6,716	1271	52	65	104	146	173	231
44	Magazine Luiza	MGLU	retail	y	742	903	301	303	314	323	340	386
45	Pão de Açucar CBD	PCAR	retail	y	7,966	1379	26	27	27	28	28	27
46	Profarma	PFRM	retail	n	194	203	7	7	8	8	7	7
47	Raia Drogasil	RADL	retail	y	3,551	1640	81	81	95	95	100	104
48	Saraiva Livrarias	SLED	retail	y	67	144	1	1	1	1	1	1
49	Viavarejo	VVAR	retail	n	2,644	1178			102	110	110	110
50	Azul S.A.	AZUL	service	y	3,614	108						21
51	BR Brokers	BBRK	service	n	184	118	2	3	2	1	0	0
52	CVC Brasil	CVCB	service	y	1,175	459			34	35	46	52
53	Gol	GOLL	service	y	1,237	93	10	10	9	10	10	10
54	IMC S/A	MEAL	service	n	453	188	13	14	14	14	14	13
55	Localiza	RENT	service	y	2,731	484	43	49	56	60	69	72
56	Lopes Brasil	LPSB	service	y	284	50	1	1	1	1	1	1
57	Movida	MOVI	service	y	696	200						6
58	Oi	OIBR	service	y	2,362	335	39	67	72	72	71	68
59	Tam S/A	TAMM	service	y	2,089	82	5					
60	Telefônica Brasil	VIVT	service	n	22,155	655	20	18	19	19	19	12

Note: this table presents the 60 retail firms that sell goods or services to individuals through local stores and are listed in the Brazilian equity market in some year between 2012 and 2017. They are sorted by sector. All these 60 firms are included in our regressions in the months they can be traded in the stock market missing values in the last columns of the table appear in the years the firm is not listed in the equity market. Among the 60 firms, 11 firms have zero local stores in small cities during the whole period.

To empirically show that local stores in small cities do not provide useful information for day-trading, we construct a dataset with all day-trades performed by individuals who live in small cities in all the 60 brick-and-mortar firms during 2012-2017 a total of 246,858 day-trades. We then estimate a daytrade-by-daytrade regression

$$Ret_{i,s,t} = \beta_1 Store_{i,s,t} + \gamma_{s,t} + \epsilon_{i,s,t} \tag{1}$$

where $Ret_{i,s,t}$ is the return of the day-trade performed by individual i on stock s on day t (computed as the total daily volume sold minus the total daily volume purchased divided by the total daily volume purchased), $Store_{i,s,t}$ is a dummy variable that is one if individual i lives in a small city that has a local store of firm s on day t, and $\gamma_{s,t}$ are stock-day fixed-effects (a constant for each pair stock-day).

The stock-day fixed-effects allow us to compare, for a given stock and on a given day, the returns obtained by all individuals who live in small cities and decided to day-trade that stock on that day. Coefficient β_1 is then comparing the result obtained by the average individual who lives in a small city that has a local store of the firm with the result obtained by the average individual who lives in a small city without a store. If local stores can give valuable information for day-trading, we should find $\beta_1 > 0$. In contrast, if day-traders cannot extract useful information from local stores, we should find $\beta_1 = 0$. As expected, this is indeed what we find. The estimated coefficient is equal to -0.0001 with t-statistic of -0.28.

To control for a composition effect, we extend equation (1) and include an individual investor fixed-effect γ_i (a constant for each investor). Day traders can be a diverse group, and day traders living in cities where stores are more likely to be located could be different from day traders living in cities less prone to have a store. Coefficient β_1 is then comparing the result obtained by the average day trade of an individual in a firm which has a local store in her small city with the result obtained by the average day trade of the same individual in a firm without a store in her small city. The estimated coefficient is equal to -0.0004 with t-statistic of -2.87, which means that day traders have a small worse performance when day trading stocks from firms with a local store in their small city (stocks salient to them), when compared to day trading stocks with no store in their small city. This is consistent with the story that place attention grabbing and salience as a behavior bias that is detrimental to the investor performance.

3. Main empirical analyses.

In principle, salience can affect the trading behavior of individuals (Barber and Odean, 2008). However, empirically showing that salience, in itself, affects retail investors is challenging, salience often covary with the arrival of new information.

We now document that an individual who lives in a small city has a significantly higher probability of *day-trading* a stock of a brick-and-mortar firm that has a store in that city. This evidence is consistent with salience, in itself, affecting the trading behavior of individuals. First, a local store in a small city clearly increases the visual salience of a brick-and-mortar firm for the city residents. Second, a local store in a small city provides no information that can be used for day-trading, as documented in the previous section.

Importantly, as we now carefully discuss, we explore the granularity of our dataset to control for confounding effects that can make retail day-trading activity to be indirectly related to the existence of local stores, such as, socioeconomic variables that vary across cities and over time, and firm-specific variables that vary regionally and over time.

3.1 Day-trading and stores: within a firm-month, across different cities.

We first compare day-trading activity across different cities within a given pair firm-month. Is the chance of, in a given month, individuals day-trading stocks of a given brick-and-mortar firm higher in a small city where the firm has a local store compared to another small city where the firm has no local store?

To answer this question, we construct a stock-city-month panel dataset that is balanced across i) the 60 brick-and-mortar firms from Table 2, ii) the 5,270 small cities in Brazil, and iii) all months in which the firm is listed in the Brazilian equity market between 2012 and 2017 (a total of 19,625,480 observations). We then run the following regressions

$$DT_{s,c,t} = \beta_1 Store_{s,c,t} + \gamma_{s,t} + \epsilon_{s,c,t} \quad (2)$$
$$DT_{s,c,t} = \beta_1 Store_{s,c,t} + \beta_2 Pop_{c,t} + \beta_3 Inc_{c,t} + \beta_4 Inv_{c,t} + \gamma_{s,t} + \epsilon_{s,c,t} \quad (3)$$
$$DT_{s,c,t} = \beta_1 Store_{s,c,t} + \beta_2 Pop_{c,t} + \beta_3 Inc_{c,t} + \beta_4 Inv_{c,t} + \gamma_{s,mr,t} + \epsilon_{s,c,t} \quad (4)$$
$$DT_{s,c,t} = \beta_1 Store_{s,c,t} + \gamma_{c,t} + \gamma_{s,t,mr} + \epsilon_{s,c,t} \quad (5)$$

where $DT_{s,c,t}$ is a dummy variable indicating whether we observe a day-trade on stock s in month t executed by individuals living in city c, $Store_{s,c,t}$ is a dummy variable indicating whether there is a local store from firm s in city c in month t, $\gamma_{s,t}$ are stock-month fixed-effects (a constant for each pair stock-month), $Pop_{c,t}$ is the log of the number of residents in city c in month t, $Inc_{c,t}$ is a proxy of the per capita income in city c in month t in thousands of reais,[12] $Inv_{c,t}$ is the number of individuals, divided by 100, who live in city c and who have traded (buy, sell, or day-trade) any stock in the Brazilian stock market in the 12-month period before month t,[13] $\gamma_{s,mr,t}$ are stock-microregion-month fixed-effects (a constant for each triple stock-microregion-month), and $\gamma_{c,t}$ are city-month fixed-effects (a constant for each pair city-month). In all regressions in the paper standard-errors are clustered by stock, by city, and by month.

Our baseline estimate for β_1 comes from equation (5), the one with the finest controls. However, the other specifications are also helpful to build intuition about the existing biases in the estimation of β_1. Table 3 presents the results.

Table 3 - Day-trading and local stores: within a pair firm-month, across cities.

	$DT_{s,c,t}$			
	(1)	(2)	(3)	(4)
$Store_{s,c,t}$	0.038***	0.025***	0.023***	0.020***
	(6.63)	(4.39)	(4.32)	(4.38)
$Pop_{c,t}$		0.001**	0.002**	
		(1.99)	(2.18)	
$Inc_{c,t}$		0.002	0.001	
		(1.57)	(0.83)	
$Inv_{c,t}$		0.078***	0.076***	
		(6.55)	(6.95)	
stock-month FE	✓	✓		
stock-microregion-month FE			✓	✓
city-month FE				✓
Obs.	19,625,480	16,574,098	16,574,098	19,625,480
Adj-R2	0.02	0.07	0.10	0.18

Note: this table shows the estimates of stock-city-month panel regressions of $DT_{s,c,t}$, a dummy variable that is one if there is at least one day trade in stock s, in month t, by an individual living in city c, on $Store_{s,c,t}$, a dummy variable that is one if there is a local store from firm s, in city c, in month t. Control variables are $Pop_{c,t}$, the log of the number of residents in city c in month t, $Inc_{c,t}$, the per capita income in city c in month t in thousands of reais, and $Inv_{c,t}$, the number of individuals, divided by 100, that live in city c who presented any trade (buy, sell, or day-trade) of any stock in the Brazilian stock market in the 12-month period before month t. Standard-errors

---

[12] We compute this proxy from the RAIS dataset by adding all wage incomes from all formal establishments in the city.

[13] When we use this control we must drop the year of 2012 from our regressions.

are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

In equation (2), by including stock-month fixed-effects, we are comparing the probability of observing a day-trade of a given firm s in a given month in a small city A, in which the firm has a local store, with this probability in a small city B in which the firm does not have a local store. The estimate of β_1 is significant at 1% and suggests that the probability is 3.8 percentage points higher in city A. This estimate should be upward biased, however. For instance, small cities that are richer have a higher chance of having more people day-trading in the stock market and, also, a higher chance of having a local store of firm s. The same concern should be valid for more populous small cities, for instance.

In equation (3), we then include three controls at the level city-month: the population, the per capita income and the number of individuals that invest in the stock market. Accordingly, we are now comparing day-trading of stock s across different small cities holding those variables fixed. The estimate of β_1 reduces to 0.025, significant at 1%.

Although cities are now comparable with respect to these important socioeconomic dimensions, we may still be comparing cities that are located in very different places. Brazil is a big country with some heterogeneities across its different regions. Hence, it may well be the case that a firm is popular in a region in the North but almost unknown in another region in the South. This unobserved variable, regional popularity, could also bias β_1 upwards; in a city located in a region where the firm is popular, there can be both more day-traders and stores.

In equation (4), we hence substitute the stock-month fixed-effects for stock-microregion-month fixed-effects (the 558 micro-regions that are shown in the bottom map of Figure 1). Now, we are comparing only small cities in a given month which are located in the same micro-region. The estimate for β_1 with this finer control is 0.023, still significant at 1%.

Finally, we employ city-month fixed effects instead of controlling for $Pop_{c,t}$, $Inc_{c,t}$, and $Inv_{c,t}$. Now, the regression controls for any dynamics in each city that may affect day-trading and were not being captured by those three observables. The estimate for β_1 is now 0.020, significant at 1%.

By employing stock-microregion-month and city-month fixed effects, equation (5) controls the effect of local stores on day-trading for important alternative indirect channels. We are comparing the day-trading activity on a given stock s in a given month across two small cities that i) are close to each other (are in the same microregion of Brasil) and ii) have the same city-month level characteristics, but one city has a local store of firm s and the other does not. We find that the probability of day-trading occurring is 2 percentage points higher in the small city with the local store. According to Section 2, the probability of occurring a day-trade for our average stock in our average small city is lower than 1%. This shows that the 2 percentage points increase is, indeed, very large.

3.2 Day-trading and stores: within a city-month, across different firms.

We now compare day-trading activity across different firms within a given small city in a given month. Do individuals who live in a given small city are more likely to day-trade the stock of a brick-and-mortar firm that has a store in the city than the stock of another brick-and-mortar firm that has no store?

To investigate this, we run the following stock-city-month panel regressions,

$$\begin{aligned}
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{c,t} + \epsilon_{s,c,t} && (6)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \beta_2 MktCap_{s,t} + \beta_3 TotalStores_{s,t} + \gamma_{c,t} + \epsilon_{s,c,t} && (7)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \beta_2 MktCap_{s,t} + \beta_3 TotalStores_{s,t} + \gamma_{c,sec,t} + \epsilon_{s,c,t} && (8)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{s,t} + \gamma_{c,sec,t} + \epsilon_{s,c,t} && (9)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{s,mr,t} + \gamma_{c,sec,t} + \epsilon_{s,c,t} && (10)
\end{aligned}$$

where $DT_{s,c,t}$ is the same dummy variable indicating whether we observe a day-trade on stock s in month t executed by individuals living in city c, $Store_{s,c,t}$ is the same dummy variable indicating whether there is a local store from firm s in city c in month t, $\gamma_{c,t}$ are city-month fixed-effects (a constant for each pair city-month), $MktCap_{s,t}$ is the log of the market capitalization of firm s in month t (the median value in the month), $TotalStores_{s,t}$ is the log of the total number of local stores that firm s has in Brazil in month t, $\gamma_{c,sec,t}$ are city-sector-month fixed-effects (a constant for each triple city-sector-month), $\gamma_{s,t}$ are stock-month fixed-effects (a constant for each pair stock-month), and $\gamma_{s,mr,t}$ are stock-microregion-month fixed-effects (a constant for each triple stock-microregion-month).

As before, our baseline estimates for β_1 comes from the last equation, the one with the finest controls, but we discuss the other equations to highlight the potential confounding effects for which we are controlling. Table 4 presents the results.

Table 4 - Day-trading and local stores: within a pair city-month, across firms.

			$DT_{s,c,t}$		
	(1)	(2)	(3)	(4)	(5)
$Store_{s,c,t}$	0.026***	0.023***	0.018***	0.016***	0.015***
	(3.85)	(3.64	(3.22)	(3.33)	(3.39)
$MktCap_{s,t}$		0.001	0.001		
		(1.16)	(0.78)		
$TotalStores_{s,t}$		0.001*	0.001**		
		(1.68)	(2.61)		
city-month FE	✓	✓			
city-month-sector FE			✓	✓	✓
stock-month FE				✓	
stock-microregion-month FE					✓
Obs.	19,625,480	19,625,480	19,625,480	19,625,480	19,625,480
Adj-R2	0.14	0.14	0.14	0.17	0.19

Note: this table shows the estimates of stock-city-month panel regressions of $DT_{s,c,t}$, a dummy variable that is one if there is at least one day trade in stock s, in month t, by an individual living in city c, on $Store_{s,c,t}$, a dummy variable that is one if there is a local store from firm s, in city c, in month t. Control variables are $MktCap_{s,t}$, the log of the market capitalization of firm s in month t (the median value in the month) and $TotalStores_{s,t}$, the log of the total number of local stores that firm s has in Brazil in month t. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

In equation (6), by including city-month fixed-effects, we are comparing, in a given pair city-month, the likelihood of observing a day-trade in the stock of a firm A that has a local store in the city in that month with the likelihood of observing a day-trade in the stock of firm B that has no local store in the city in that month. The estimate of β_1 is 0.026, significant at 1%, indicating a difference in these likelihoods of 2.6 percentage points. However, comparing firms of different sizes, for instance, should bias β_1. If firm A is larger than firm B, firm A has a higher chance of being day-traded (may be more popular in the country) and, also, of having a local store in city c.

To avoid this potential bias, we include two controls at the stock-month level in equation (7), namely, the firm's market capitalization and the total number of stores the firm has in Brazil. Thus, we are now comparing day-trading in city c in month t across different firms, with and without a local store, holding those important firm-level characteristics fixed. The estimate of β_1 reduces to 0.023, but is still significant at 1%.

Although market capitalization and total number of stores are controlled for, we may still be comparing firms from very distinct sectors, for instance, a commercial bank and a drugstore. If banks are, for some reason, more popular than drugstores, this could affect the estimation. In equation (8), we then substitute the city-month fixed-effects for city-sector-month fixed-effects. We are now comparing only firms of the same size and from the same sector. With these set of controls, the estimate of β_1 becomes 0.018, significant at 1%.

Next, we substitute the controls $MktCap_{s,t}$ and $TotalStores_{s,t}$ by general stock-month fixed-effects. By doing this, we control for any stock-specific time dynamics that may be correlated with day-trading. The estimate for β_1 becomes 0.016, significant at 1%.

In the specification of equation (9), there is one important potential bias still unaddressed. Bank B may have no branch in the city because it is not popular *in that specific region*. That is, the regional popularity of the firm may be affecting both the existence of a local store and the day-trading activity in the city. In the analysis of the previous section (within a given firm, across different cities) we addressed this concern by comparing only cities that are close to each other, i.e., in the same micro-region. In this section, we substitute the stock-month fixed-effects for stock-microregion-month fixed-effects. We are now comparing, within a given city-month, the likelihood of observing a day-trade of, say, a bank A that has a local branch in the city with the likelihood of observing a day-trade of a bank B that has no local branch in the city, and both banks are comparable regarding all characteristics that can vary at the microregion-month level, for instance, their regional popularity. The estimate for β_1 becomes 0.015, still significant at 1%.

The coefficient from equation (10) is our baseline estimate in this sub-section. That is, in a given pair city-month, the probability of individuals day-trading stocks from a firm with a local store is 1.5 percentage point higher than the probability of individuals day-trading stocks from a firm without a local store. Due to the set of fixed-effects included, both firms belong to the same sector and are comparable across all characteristics that vary at the firm-microregion-month level. Again, the magnitude of the coefficient is very large compared with the unconditional probability of observing a day-trade of below 1% reported in Section 2.

3.3 Day-trading and stores: within a pair city-firm, across different months.

Finally, we fix both the city and the firm. Do individuals who live in a small city day-trade more in the stock of a firm in the months when the firm has a local store in the city compared to the months when there is no local store? As we can see from the last six columns in Table 2, we have many instances of stores openings and closures across the small cities during our sample periods.

To answer this question, we now run the following stock-city-month panel regressions,

$$
\begin{aligned}
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{c,s} + \epsilon_{s,c,t} && (11)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \beta_2 MktCap_{s,t} + \beta_3 TotalStores_{s,t} \\
&\quad + \beta_4 Pop_{c,t} + \beta_5 Inc_{c,t} + \beta_6 Inv_{c,t} + \gamma_{c,s} + \epsilon_{s,c,t} && (12)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{s,t} + \gamma_{c,t} + \gamma_{c,s} + \epsilon_{s,c,t} && (13)\\
DT_{s,c,t} &= \beta_1 Store_{s,c,t} + \gamma_{s,mr,t} + \gamma_{c,t} + \gamma_{c,s} + \epsilon_{s,c,t} && (14)
\end{aligned}
$$

where $DT_{s,c,t}$ is the same dummy variable indicating whether we observe a day-trade on stock s in month t executed by individuals living in city c, $Store_{s,c,t}$ is the same dummy variable indicating whether there is a local store from firm s in city c in month t, $\gamma_{c,s}$ are city-stock fixed-effects (a constant for each pair city-stock), $\gamma_{s,t}$ are stock-month fixed-effects (a constant for each pair stock-month), $\gamma_{c,t}$ are city-month fixed-effects (a constant for each pair city-month), $\gamma_{s,mr,t}$ are stock-microregion-month fixed-effects (a constant for each triple stock-microregion-month), and the control variables in equation (12) are the same ones already used. Note that specification (13) and (14) are a triple difference model, given that it presents three double interaction fixed-effects (stock-time, city-time, and city-stock), and store need to vary in all three dimensions for the estimation of β_1.

In equation (11), by including city-stock fixed-effects, we are comparing, in a given city c and for a given firm s, the likelihood of observing a day-trade in a month when the firm has a local store in the city with the likelihood of observing a day-trade in a month when the firm does not have a local store in the city. The estimate of β_1 is 0.015, significant at 5%.

Equation (11) may be comparing very different months, however. First, cities change over time, what can affect both the number of local stores and the intensity of the retail day-trading. Second, firms also change over time, also affecting both their local stores and day-trading on their stocks. Accordingly, in equation (12), we include the same city-month and firm-month controls used before. The estimate of β_1 is 0.011, significant at 10%. Naturally, a more flexible way to control for city-month and firm-month characteristics is to use the city-month and the stock-month fixed-effects instead of these five control variables. This is what we do in equation (13), where the estimate of β_1 reduces to 0.009, significant at 10%.

Table 5 presents the results. As before, we begin with equation (11) and discuss all equations to highlight the potential alternative channels for which we are controlling.

Table 5 - Day-trading and local stores: within a city-firm, across months.

	$DT_{s,c,t}$			
	(1)	(2)	(3)	(4)
$Store_{s,c,t}$	0.015**	0.011*	0.009*	0.009**
	(2.10)	(1.78)	(1.77)	(2.03)
$Pop_{c,t}$		0.013		
		(1.61)		
$Inc_{c,t}$		0.003***		
		(3.52)		
$Inv_{c,t}$		0.064***		
		(4.82)		
$MktCap_{s,t}$		0.001		
		(1.56)		
$TotalStores_{s,t}$		0.002		
		(1.33)		
city-stock FE	✓	✓	✓	✓
stock-month FE			✓	
city-month FE			✓	✓
stock-microregion-month FE				✓
Obs.	19,625,480	19,625,480	19,625,480	19,625,480
Adj-R2	0.22	0.23	0.26	0.28

Note: this table shows the estimates of stock-city-month panel regressions of $DT_{s,c,t}$, a dummy variable that is one if there is at least one day trade in stock s, in month t, by an individual living in city c, on $Store_{s,c,t}$, a dummy variable that is one if there is a local store from firm s, in city c, in month t. Control variables are $Pop_{c,t}$, the log of the number of residents in city c in month t, $Inc_{c,t}$, the per capita income in city c in month t in thousands of reais, and $Inv_{c,t}$, the number of individuals living in city c who presented any trade (buy, sell, or day-trade) of any stock in the Brazilian stock market in the 12-month period before month t, $MktCap_{s,t}$, the log of the market capitalization of firm s in month t (the median value in the month), and $TotalStores_{s,t}$, the log of the total number of local stores that firm s has in Brazil in month t. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

There is, however, another potential indirect channel by which local stores may correlate to day-trading within a given pair city-stock. Suppose that the way the popularity of a given firm changes over time in Brazil is heterogeneous across different regions of the country, what could affect both local stores and day-trades differently in each region. This would not be captured by the stock-month fixed effect because of the heterogeneity across regions. To account for that, in equation (14) we use stock-microregion-month fixed-effects instead of stock-month fixed-effects. The estimate of β_1, however, does not change (it is now significant at 5%).

Summing up, for a given pair city-firm, the probability of individuals day-trading the stock in a month when there is a local store is 0.9 percentage point higher than the probability of individuals day-trading the stock in a month without a store. As before, the economic magnitude is important, as the unconditional probability of observing a day-trade in the average city and average firm is below 1% during our sample.

3.3.1 Openings and closings.

The parameter obtained from equation (14) captures the dynamic effect from both openings and closings of stores in a given pair city-firm. We now investigate whether this dynamic effect indeed comes from both openings *and* closings.

In our sample, there are 357 city-firm pairs, the "opening group", in which i) the number of local stores goes from 0 to 1 and ii) we observe at least 12 months with no store and 12 months with the store during our sample period.[14] In turn, there are 175 city-firm pairs, the "closing group", in which i) the number of local stores goes from 1 to 0 and ii) we observe at least 12 months with the store and 12 months without the store.[15]

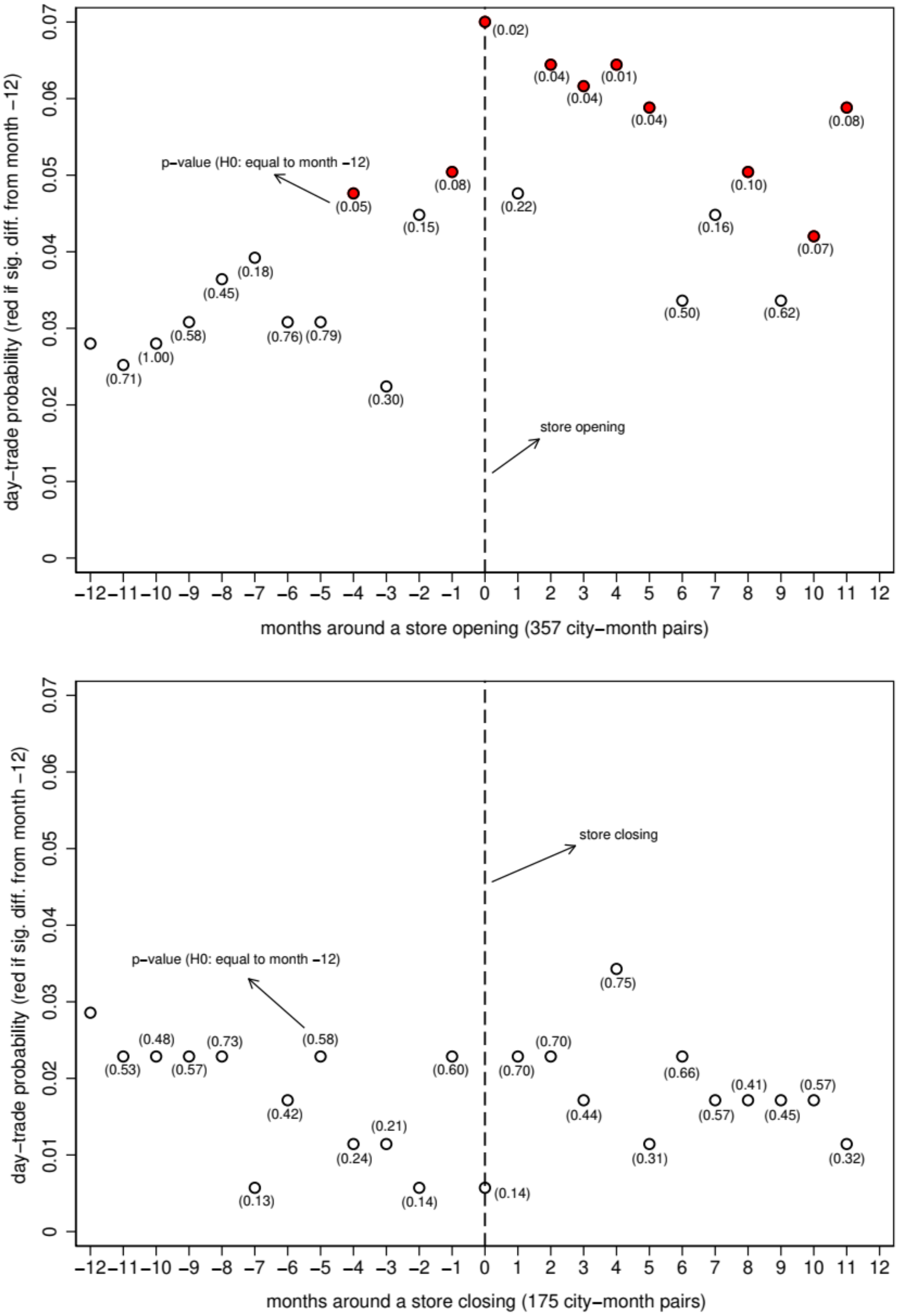
For these two groups of city-firm pairs, we estimate how the day trade probability evolves in the months around the opening or closing event. We evaluate how the day trade probability in months $-11, -10, \ldots, 0, 1, \ldots, 12$ compares with the day trade probability in month -12 (i.e., 12 months before the opening or closing event).

---

[14] These 357 city-firm pairs come from 26 different firms and 292 different small cities.

[15] These 175 city-firm pairs come from 19 different firms and 171 different small cities.

Figure 3 - Store opening vs. closing in a city-firm pair



Note: these figures show the probability of a retail day trade in 12 months before and 12 months after the opening or closing of a local store in a city-firm pair. In the top plot (openings) we focus on city-firm pairs in which the number of local stores goes from 0 to 1 and we have at least 12 months with no store and 12 months with the store (357 city-firm pairs). To construct the bottom plot (closings) we focus on city-firm pairs in which the number of local stores goes from 1 to 0 and we have at least 12 months with the store and 12 months with no store (175 city-firm pairs). Standard-errors are clustered by stock, by city and by month. For each month we present the p-value of the null hypothesis that the probability of a retail day trade in that respective month is equal to the probability of a retail trade in month -12 (i.e., 12 months before the opening or the closing event). If the p-value is below 5%, the circle is presented in red.

The top plot in Figure 3 presents the average day trade probability across the 357 city-firm pairs in the opening group in each one of the 25 months around the opening event (12 months before, the month of the store opening, and 12 months after). For each month we present the p-value of the null hypothesis that the probability of a retail day trade in that respective month is equal to the probability of a retail trade in month -12. If the p-value is below 5%, the circle is presented in red. In the period before the store opening, we see that the probability of retail day-trading is statistically stable, except for months -4 and -1. We believe that this anticipation of the effect may be because the local store construction should begin before month 0 which is the month when we observe workers already hired by the firm in the RAIS dataset. Therefore, it could be that the visual salience of that local store begins a few months before our indicator of store opening. In turn, in the month of the store opening (month 0) and in the 12 following months, the probability of retail day trade is in general significantly higher than in month -12.

The bottom plot in Figure 3 presents the average day trade probability across the 175 city-firm pairs in the closing group in each one of the 25 months around the closing event. In this case, the probability of retail day-trading is statistically stable in all months, i.e., significantly equal to month -12.

The dynamic effect coming from local stores openings (and not closings) seems consistent with the salience channel. Once a store opens, salience is affected in a sharp, discontinuous way. On the other hand, once a store closes, the firm should continue to be present in people’s minds, at least during the following months.

In the next section, we further study the dynamic effect of stores openings focusing on an important expansion plan of a brick-and-mortar firm that occurred in Brazil during our sample period.

3.4 Lojas Americanas case

The retailer “Lojas Americanas” provides us with an interesting case to study. During our sample, the firm pursued an aggressive expansion plan in small cities. We explore this to see how the effect of local stores on day-trading evolves over time.

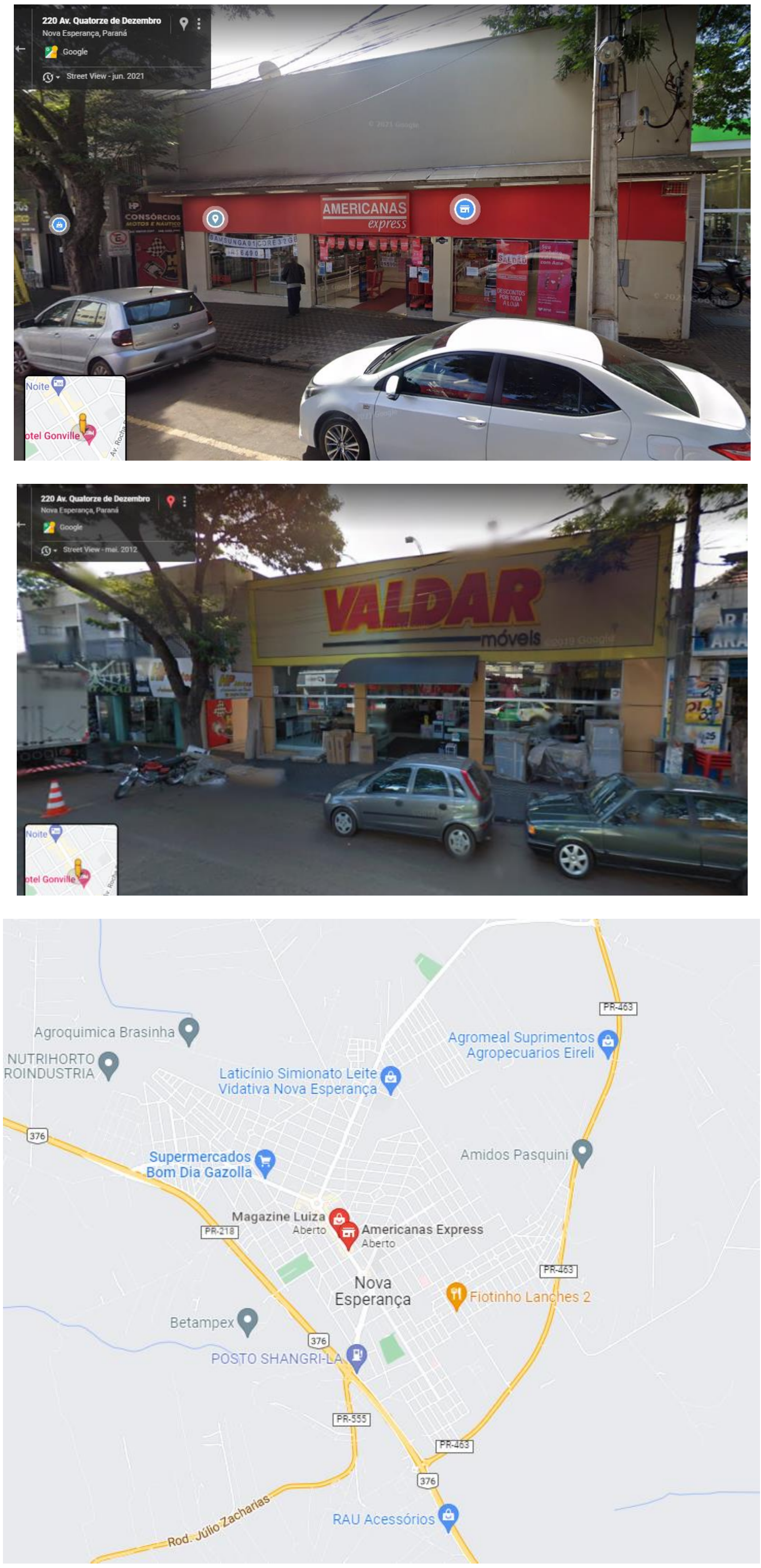
The firm opened its first store in 1929 in Rio de Janeiro and went public in 1940, having its shares traded on the Brazilian stock market since then. In 2014, the firm started an expansion plan named “85 years in 5.” As we can see in Table 2, in 2012, 52 small cities had a local store of the firm. This number increases to 65 in 2013, 104 in 2014, 146 in 2015, 173 in 2016, and 231 in 2017. Figure 4 shows a map of the small cities that received a local store between 2012 and 2017.

Figure 4 - Small cities that received a Lojas Americanas store between 2012 and 2017

Note: this figure shows all 107 small cities (the ones with less than 100 thousand people) that received a local store from Lojas Americanas in the years between 2012 and 2017.

To illustrate the salience mechanism, Figure 5 shows a photo of the front of a Lojas Americanas store taken from Google Streets in the city of Nova Esperança, a small city from the State of Paraná. We also show a photo from Google Streets of the same location taken in May 2012, when there was another store (a local furniture business) in the same location. The Americanas store is located in the main street of the city, which is usually the case since this is the place where all residents shop in these cities. Apart from its strategic location, the storefront clearly displays the firm’s name in white and red, increasing the firm’s salience in the city.

Figure 5 - Store front of a Lojas Americanas local store.



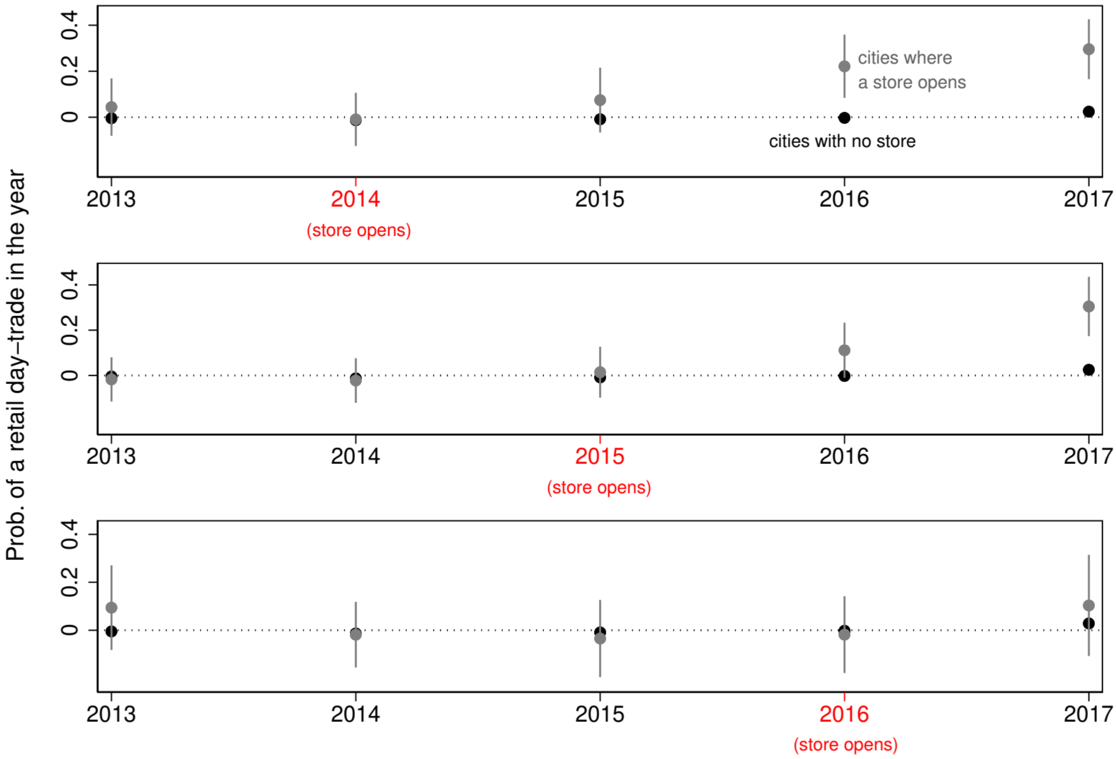
Note: the top photo shows the front of a Lojas Americanas store in the city of Nova Esperança, in the state of Paraná; the photo in the middle shows the same location in May 2012 (with a store from a non-listed firm). The map shows the city of Nova Esperança.

To investigate the dynamics of the relation between local stores and retail day-trading in the case of Lojas Americanas, we proceed as follows. We select all small cities in which a local store was opened in 2014 (39 cities across 36 micro-regions). For each one of these 39 cities we compute: i) a dummy variable $DT_{c,y}$ that is one in case we observe a day-trade of Lojas Americanas by an individual from this city in year y (2013-2017) and zero otherwise, ii) $Pop_{c,y}$, the monthly average of the log of the population of the city during year y, iii) $Inc_{c,y}$, the monthly average of the per capita income in the city during year y, and iv) $Inv_{c,y}$, the monthly average during year y of the number of individuals who live in city c and traded any stock in the previous 12 months. We also compute these four variables for all small cities in Brazil that have no stores of Lojas Americanas during the complete period between 2013 and 2017 (5,039 cities). With the city-year panel and both groups of cities, we then regress $DT_{c,y}$ on $Pop_{c,y}$, $Inc_{c,y}$, $Inv_{c,y}$ and γ_{mr} (micro-region fixed-effects), to obtain $ResDT_{c,y}$, the residual of this regression.

After obtaining $ResDT_{c,y}$, we then compute its average within each year across all 39 small cities in which a local store was opened in 2014, defining it as $\overline{ResDT_y}(1)$. Separately, we also compute the average of $ResDT_{c,y}$ within each year across all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017, defining it as $\overline{ResDT_y}(0)$. By doing that, we are estimating for these two groups of cities the probability of a retail day-trade of Lojas Americanas in each year, after controlling for population, income, number of investors, and location of the city. When focusing on openings within each year, we believe we can reduce the anticipation effect that might happen because the store is considered open only when we observe workers already hired in the RAIS dataset. Visual salience could start a few months earlier with the construction and advertisement of the new store.

The top plot of Figure 6 shows how $\overline{ResDT_y}(1)$ and $\overline{ResDT_y}(0)$ evolve over time. The gray circle displaying the 95% confidence interval refers to the 39 small cities in which a local store was opened in 2014. The black circle refers to the 5,039 small cities with no local store in the whole period given the large number of small cities in this group, the 95% confidence band is too narrow to be shown in the plot. In 2013 and 2014, the probability of retail day-trading, controlled for the population, income, number of stock market investors, and location of the city, was not statistically different between the two groups of cities. In turn, in the years after the store opening (2015, 2016, and 2017), the probability of a day-trade in a small city with the local store begins to sharply increase. In 2017, it reaches about 30%.

Figure 6 - The dynamics of the relation between stores and day-trading



Note: this figure illustrates the dynamics of the relation between local stores and retail day-trading in the case of Lojas Americanas. In the top plot, we select all small cities in which a local store was opened in 2014 (39 cities). For each one of these 39 cities we compute a dummy variable $DT_{c,y}$ that is one in case we observe a day-trade of Lojas Americanas by an individual from this city in year y (2013-2017) and zero otherwise, $Pop_{c,y}$, the monthly average of the log of the population of the city during year y, $Inc_{c,y}$, the monthly average of the per capita income in the city during year y, and $Inv_{c,y}$, the monthly average during year y of the number of individuals in city c that traded (buy, sell, or day-trade) any stock in the previous 12 months. We also compute these four variables for all small cities in Brazil that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (5,039 cities). With the city-year panel with both groups of cities, we then regress $DT_{c,y}$ on $Pop_{c,y}$, $Inc_{c,y}$, $Inv_{c,y}$, and Brazilian micro-regions fixed-effects. Define $ResDT_{c,y}$ as the residual of this regression. After obtaining $ResDT_{c,y}$, we compute its average within each year across all 39 small cities in which a local store was opened in 2014, defining it as $\overline{ResDT}_y(1)$. Separately, we also compute the average of $ResDT_{c,y}$ within each year across all 5,039 small cities in Brazil that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 ($\overline{ResDT}_y(0)$). The top plot shows how $\overline{ResDT}_y(1)$ and $\overline{ResDT}_y(0)$ evolve over time. The gray circles refer to $\overline{ResDT}_y(1)$ with its 95% confidence band. The black circle refers to $\overline{ResDT}_y(0)$ (given the large number of small cities in this group, the 95% confidence band is too narrow to appear in the plot). The same exercise is performed using all small cities in which a local store was opened in 2015 and 2016 (middle and bottom plots).

We also run the same exercise described above using all small cities in which a local store was opened in 2015 (41 cities in 39 micro-regions) and 2016 (27 cities in 27 micro-regions). The middle and bottom plots in Figure 6 present the results. As before, in the years before the store opening, the probability of retail day-trading, controlled for the population, income, number of stock market investors, and location of the city, is not statistically different between the cities where the store was opened and all other cities where no store was opened. In turn, after the store opening, the probability of a day-trade occurring begins to increase.

To complete this section, we summarize in a regression the information contained in Figure 6. We estimate the city-year panel regression

$$\begin{aligned} DT_{c,y} &= \beta_1 Treat_c + \beta_2 After_y + \beta_3 Treat_c \times After_y \\ &\quad + \beta_4 Pop_{c,y} + \beta_5 Inc_{c,y} + \beta_5 Inv_{c,y} + \gamma_{mr} + \epsilon_{c,y} \end{aligned} \tag{15}$$

Analogous to the top plot of Figure 6, we estimate regression (15) considering all 39 small cities in which a local store was opened in 2014 (for which $Treat_c = 1$) and all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (for which $Treat_c = 0$). The dummy variable $After_y$ is equal to one for the years 2015, 2016, and 2017, and zero before that.

Finally, analogous to the bottom plot of Figure 6, we also estimate regression (15) considering all 27 small cities in which a local store was opened in 2016 (for which $Treat_c = 1$) and all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (for which $Treat_c = 0$). The dummy variable $After_y$ is equal to one for the year 2017, and zero before that.

Table 6 presents the results. Column 1 refers to the regression with the 39 small cities in which a local store was opened in 2014. The coefficient β_3 indicates that the probability of a day-trade in 2015, 2016, or 2017 is 17.0 percentage points higher in a city that received a local store compared to a city that did not receive any local store. Column 2 refers to the regression with the 41 small cities in which a local store was opened in 2015. The coefficient β_3 indicates that the probability of a day-trade in 2016 or 2017 is 19.9 percentage points higher in a city that received a local store compared to a city that did not receive any local store. Finally, column 3 refers to the regression with the 27 small cities in which a local store was opened in 2016. The coefficient β_3 indicates that the probability of a day-trade in 2017 is 6.3

percentage points higher in a city that received a local store compared to a city that did not receive any local store.

Table 6 - Lojas Americanas: city-year panel regression.

	$DT_{c,y}$		
	(1)	(2)	(3)
$Treat_c$	0.054	0.020	0.021
	(1.23)	(0.71)	(0.39)
$After_y$	0.018	0.028*	0.042***
	(1.85)	(2.34)	(9.30)
$Treat_c \times After_y$	0.170**	0.199**	0.063**
	(3.19)	(2.76)	(4.46)
$Pop_{c,y}$	0.027	0.027	0.038*
	(2.11)	(2.02)	(2.52)
$Inc_{c,y}$	0.023*	0.014**	0.017**
	(2.19)	(3.82)	(3.93)
$Inv_{c,y}$	0.396***	0.402***	0.266**
	(8.80)	(8.04)	(3.98)
microregion FE	✓	✓	✓
Obs.	25,575	25,585	25,515
Adj-R2	0.24	0.24	0.21

Note: the table shows city-year panel regressions using the expansion of Lojas Americanas. We regress $DT_{c,y}$, a dummy variable that is one in case we observe a day-trade of Lojas Americanas by an individual from city c in year y and zero otherwise, on $Pop_{c,y}$, the monthly average of the log of the population of city c during y, $Inc_{c,y}$, the monthly average of the per capita income in city c during year y, $Inc_{c,y}$, the monthly average during year y of the number of individuals in city c who traded (buy, sell, or day-trade) any stock in the 12 previous months, microregions fixed-effects, and on $Treat_c$, $After_y$, and their interaction. In column 1, we consider all 39 small cities in which a local store was opened in 2014 (for which $Treat_c = 1$) and all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (for which $Treat_c = 0$), and the dummy variable $After_y$ is equal to one for the years 2015, 2016, and 2017, and zero before that. In column 2, we consider all 41 small cities in which a local store was opened in 2015 (for which $Treat_c = 1$) and all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (for which $Treat_c = 0$), and the dummy variable $After_y$ is equal to one for the years 2016 and 2017, and zero before that. In column 3, we consider all 27 small cities in which a local store was opened in 2016 (for which $Treat_c = 1$) and all 5,039 small cities that have zero local stores from Lojas Americanas during the complete period between 2013 and 2017 (for which $Treat_c = 0$), and the dummy variable $After_y$ is equal to one for 2017 and zero before that. Standard-errors are clustered by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

3.5 Local stores that do not increase firm's salience.

The local stores of 19 firms from Table 2 are arguably less likely to increase the salience of their respective firms. This is because the names that appear on the storefront differ from the names under which these firms are listed in the stock market. An example is the pharmaceutical retailer Profarma (stock PFRM3). The firm has local drugstores with four different flags, “Drogasmil”, “Farmalife”, “Tamoio”, and “Rosário”, which does not resemble the firm’s actual listing name, Profarma.[16] Column 5 in Table 2 describes whether each firm and their local stores have the same name.

When the storefront name does not clearly increase the firm’s salience, we expect the effect on day-trading to be smaller or null. That is, the estimates of β_1 in equations (4), (10), and (14), our baseline regressions, should be smaller or even insignificant for these 19 firms. Indeed, this is what Table 7 shows. Column 1 shows a smaller point estimate for β_1 equal to 0.008, although statistically significant at 1%. In columns 2 and 3, the estimates are statistically equal to zero. In turn, when we run these regressions with the other 41 firms that have the same names that appear on their storefronts, we find a positive effect. Columns 4, 5, and 6 show estimates equal to 0.022 (significant at 1%), 0.018 (significant at 1%), and 0.010 (significant at 10%), respectively.

---

[16] It is possible to see the four different flags here: https://grupoprofarma.com.br/en/our-flags/.

Table 7 – When the firm name is not salient.

	$DT_{s,c,t}$					
	store name ≠ firm name			store name = firm name		
	(1)	(2)	(3)	(4)	(5)	(6)
$Store_{s,c,t}$	0.008***	0.010	0.004	0.022***	0.018***	0.010*
	(5.85)	(1.43)	(0.67)	(5.15)	(3.51)	(1.71)
stock-microregion-month FE	✓	✓	✓	✓	✓	✓
city-stock FE			✓			✓
city-sector-month FE		✓			✓	
city-month FE	✓		✓	✓		✓
Obs.	6,448,760	6,448,760	6,448,760	13,139,480	13,139,480	13,139,480
Adj-R2	0.18	0.12	0.27	0.19	0.19	0.28

Note: the first three columns of this table show the estimates of equations (4), (10), and (14) when we use only the 19 firms that do not have their names in their local stores. The last three columns show the estimates of equations (4), (10), and (14) when we use only the 41 firms that have their names in their local stores. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

3.6 Local stores increase firm's salience on attention grabbing events.

Our low frequency measure of visual salience can coexist or intensify the effects of other salience and attention-grabbing events. To investigate this hypothesis, we remodel our panel to the daily frequency, the same frequency as popular attention-grabbing proxies. We also take advantage of this new data format to tighten even further our controls over possible sources of information for day-traders. On our last specification we include stock-microregion-day fixed-effects, which accounts for every news and information happening at the daily level for each stock and microregion, including daily attention-grabbing events.

We adapt Barber and Odean (2008) attention-grabbing proxies for our exercise. We use corporate news announcements from the public CVM database of "Consultation of Listed Company Documents" (*Consulta de Documentos de Companhias Abertas*)[17]. We also use previous day absolute return and abnormal volume - which is the volume on day t of stock s, divided by the average volume of stock s in the past 252 days - as proxies. Corporate news proxy is an indicator, one when there is news for firm s on day t, zero otherwise. While the previous day absolute return and abnormal volume proxies are standardized for each day over the 60 firms.

For computational purposes we exclude very small cities (less than 10 thousand individuals) from the following regressions. This does not affect the general results (see section 4.2). We use as benchmark the fixed-effects of equation (14), the one with the finest controls[18]. We estimate the stock-city-day panel regressions as follows:

$$DT_{s,c,t} = \beta_1 Salience\ Proxy_{s,t} + \gamma_{s,mr,m} + \gamma_{c,m} + \gamma_{c,s} + \epsilon_{s,c,t} \tag{16}$$

$$\begin{aligned} DT_{s,c,\mathrm{t}} &= \beta_1 Store_{s,c,\mathrm{m}} + \beta_2 Salience\ Proxy_{s,t} + \\ &\beta_3 Store_{s,c,m}\ \mathrm{X}\ Salience\ Proxy_{s,t} + \gamma_{s,mr,m} + \gamma_{c,\mathrm{m}} + \gamma_{c,s} + \epsilon_{s,c,t} \end{aligned} \tag{17}$$

$$DT_{s,c,t} = \beta_1 Store_{s,c,m} + \gamma_{s,mr,t} + \gamma_{c,m} + \gamma_{c,s} + \epsilon_{s,c,t} \tag{18}$$

---

[17] Specifically, we select five different types of announcements on the CVM documents database: "*DFP - Demonstrações Financeiras Padronizadas*", "*ITR - Informações Trimestrais*", "*Fato Relevante*", "*Aviso aos Acionistas*", "*Comunicado ao Mercado*". We use only the announcement date of the first document version available, that is, the first time the information became public. For documents made public after trading hours, weekends, and holydays, we shift their announcement date to the next trading day

[18] Using the fixed-effects of equation (4) and (10) does not affect the results.

where $DT_{s,c,t}$ is the same dummy variable indicating whether we observe a day-trade on stock s executed by individuals living in city c, but now on day t. $Store_{s,c,m}$ is the same dummy variable indicating whether there is a local store from firm s in city c in month m. $Salience\ Proxy_{s,\mathrm{t}}$ is an attention-grabbing proxy for firm s in day t - corporate news indicator, abnormal volume, and previous day absolute returns. $\gamma_{c,s}$ are city-stock fixed-effects, $\gamma_{c,t}$ are city-month fixed-effects, $\gamma_{s,\mathrm{mr,m}}$ are stock-microregion-month fixed-effects, and $\gamma_{s,\mathrm{mr,t}}$ are stock-microregion-day fixed-effects.

Table 8 shows several interesting results. First, columns (1), (3), and (5) show that the probability of day-trade firm s in a small city responds as expected (with an increase) on days of attention-grabbing events for firm s. Day-trades being affected by attention-grabbing events corroborates that limited attention is binding for investors decision about which stock to day-trade, like it is for the decision to buy stocks (Barber and Odean, 2008). Second, our cleaner and low frequency measure of visual salience coexist with salience proxies at higher frequency, columns (2), (4), and (6). Both positively affect the probability of day-trading, but only store coefficients remain unchanged when considering both measures, with the other proxies marginally reducing their magnitude and statistical significance. Third, the interaction term in columns (2), (4), and (6) show that high frequency and low frequency measure of salience reinforce themselves, increasing even more the probability of day-trading stocks which are salient to the investor. Last, column (8) shows that a local store continues to affect day-trade even after controlling for any event, news or other, that might affect a firm at the daily level.

Table 8 – Local stores intensifying the salience of attention-grabbing events, and daily level fixed-effects.

$Dep.Variable$	$DT_{s,c,t}$							
$Salience\ proxy$	$Corporate\ News_{s,t}$		$Abnormal\ Volume_{s,t}$		$Absolute\ Return_{s,t-1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Store_{s,c,t}$		0.001**		0.001**		0.001**	0.001**	0.001**
		(2.08)		(2.09)		(2.26)	(2.24)	(2.16)
$Salience\ proxy_{s,c,t}$	0.0004***	0.0003***	0.0003***	0.0002***	0.0002***	0.0001***		
	(6.37)	(5.04)	(4.17)	(3.86)	(4.07)	(3.66)		
$Proxy_{s,c,t} * Store_{s,c,t}$		0.001***		0.002***		0.0006***		
		(3.52)		(4.64)		(3.49)		
stock-microregion-day FE								✓
stock-microregion-month FE	✓	✓	✓	✓	✓	✓	✓	
city-stock FE	✓	✓	✓	✓	✓	✓	✓	✓
city-month FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	215,686,024	215,686,024	211,065,908	211,065,908	212,682,808	212,682,808	215,686,024	215,686,024
Adj-R2	0.15	0.15	0.15	0.15	0.15	0.15	0.13	0.33

Note: columns (1), (3), and (5) of this table show the estimates of equation (16). Columns (2), (4), (6), and (7) of this table show the estimates of equation (17). The last column show the estimates of equation (18) when we use stock-microregion-day fixed-effects. Corporate news is an indicator, one when there is news for firm s on day t, zero otherwise. Absolute return and abnormal volume proxies are standardized for each day of the sample over the 60 firms. Standard-errors are clustered by stock, by city and by day and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respective.

4. Robustness analyses.

In this section we show that the previously documented results remain qualitatively the same under some alternative definitions and samples. We use new definitions for what constitute a day-trade and small cities, focus the analyses on small cities of Sao Paulo, the richest State of Brazil, and look at the relation between stores and day-trading in big cities[19].

4.1 Alternative definitions for day-trade.

We identified a day-trade as a day in which the investor purchased and sold the same stock in the exact same quantities. In this section, we only require the investor to have purchased and sold the same stock on the same day, not necessarily in the same quantities.

The dashed-line in Figure 2 shows the probability of a day-trade, under this less restrictive definition, in each month between 2012 and 2017 for our average stock in our average small city. As expected, the probability is now slightly higher.

We replicate equations (4), (10), and (14) using the alternative definition for day-trade to compute the dependent variable $DT_{s,c,t}$. The results presented in Table 9 are qualitatively the same as before.

Table 9 - Day-trade alternative definition.

	$DT_{s,c,t}$		
	(1)	(2)	(3)
$Store_{s,c,t}$	0.023***	0.017***	0.011**
	(4.49)	(3.45)	(2.13)
stock-microregion-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	19,625,480	19,625,480	19,625,480
Adj-R2	0.18	0.19	0.29

Note: this table shows the estimates of equations (4), (10), and (14) when we use the alternative definition of day-trade to construct $DT_{s,c,t}$. Here, we define as a day-trade an individual-stock-day observation with a positive quantity purchased and a positive quantity sold, which do not have to be the same . Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

---

[19] On additional exercises, available upon request, we also observed that results remain when excluding the financial sector, the sector with the highest number of local stores in small cities.

We also verify whether individuals who already day-trade are also affected by the presence of a local store. To do that, we compute the dummy variable $DT_{s,c,t}$ considering only day-traders: individuals who have at least 10 other day-trades (in any stock, not necessarily in the 60 brick-and-mortar firms) in the previous 12 months. We replicate equations (4), (10), and (14). Again, the results presented in Table 10 remain qualitatively the same.

Table 10 - Day-trades by individuals who already day-trade.

	$DT_{s,c,t}$		
	(1)	(2)	(3)
$Store_{s,c,t}$	0.018***	0.014***	0.008*
	(4.30)	(3.39)	(1.78)
stock-microregion-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	19,625,480	19,625,480	19,625,480
Adj-R2	0.20	0.21	0.33

Note: this table shows the estimates of equations (4), (10), and (14) when we compute our dependent variable $DT_{s,c,t}$ considering only day-trades by individuals who have already made at least 10 other day-trades (any stock) in the previous 12 months. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

Another robustness exercise was to verify whether local stores affect both, the likelihood of the first day trade of an individual on that firm (extensive margin), and additional day trades from individuals who already day traded the stock of that same firm (intensive margin). To do that, we compute two different dummy variables $DT_{s,c,t}$. One considering only day-trades which were the first one done by some individual on that stock (intensive margin), and another considering all other repeated day-trades of the individual on that same stock (extensive margin). We replicate equations (4), (10), and (14). Again, the results presented in Table 11 remain qualitatively the same. Although estimated magnitudes between extensive and intensive margin are different, the relative magnitude (increase in probability when compared to the unconditional probability of each margin) is remarkably similar.

Table 11 – First and repeated day-trades.

	$DT_{s,c,t}$					
	Intensive Margin			*Extensive Margin*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Store_{s,c,t}$	0.018***	0.013***	0.007**	0.009***	0.007***	0.005**
	(4.37)	(3.37)	(2.16)	(5.29)	(3.80)	(1.94)
stock-microregion-month FE	✓	✓	✓	✓	✓	✓
city-stock FE			✓			✓
city-sector-month FE		✓			✓	
city-month FE	✓		✓	✓		✓
Obs.	15,962,830	15,962,830	15,962,830	15,962,830	15,962,830	15,962,830
Adj-R2	0.30	0.41	0.43	0.19	0.30	0.22

Note: the first three columns of this table show the estimates of equations (4), (10), and (14) when we compute our dependent variable $DT_{s,c,t}$ considering only day-trades which were the first one done by some individual on that stock (intensive margin). The last three columns show the estimates of equations (4), (10), and (14) when we compute our dependent variable $DT_{s,c,t}$ considering only repeated day-trades of some individual on that same stock (extensive margin). We use the first year of the sample (2012) to gather information about which stocks each individual investor was already day trading, therefore this year is not included in the regression. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

4.2 Alternative thresholds for small cities.

Our baseline definition for a small city is having a population of less than 100 thousand people in 2017. In this subsection, we first increase the number of cities in our regressions by changing this threshold to 250 thousand individuals. We replicate the baseline regressions (4), (10), and (14) using all cities with less than 250 thousand individuals (5,460 cities). The results presented in Table 12 are robust to the inclusion of more cities in the sample.

Table 12 - Including more cities.

	$DT_{s,c,t}$		
	(1)	(2)	(3)
$Store_{s,c,t}$	0.026***	0.022***	0.011**
	(4.06)	(3.20)	(2.41)
stock-microregion-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	20,333,040	20,333,040	20,333,040
R2	0.23	0.24	0.35

Note: this table shows the estimates of equations (4), (10), and (14) when we include all cities with less than 250 thousand individuals (5,460 cities) in the regressions. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

Table 13 - Excluding very small cities.

	$DT_{s,c,t}$		
	(1)	(2)	(3)
$Store_{s,c,t}$	0.015***	0.010**	0.010**
	(3.17)	(2.03)	(2.02)
stock-microregion-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	10,501,680	10,501,680	10,501,680
R2	0.20	0.20	0.29

Note: this table shows the estimates of equations (4), (10), and (14) when we exclude cities with less than 10 thousand individuals (3,120 cities) from the sample. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

The results are also robust to the exclusion of very small cities from the sample, i.e., 3,120 cities with less than 10 thousand individuals. We replicate regressions (4), (10), and (14) using all cities with population between 10 thousand and 100 thousand individuals (2,820 cities). Results are presented in Table 13.

4.3 Focusing on the State of Sao Paulo.

The State of Sao Paulo is the richest state in Brazil, being responsible for about 30% of the Brazilian GDP. It has 645 cities, 569 with less than 100 thousand people (i.e., small cities according to our baseline classification). Sao Paulo is the state with the highest density of small cities with local stores (see blue dots in Figure 1). We can thus estimate equations (4), (10), and (14) focusing only on the small cities in the State of Sao Paulo.

Table 14 presents the results. Results are qualitatively the same as the ones obtained when we look at the entire country.

Table 14 - Focusing on the State of Sao Paulo.

	$DT_{s,c,t}$		
	(1)	(2)	(3)
$Store_{s,c,t}$	0.030***	0.021**	0.009*
	(3.10)	(2.39)	(1.67)
stock-microregion-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	2,118,956	2,118,956	2,118,956
R2	0.23	0.24	0.34

Note: this table shows the estimates of equations (4), (10), and (14) when we focus only on the small cities of the state of Sao Paulo. Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

4.4 Medium and large cities.

Instead of studying the relation between day-trading and local stores using small cities, we could, in principle, focus on one very large city, divide it into neighborhoods, and relate the day-trading activity in each neighborhood with the existence of a local store close by. Unfortunately, this is not possible because we do not have the exact addresses of individuals (we only observe the cities where they live).[20] In this sub-section, however, we show that if we estimate equations (4), (10), and (14) considering only cities with more than 250 thousand individuals (a total of 110 cities), we obtain results that are consistent with the ones we have been presenting.

Given that we are now looking at medium and large cities, instead of using dummy variables for the existence of a local store and for whether we observe a day-trade, we use the number of local stores of firm s in city c in month t ($NumStores_{s,c,t}$) and the number of retail day-trades in firm s in city c in month t ($NumDT_{s,c,t}$). Moreover, instead of using micro-regions to control for regional unobserved effects we use states fixed-effects. We then run equations (4), (10), and (14) with those variables. Table 15 presents the results.

Table 15 - Large cities.

	$NumDT_{s,c,t}$		
	(1)	(2)	(3)
$NumStores_{s,c,t}$	14.79***	17.61***	73.42**
	(7.21)	(8.97)	(2.14)
stock-state-month FE	✓	✓	✓
city-stock FE			✓
city-sector-month FE		✓	
city-month FE	✓		✓
Obs.	409,640	409,640	409,640
R2	0.23	0.24	0.55

Note: this table shows the estimates of equations (4), (10), and (14) when we consider only cities with more than 250 thousand individuals (110 cities) in the regressions. Here, instead of using dummy variables for the existence of a local store and the existence of day-trade, we use the number of local stores of firm s in city c in month t ($NumStores_{s,c,t}$) and the number of retail day-trades in firm s in city c in month t ($NumDT_{s,c,t}$). Standard-errors are clustered by stock, by city and by month and the t-statistics are presented in parentheses. *, **, and *** indicate significance levels of 0.10, 0.05, and 0.01, respectively.

According to column 1 in Table 15, one additional local store of a given firm across two different cities with more than 250 thousand individuals increases in 14.79 the number of monthly day-trades of this firm in these cities, after controlling for stock-state and city-month fixed-effects (the average number of monthly day-trades in these cities is 96.32). According to

---

[20] We would like to note that, even if we observed the exact location of individuals, we believe the small cities may still offer a cleaner empirical exercise. In a big city, people circulate a lot, usually living far from where they work. As such, then can see many stores which are not close to where they live.

column 2, a firm that has one additional local store in a given city will have 17.61 more monthly day-trades compared to another firm after controlling for city-sector and stock-state-month fixed-effects. Finally, according to column 3, if a given firm builds an additional local store in a given city, the number of monthly day-trades should increase in 73.42 after controlling for city-month and stock-state-month fixed-effects.

5. Conclusion.

Day-trading is an extremely short-lived trading strategy that lasts, at most, a few hours. As such, it is very unlikely that day-trading can benefit from any piece of information that one could gather from a local store in a small city. Therefore, if the presence of a local store in a small city increases the chances of individuals day-trading the respective firm, it should be because of the increased visual salience, as the local store naturally makes the firm known to the city residents.

We find in this paper a robust and strong positive relation between the presence of local stores in small cities from brick-and-mortar firms and the propensity for day-trading the respective stock among the individuals who live in the city. Importantly, the granularity of our dataset allows us to control for many indirect channels that could be behind this relation. The relation stands even when controlling for anything that happened in a firm at the daily level (i.e., other salience shocks). We also perform several robustness exercises. In particular, we examine the case of firms that have a different storefront name from their listing names in the stock exchange and we find that the effects are weaker or null, as expected.

We believe that our results increase the understanding about the direct effects of salience on the behavior of retail investors. By combining a low-frequency salience measure the existence of a local store in a small city with a high-frequency trading activity that is becoming increasingly common among individuals' day-trading we can plausibly isolate the potential information-related channel that often pollutes the relation between stock salience and trading decisions.

3. From Poor to Rich: Assessing Wealth Transfer through Trading in the Stock and Derivatives Market

Abstract

This paper provides a life like monetary measure of gains and losses for all investors participating in the Brazilian stock and derivatives market. We document that most retail investors experience monetary gains when trading, but that the investors with negative outcomes lose higher amounts, and drive volume-weighted and aggregate results. Investors with smaller accounts have higher propensity of losing money and lose relatively higher amounts. Through a contrafactual exercise, we show that this led to wealth transfer from poor to rich and increased wealth concentration in equity holdings. We show that same day returns, and trading derivatives contracts have significant participation in retail investor bad performance and contribute to wealth redistribution, with poorer investors doing worse on these trades. This suggests that studies not taking these factors into account could be seen as a lower bound for the individual trading loss. Under-diversification and contrarian trading pattern are both negatively correlated with wealth and are associated with worst monetary outcomes, even when controlling for account size. The evidence in this study is more relevant in nowadays context of the growing popularity of commission-free brokerage apps, which incentivize trading in the stock and derivatives market.

Keywords: retail investor; return heterogeneity; contrarian behavior; wealth redistribution.

1. Introduction.

Active trading in the stock market by retail investors is detrimental to their wealth. Twenty years have passed since the first articles revealed that individual investors lose by trading, with many studies that followed verifying the same pattern[21]. Still, with technology reducing the cost of stock market participation, and the popularity growth of commission-free brokerage apps, the number of retail investors actively trading in the stock market have been rising[22]. On one hand, easier access seems beneficial given the evidence of household limited participation, especially for the lower strata of the wealth distribution[23]. On the other hand, if individuals trade "too much" when directly participating in the stock market, there could be harmful consequences from this participation. Importantly, the consequences may be different among distinct investors. On this paper we investigate how heterogeneity in trading outcome results in wealth transfer and impact the dynamics of wealth concentration in equity holdings.

We document that individual investors lose to institutions when they trade, with poorer individuals doing significantly worse. Our analysis reveals that individual active trading led to a wealth transfer from poor to rich and increased wealth concentration in equity holdings. To the best of our knowledge, this study has the most complete setting to study how trading in the equity market leads to monetary gains and losses for each participant. Our investigation considers all investors and all trades in the stock and derivatives market, considers the investor's actual execution price and holding period, and measures their cumulative monetary gain or loss from each trade. All these elements allow us to be closer to the true investor experience and wealth transfer among investors.

---

[21] Odean (1999), Barber and Odean (2000), Barber, Lin and Odean (2021) for the United States, Grinblatt and Keloharju (2000) and Grinblatt, Keloharju, and Linnainmaa (2012) for Finland, Barber et. al. (2009) for Taiwan, Anderson (2013) for Sweden, and Chen et. al. (2007) and Jones et. al. (2021) for China, all document that individual investors underperform when trading.

[22] There is evidence of increase in retail trading activity and participation in several stock exchanges, such as in the USA, UK, Europe, India, and Brazil; https://www.wsj.com/articles/individual-investor-boom-reshapes-u-s-stock-market-11598866200; https://www.bloomberg.com/news/articles/2022-01-27/u-k-retail-trading-platforms-see-renewed-demand-in-wild-markets; https://www.reuters.com/business/retail-investor-base-doubles-europe-us-meme-stock-mania-spreads-euronext-2021-06-11/; https://www.bloomberg.com/news/articles/2022-02-09/retail-investing-boom-in-india-stays-strong-as-foreigners-flee; https://www.economist.com/finance-and-economics/2021/01/14/first-time-investors-are-flooding-brazils-stockmarket.

[23] The literature on limited participation is extensive, not only documenting but trying to explain why a large fraction of households do not invest in risky financial assets. Haliassos and Bertaut (1995), Vissing-Jørgensen and Attanasio (2003), Calvet, Campbell and Sodini (2007), Grinblatt, Keloharju and Linnainmaa (2011) and Bach, Calvet and Sodini (2020) are studies for different countries that document the limited participation of households in the stock market.

Our setting under study is the Brazilian stock and derivatives market from 2012 to 2019, the largest of Latin America in market capitalization and trading volume. We rely on high-detail administrative data, with all transactions (including options and future contracts) and all investors (including institutions and individuals). The investigation is at the daily level and considers the average execution price of the trade. Hence, we obtain the precise monetary flow of each trade at an optimal frequency to study trading flows and its monetary outcome. We use investors holdings to rank investors and divide them into groups of wealth.

Our main findings are the following. Retail investors lose R$ 17.1 billion to institutions by trading, U$ 5.7 billion using the sample average exchange rate of 3.01 reais per dollar. This is approximately 1.3% of their holdings every year, while poorer investors lose over 2% of their holdings every year. Considering the same day return and derivatives market increased retail loss in 57% and considering the actual holding period of the investor increased retail loss in 9%. In every year of our study, we observe an increase in inequality of equity holdings, with investor active trading contributing 43% for the increase in inequality. These findings suggest that active trading is more harmful to the investors who can least afford them.

To estimate wealth transfer we expanded the methodology of An, Lou and Shi (2022). Their cumulative measure of trading gain and losses has the advantage of being a dollar-weighted measure considering the true holding period of the investor. This estimation reflects a lifelike experience of investors. Additionally, money transfers are what matters when studying wealth distribution. We adjust the measure to capture flows from the derivatives market and the trading execution prices, getting a complete picture of the trading experience. The measure gives us an accurate estimate of the monetary quantity that each investor gained and lost in the stock and derivatives market.

With those tools we surprisingly find that most individual investors have positive results when trading. However, aggregate results for each quintile of wealth show that individuals as groups lost money, which implies that when investors lose by trading, they lose larger amounts of money than the investors who win by trading. Our findings suggest that results for individual investors obtained with aggregated and volume-weighted data deserve a caveat. These studies are much more representative of a small group of wealthy and active investors and cannot represent the broad individual investor experience.

Also, the number of investors winning when trading increases with the investor wealth, with the poorest quintile having most individuals losing. The characterization that individuals with lower wealth underperform and lose money is persistent in time and robust to different subsamples. To reduce concerns that our results come from investors adjusting their wealth share in stocks based on their experience in trading, we sort investors by their first month net inflow to the stock market. Investors with smaller first month inflows have *ex-post* worse results. In all eight years under study, we observe new investors with lower wealth under-performing and losing proportionally large sums of money.

When investigating trading patterns that can contribute to underperformance, lack of diversification is the big “villain”. Investors with low relative diversification within each wealth group lose higher amounts of money, while lack of diversification is higher for investors with lower wealth. A contrarian trading pattern - buying stocks with recent negative changes in price - is negatively correlated with wealth and led retail investors to lose even more money. We find that the level of activity (monthly turnover) is well distributed among wealth groups, with high turnover within each wealth group also increasing losses, although in smaller magnitude, which endorses the first findings of Odean (1999).

Since we estimate monetary gain for everyone, we can conduct a wealth redistribution analysis. Through a counterfactual exercise we describe how these wealth transfers among investors affect inequality in equity holdings. In the exercise we shut down different trading channels that affect wealth distribution and compare how inequality would have evolved under different assumptions. The general idea for the wealth transfer measure and counterfactuals is to compare the wealth of an investor with a no-trade strategy against the wealth of this same investor after trading. We show that trading contributes significantly to the concentration of wealth in equity holdings, with the same day stock market returns and derivatives contracts further increasing inequality in equity holdings.

The paper that most relates to our study is An, Lou and Shi (2022), ALS from now on. Our paper also focuses on trading monetary gains and losses and the resulting wealth transfers among investors. The authors reveal that there was an intense wealth transfer during a stock market bubble in China from the investors on the bottom 85% of wealth distribution to the investors on the top 0.5%. However, we do not focus on how all other investors fair against the ultrawealthy investors. We calculate monetary gains and losses for each investor, having a better picture of what happens in the lower strata of the wealth distribution, and allowing us to assess changes in inequality through a counterfactual exercise. Our framework is also different, participation in trading volume by retail investors in the Brazilian stock market is much smaller than other studied emerging markets, such as China and Taiwan, and is closer to studies of developed countries, like US and Finland. Hence, our study also documents wealth transfer from individuals to institutions. Finally, with a longer sample we can identify the dynamics and persistency of the results, showing that growth in inequality was persistent through time.

Our paper contributes to the debate on the performance of individual investors in the equity market, primarily to the literature on the heterogeneity of performance across investors. We show evidence of cross-sectional variation due to account size, diversification, and contrarian investment style. Cross-sectional variation in performance could happen because of differences in investor experience (Seru, Shumway and Stoffman, 2010), cognitive ability (Grinblatt and Keloharju 2021), financial literacy (Bianchi, 2018), or differences in investment style, such as activity (Barber and Odean, 2000), diversification (Goetzmann and Kumar, 2008), attention-based trading (Barber, Lin and Odean, 2021), and gambling preferences (Kumar, 2009). Further, these variations tend to be persistent over time, as shown by Coval, Hirshleifer, and Shumway (2005).

First studies on retail performance in the stock market, such as Barber and Odean (2000) and Grinblatt and Keloharju (2000), did not find variation in performance due to differences in account sizes. However, recent studies for China as in Li et al (2017) and Jones et al (2021) associate outperformance and investor wealth, especially for the very wealthy, top 0.5% of the distribution. Also, Barber, Lin and Odean (2021) show evidence that smaller retail trades in the US, which are likely made by investors with lower wealth, underperform in comparison with larger retail trades. Our study complements this literature by estimating the actual monetary return experienced by investors after a trade, showing how investment strategies relate to account size and performance, and showing that retail investors in the bottom quintile of wealth are the only quintile with most of its investors losing.

This paper also adds to the literature of how individual investors trade and perform in the derivative markets (Bauer, Cosemans and Eichholtz 2009, and Li, Subrahmanyam and Yang 2021). We report that individuals as a group lose money when trading derivatives, accounting for 19% of the total loss. Also, poorer investors that trade derivatives lose impressive amounts of money, which makes derivatives trading account for 34% of the growth in inequality coming from active trading..

Our findings also relate to studies that account for the return on the same day of the trade. Barrot, Kaniel, and Sraer (2016) and Barber, Lin and Odean (2021) document that individuals underperform from the time they trade until the closing prices of that same day. Disregarding this same day return, like in hypothetical trading strategies based on retail order imbalance, could lead to a major difference from investor actual investment experience. Our results reveal that 20% of retail trading loss comes from the same day return, with this number going up to 40% for lower wealth strata.

Lastly, our paper also contributes to the recent discussion on how return heterogeneity affects wealth inequality. Piketty (2014) and Saez and Zuchman (2016) highlight that return differentials between the poor and wealthy may raise wealth inequality. Recent studies using annual tax records from Nordic countries support this view (Bach, Calvet and Sodini, 2020 and Fagereng *et. al.*, 2020). Tax records are comprehensive over different asset classes but lack the high frequency to investigate how active trading contributes to return heterogeneity and concentration of wealth. Campbell, Ramadorai and Ranish (2019) study monthly stock portfolios in India and find that average raw returns are decreasing in wealth, but that average log returns, which depend on diversification, are increasing with wealth. This study shows that within a wealth group, raw returns are also increasing with diversification, and that active trading was as important as portfolio heterogeneity for the inequality increase in stock holdings in Brazil.

The rest of the paper is organized as follows. Section 2 describes the data and shows descriptive statistics across different brackets of wealth in equity holdings. Section 3 investigates the trading gains and losses of retail investors and the wealth transfer across distinct groups of wealth. Section 4 assesses how trading results hold under different subsamples and its dynamics in time. Section 5 discusses the relation between account size, different investment styles and wealth transfers. Section 6 assesses how wealth transfer contributes to changes in inequality in equity holdings. Section 7 concludes.

2. Data.

We rely on daily investor-level administrative data from *the Comissão de Valores Mobiliários* (CVM), the Brazilian equivalent to the Securities and Exchange Commission (SEC)[24]. The data is highly reliable and contain all transactions (including options and single-stock future contracts) between January 2012 and December 2019 on the B3, the only Brazilian

---

[24] Chague *et. al.* (2018) and Birru *et. al.* (2022) study a different extraction of this same data.

stock exchange. We observe all investors, including foreign investors, mutual funds, and individuals. Data are de-identified, with investors assigned permanent reference numbers that allow us to follow them over time. This data also provides quarterly snapshots of all investor's holdings in the stock market starting in December 2016[25].

For each investor-day pair we observe which stocks were bought and sold, number of transactions, total quantity, and monetary volume. Thus, for each investor-day-stock we have the average execution price for that transaction. Important to this paper analysis, we also observe the derivatives contracts bought and sold, their monetary volume, and when, or if, the contract was executed. Investors may have accounts on different brokerages, and we observe their aggregate daily transactions in all brokerages.

We also rely on other complementary data. First, we use B3 quote history to collect information on the derivatives traded in the Brazilian Stock Market. We collect data on each contract type, underlying assets, expiration date, and execution/strike price. Other complementary data comes from the Economatica database. We collect general company information as well as daily (un)adjusted prices, volume, outstanding shares, and historical events (dividends, splits, and follow-ons). Lastly, we also gather information on other events such as IPO`s, OPA`s, mergers, incorporations, and spin-offs. The complementary data allows us to accurately estimate gains, losses and redistribution coming from trading in the stock and derivatives market, as well as provide a precise mapping from transactions to daily portfolio holdings[26].

This paper objective is to estimate monetary redistribution through trading; therefore, we focus on stocks which traded at least 99% of the days they were listed. The final sample has 226 stocks, which account for 98.7% of all trades and 99.1% of all monetary volume traded in individual stocks. There are 343,554 derivatives contracts traded on those stocks. The final data has 156 million daily observations with over 4.2 billion trades performed by 1,701,459 individual investors and 51,150 institutions.

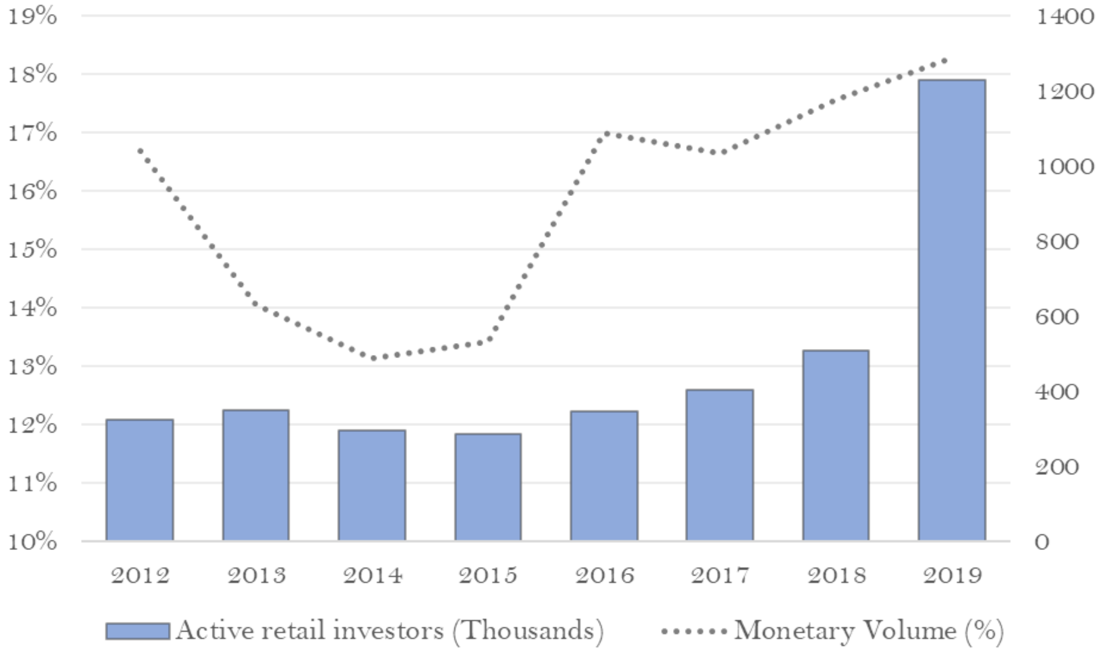
---

[25] Single positions larger than 5% of the total outstanding shares are public available. Therefore, these observations are suppressed for the data to remain de-identified.

[26] CVM transaction data only record transactions which happen through electronics orders. Thus, investor portfolio may change because of over-the-counter transactions, block trades, inheritance, donations, follow-ons, or any other transaction not recorded electronically. Nevertheless, those different operations are a minority. From one quarter snapshot on holdings to another quarter snapshot we can recover the investor's portfolio through the trading data with a 99% precision.

Brazilian stock market is the largest in Latin America with a market capitalization that fluctuated around U$ 1 trillion between 2012 and 2019. Retail investors represent on average 16% of total monetary volume, with their highest share of 18.3% in 2019, the last year of the sample (see Figure 1). The number of active retail investors (trading at least one stock in the year) was stable until 2015, varying between 287,000 and 352,000 active investors. After 2015, the number of individuals trading in the stock market started growing steadily, with a sharp increase in 2019. The substantial increase seen in 2019 continued in 2020 and 2021[27].

Figure 1 – Number of active retail investors and retail investors share in total monetary traded volume.



Notes: The right axis refers to the number of active retail investors in thousands. Individuals who traded at least one stock or one derivative contract on a year are considered active for that year. The left axis has the percentage monetary volume participation, which is calculated by the sum of the value of all purchases and sales of individual investors divided by the total value of purchases and sales in the stock market.

There are two relevant limitations of the data. First, we do not observe the social demographic characteristics of the investors. Thus, any heterogeneity analysis comes from information that we can extract from holdings and transactions, such as account size, diversification, turnover, or other trading patterns.

Second, we can only observe direct equity holdings. We do not observe other components of total wealth, such as cash holdings, investments in bonds, real state, or holdings of mutual funds (which account for part of the wealth in equity). Although total wealth and

---

[27] https://www.b3.com.br/data/files/D9/D6/36/B4/B26DE7106C31DCE7AC094EA8/Book_PF_ultimo%20tri%202021.pdf/.

equity wealth are positively correlated[28], we emphasize that this paper results on redistribution relate to wealth in risky assets, more precisely to direct holdings in equity. This study focuses on how trading, which is heavily concentrated on direct equity holdings[29], will affect this distribution. The data is very complete for transactions, with daily execution prices and derivative contracts.

Throughout this paper we refer to individual wealth in risky assets and wealth in equity market as the monetary value we can observe due to direct holdings. When classifying individuals by their wealth we rank them by i) the value of their holdings in the beginning of the period for existing accounts, and by ii) the maximum wealth in a three-year period for new accounts opened in the same year[30]. Hence, new accounts are classified by yearly cohorts. Classifying all new accounts in the same cohort lead to bias grouping, given that most of new accounts opened in the last year of the sample, and there would be less than one year of inflow to help the ranking.

Panel A of Table 1 shows that institutions have significantly larger holdings than individuals and are on average more diversified. This paper focuses on the heterogeneity within individuals, but from Table 1 we also observe a large heterogeneity within institutions. The average portfolio value for institutions is larger than the 90th percentile, which indicates a substantial concentration of holdings. Our data consists of every participant in the equity market, the surprisingly high number of trades of institutions is due to the inclusion of market makers and other high frequency trading (HFT) institutions in our data. Given the complexity of a derivative instrument, it also surprises that the proportion of institutions trading derivatives is close to the proportion of individuals trading derivative contracts.

---

[28] Brazil has no surveys or public data on total household wealth composition. Available aggregated tax data from the Brazilian IRS show that income, total wealth, and income from equity (dividends and other) are highly correlated. The data also shows a high concentration of wealth and equity income in the top income brackets. In 2020, the households in the top 1% in total earnings distribution, who earn above R$ 750 thousand a year (U$ 200 thousand), concentrated 22% of total income, but had 60% of the total equity income, and 32% of total wealth.

[29] Between 2012 and 2019, monetary flow in/out of mutual funds trading equity had an average volume of 19.7% of the average volume of direct trading by individual investors in the equity market.

[30] Choosing the maximum wealth in a two-year period or using new accounts three-year net inflow to equity market led to qualitatively equivalent results. We also perform robustness exercises on section 4, for new and old accounts, and with investors being classified by their size of the first trade.

Table 1 – Descriptive statistics by investor groups.

Investors group	**Wealth (Thousands R$)**			**# Stocks in portfolio**			**Trades (monthly)**			**Turnover (% monthly)**			**Trading derivatives (%)**	**Obs.**
	Mean	**Median**	**90 Pct.**	**Mean**	**Median**	**90 Pct.**	**Mean**	**Median**	**90 Pct.**	**Mean**	**Median**	**90 Pct.**		
Panel A														
Institution	24719	206	24352	6.1	2.0	14.6	1052	8.0	654	34%	11%	70%	17%	51150
Retail	95	6.3	87	3.5	2.1	7.7	14	2.0	20	85%	24%	200%	14%	1701459
Panel B - Within Retail														
Top 1%	5405	724	9521	7.0	4.6	15.5	135	27	268	100%	27%	253%	34%	16581
5th quintile	169	65	341	5.7	4.2	12.1	31	5.8	52	94%	27%	231%	23%	323350
4th quintile	29.6	17.7	59.0	4.1	3.0	8.6	13	2.7	23	88%	25%	214%	17%	341321
3th quintile	13.0	7.4	22.5	3.2	2.3	6.5	8.9	1.8	15	86%	25%	209%	13%	340613
2sd quintile	6.3	3.0	9.2	2.5	1.8	5.0	6.4	1.3	10	84%	23%	195%	10%	340466
1st quintile	2.0	0.5	2.5	1.7	1.0	3.0	3.6	0.7	6	75%	17%	152%	7%	339128

Note: For each investor and for the period they have positive direct equity holding, we calculate her time-series average of total value of portfolio (wealth), number of stocks in her portfolio, total number of trades in the month, and turnover (monetary volume of buy and sell orders divided by two times the average monthly holding). Table 1 then reports the cross-section averages and percentiles of these time-series averages for different subgroups of investors. Institutions include Firms, Banks, and other Asset managements, either domestic or foreigner. Retail investors are ranked i) by their total wealth in January 2012 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019).

Panel B of Table 1 shows statistics for six distinct groups of retail investors classified by their stock holdings' value, with the top 1% of the distribution excluded and shown separate from the top quintile. Wealthier investors are more active, more diversified and participate more in the derivatives market. There is also a large variation within the groups. The median wealthy investor has higher turnover than the median poorer investor, however, the 90th percentile of poorer investor have a much higher turnover than the median wealthy investor. This suggests that higher activity is widespread on various levels of wealth and diversification. On the next section we start the investigation on how trading transfers wealth and may affect concentration of wealth.

3. Empirical exercise: wealth transfer.

The main empirical exercise will estimate monetary transfers that happen in the stock and derivatives market through active trading. Investor trading will lead them to hold different portfolios and affect their performance. This could happen because of both market timing ability, with net flows in/out of equity market, and stock selection ability across individual stocks.

3.1 Wealth transfer - trading gains and losses.

We start by defining how to measure the gains and losses resulting from trading in the stock and derivatives market. We adjust and expand ALS (2022) monetary gain and losses measure to our setup, where we have derivatives market, trading execution prices, and are interested in explicit transaction orders rather than portfolio change. The general idea for this measure is to compare the wealth of an investor with a no-trade strategy (baseline) against the wealth of this same investor after trading.

The goal is to measure the total cumulative gain and losses from trading that an investor i had from day t up to day τ. We calculate this in three steps:

$$Cumulative\ Trading\ Gain_{i,t:\tau} = \begin{array}{c} 1)\ Stock\ Market\ Daily\ Gain_{i,t:\tau} \\ + \\ 2)\ Stock\ Market\ Intraday\ Gain_{i,t:\tau} \\ + \\ 3)\ Derivatives\ Market\ Gain_{i,t:\tau} \end{array} \qquad (1)$$

Concisely, the first step calculates how daily transaction orders affect the investor`s portfolio, assuming that the transaction was at closing prices. That is, given an order imbalance made by an individual on day t and stock s, we estimate how her wealth varies compared when she had no order imbalance. In the second step, we calculate the investor intraday monetary gains and losses using the difference between closing prices and her order execution price on the same day. The last step is to consider gains and losses from using derivatives contracts, here we calculate contracts premiums and the difference between stock closing prices and the stock price in the derivatives contract.

Specifically for the first step, investor i cumulative stock market daily gain from day t up to day τ is calculated by the cumulative sum of each stock s transaction $flow$ times its excess return in relation to the risk free from day t up to day τ:

$$Stock\ Market\ Daily\ Gain_{i,t:\tau} = \sum_{s} \sum_{t \leq \tau} flow_{i,s,t:\tau} \times excess\ return_{s,t:\tau} \tag{2}$$

such that $flow_{i,s,t}$ are the shares bought minus the shares sold in stock s on day t by investor i at the closing price on day t, an order imbalance. With *excess return* defined as:

$$excess\ return_{s,t} = \ total\ return_{s,t} - risk\ free_{t} \tag{3}$$

Where $risk\ free_t$ is the rate of return on the short-term treasury bond, the SELIC rate in Brazil. Using excess return instead of raw return has an important underlying hypothesis. It assumes that the opportunity cost of investing in any stock would be investing in short term bonds, and that cash from capital outflow of the equity market is invested in bonds[31].

To make things more concrete, consider an investor who buys a stock on day 1 and liquidates her position after one month, on day 21. The measure is then equal to the purchase value at closing prices times the stock excess return from day 1 to day 21. Thus, it measures her trading gains from investing in that stock, against the outside option of investing in a treasury bond. If that were her only trade in a year, the cumulative measure after those 21 days would still be the same. Now, consider a second example, where an

---

[31] For short periods of time this hypothesis should make minor difference. However, in an eight-year sample and in Brazil where the average annual short term interest rate was 9.6% for the period, considering that the opportunity cost of investing in the equity market was investing in treasury bonds appears more reasonable than considering that the opportunity cost was holding cash. We repeat the exercises of this section using raw returns for shorter intervals, results are qualitatively similar.

investor sells stock X on day 1 to buy stock Y also on day 1 (with both stocks having the same closing prices on that day), and again liquidates the entire position on day 21. After summing over the two stocks, the measure is then equal to the purchase value at closing prices times the difference between stock Y return and stock X return from day 1 to day 21. Thus, it measures her trading gains and losses from investing in stock Y, against the baseline of no-trade, which was to hold stock X.

This first measure alone does not account for gains and losses on the same day of the trade. Barber, Lin and Odean (2021) show that accounting for the same day can be important, with retail investors losing money on the day of trade. We take this into consideration and measure investor i intraday stock market gain from day t up to day τ as the following cumulative sum:

$$Stock\ Market\ Intraday\ Gain_{i,t:\tau} = \sum_{s}\sum_{t}^{\tau} \begin{matrix} (Closing\ Price_{s,t} - Execution\ Price_{i,s,t}) * Shares\ bought_{i,s,t} \\ + \\ (Execution\ Price_{i,s,t} - Closing\ Price_{s,t}) * Shares\ sold_{i,s,t} \end{matrix} \qquad (4)$$

where equation (4) can capture not only the gains and losses from the same day of the trade - difference between intraday execution and closing prices - but also, day trading gains and losses - when shares are both bought and sold in the same day. Day trading may be an important source of trading gain for institutions working with HFT and as market makers, while there is evidence that retail investors lose in this activity[32]. This measure complements the first one, and together they measure the total monetary trading gain and losses in the stock market.

The last step is to calculate trading gains and losses from the derivatives market. We observe trading on options contracts and forward contracts of individual stocks. To measure trading gains and losses from day t up to day τ we separate open and closed contracts on day τ. For open contracts, the buyer's result of a particular contract is the market value of the contract on day τ minus the premium paid for the contract, while the inverse value is attributed to the seller. For terminated options contracts, the buyer's result for call and put options are:

---

[32] Jordan and Diltz (2003) and Barber et. al. (2014) shows that individuals in general lose by day-trading, while Carrion (2013) and Baron et. al. (2019) shows that HFT, performed by institutions, is profitable.

Call option contract

$$- \text{Contract Premium} * Contracts$$

$$+ \begin{cases} (P_c - P_{strike}) * Contracts & if\,(P_c - P_{stike}) > 0 \\ 0 & otherwise \end{cases}$$

Put option contract

$$- \text{Contract Premium} * Contracts$$

$$+ \begin{cases} (P_{strike} - P_c) * Contracts & if\,(P_{strike} - P_c) > 0 \\ 0 & otherwise \end{cases}$$

where P_{strike} is the *strike* price on the contract and P_c is the underlying stock closing price at the expiration date (or at the termination of the contract for American options). Once again, the opposite value is attributed to the seller of the contract. As for terminated forwards contracts, the buyer's result is the difference between the closing price of the underlying stock on the expiration date and the contracted price, multiplied by the number of contracts. The result of each contract is summed and accumulated from day t up to day τ for each investor i.

For those investors who trade in the derivatives market, accounting for their gains and losses in all markets is essential to fully understand their performance. Investors may hedge taking various positions in the stock and derivatives. It could be the case that an individual is taking losses in the stock market but having gains in the derivatives market, or the opposite.

The three measures together give us an accurate estimate of the monetary quantity that each investor gained and lost in the stock and derivatives market. In a sense, the money estimate, which is close to a dollar-weighted return, should be more lifelike in representing individuals experience as investors (Dichev, 2007), and is what matters when studying wealth distribution[33]. Also, note that this measure does not depend on any assumption about the holding horizon of the investor and will measure investors monetary gains and losses until day τ against an initial buy & hold portfolio in t. Studies of retail performance usually assume an average holding period for their analysis which is a good hypothesis to estimate the average performance of a group of investors[34]. However, considering the actual holding horizon of the investor is key to reflect the wealth transfer through trading of each investor and estimate its impact on the distribution.

---

[33] We acknowledge that using excess returns to calculate the monetary trading gains and losses has the underlying assumption that the investor considers treasury bonds as his opportunity cost. On the one hand, monetary amount gain or lost with raw returns may better represent the experience of a naïve investor.

[34] An, Lou and Shi (2022) are an exception that considers the actual holding period of each investor. However, they do not account for same day return and derivatives market.

3.1.1 Retail investors lose.

First, we ask how individual investors as a group did on trading in the equity market. Table 2 shows that individuals lost a significant amount of money when trading against institutional investors, over 17.1 billion reais (U$ 5.7 billion). Given that we have data on all investors, the cumulative monetary trading results of institutions are a mirror of the results of individuals. Still, because individuals hold on average about 12% of all holdings, percentage returns differ drastically between the two groups.

Table 2 – Cumulative trading gains and returns for institutions and retail investors over an eight-year period (2012-2019).

Investors / Market			**Cumulative trading gains and return**			
			Institutions		**Retail**	
			Millions (R$)	Annual (%)	Millions (R$)	Annual (%)
Panel A						
All investors	All markets		17170	0.2%	-17170	-1.3%
Trade only in stock market	Stock Market		19696.7	0.3%	-10002	-1.3%
Trade also in derivatives market	All markets		-2527	-0.1%	-7167	-1.3%
		Derivatives	3461	0.1%	-3461	-0.7%
		Stock market	-5988	-0.2%	-3706	-0.6%
Panel B						
% of investors w/ negative trading result			61.3%		42.7%	
Ave. gain | Positive result (Thousand R$)			15798.6		38.2	
Ave. gain | Negative result (Thousand R$)			-9238.9		-74.5	
Ave. Wealth 2012-2019 (Millions R$)			1198853		164076	
# Investors			51150		1701459	

Note: Institutions include Firms, Banks, and other Asset managements, either domestic or foreigner. Monetary cumulative trading gains are calculated from all trades between 2012 to 2019 considering excess returns. Stock market gains combine daily and intraday gains measures, together with derivative gains they constitute the “All markets” cumulative trading gain. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative result are investors with the cumulative trading measure below zero at the end of 2019.

The annual percentage returns shown on Table 2 are an estimative of how much of the cumulative monetary gain or loss represents over the portfolio value. For each day we calculate the percentage return of the marginal gain or loss of that day, which is the difference between the cumulative gain of “today” and the cumulative gain of “yesterday”, divided by the portfolio value of “yesterday”. The annual percentage is calculated accruing the daily returns obtained from the marginal daily gain and loss.

Over an eight-year period, the amount invested in the equity market can drastically change, increasing in some periods and decreasing in others. This has an important implication for the monetary and return measures presented in Table 2. Periods when monetary holdings are larger will have bigger contributions to the total cumulative monetary measure[35]. But because the return measure is not dollar-weighted, returns will not be as affected. As shown in Figure 1, there is a substantial increase in retail investors in the last year of the sample, which also increased total holdings. Therefore, the last year made a bigger contribution to the cumulative monetary measure, but not to the return measure (see Figure A1 on the appendix).

Results reported on Table 2 labelled as “All markets” are the cumulative trading gain of all three factors described in the previous section. “Stock market” results on Table 2 include the first two factors of equation 1, daily and intraday gains. Retail lost R$ 2.7 billion due to same day return, about 20% of the total cumulative loss in the stock market. We separate investors who trade in the derivatives market from those who never did and show individually their result in each market. As a group, individual investors who traded in derivatives market lost in both markets and nearly 50% of their cumulative monetary loss comes from derivatives. Overall, they had a similar bad performance (in percentage terms) as investors which did not trade derivatives. However, when focusing only on the stock market result, retail investors which trade derivatives lose proportionally less. The disaggregation between markets also reveals that institutions which participate in the derivatives markets, although profiting from trading derivatives contracts, lose larger sums in the stock market and have aggregate negative monetary results, which may seem surprising if one expects that trading derivatives is a sign of sophistication of the institution.

---

[35] For example, suppose that in the first-year retail investor held 100 billion in equity and loss 1 billion in trading, resulting in 1% of their holdings. Now suppose that in the second-year retail investor held 300 billion and gained 2 billion in trading, which would be 0.67% of their holdings. The cumulative result of the two years would be positive in 1 billion for the monetary measure, but negative in 0.3367% for the return measure. Thus, years with higher monetary holdings and trading volume drive the monetary (dollar-weighted) result. The previous example calculated marginal trading gain/loss and returns in a yearly basis, but the same logic applies for the daily calculations used in this study.

Another surprising fact appears in Panel B of Table 2. We calculate trading gains and losses for each investor and report the proportion of investors with the cumulative trading measure below zero at the end of 2019. Panel B reveals that most individual investors had cumulative positive trading gains, and that most institutions lost by trading. Panel A results come from the fact that the few institutions who won by trading won larger sums than the many who lost, with the other way around for individual investors. This finding points out how grouped results for investors may be deceiving and not represent the true average experience of an investor in the stock and derivatives market.

Aggregate results on Panel A, however, endorse the already well-established finding that retail investors lose by trading. We add to this evidence showing that this is still true if one considers a dollar-weighted measure, the execution price of the trade, the trades in the derivatives market, and the actual holding period of the investor. When taking all those factors into account for an eight-year period, the evidence that individual investors (as a group) underperform on trading is strong. Also, losses due to same day return and trading derivates contracts accounted for 36% of the total measure, which suggests that results that do not account for those factors are only a lower bound for the true loss of retail investors.

The results on trading of Table 2 are on a comparative basis. A negative result does not mean that the participation of retail investors in the stock market is bad for them, but that active trading led to inferior performance and substantial monetary loss. The results reveal that individuals had R$ 17.1 billion less in the end of 2019 than they would have if they simply held their initial portfolios from the end of 2011 and did not trade. Despite their trading losses, individuals could well have monetary profits from participating in the equity market. In fact, the estimative of retail annual excess return was 3.94%, which amounted to R$ 57.9 billion on gains between 2012 to 2019. Table 2 then suggests that the gains could have been larger, R$ 17.1 billion larger.

3.1.2 Within retail, poor investors lose more.

Monetary results are more representative of the investors with a high level of wealth and trade volume in the equity market. Data shows that there is an extremely high inequality in equity holdings in Brazil, with the top 1% of the distribution holding 60.3% of retail wealth on equity. Thus, wealthier investors drive the results reported on the previous section and we now ask how diverse groups of individual investors did on trading in the equity market.

Panel A of Table 3 disaggregates the results of individuals of Table 2 into six different wealth groups. Investors are separated into five different quintiles, with the top 1% of the distribution excluded and shown separate from the top quintile. We rank existing accounts by their total wealth at the end of 2011, and new accounts opened in the same year by their maximum wealth in a three-year period[36]. As seen in Table 1, more than half of the total number of accounts are new accounts opened after 2011. Section 4 later investigates the dynamics of the results separating old from new accounts.

Table 3 shows that there is a large heterogeneity of results within retail investors. Individuals in the bottom quintile of wealth hold less than 0.5% of the retail wealth but have the worst overall performance when trading, having an average loss of 2% of their holdings a year. From the 4th to the bottom quintile, overall trading results monotonically get worse. However, the wealthier quintiles, especially the extraordinarily rich (Top 1%), not only lose by trading, but have proportionally worse results than the investors in the middle quintiles. This result seems in contrast with Li et. al. (2017), who find that very wealthy investors outperform when trading, and with the ex-ante conception that wealthier investors are more sophisticated.

Panel B shed light on this apparently strange and conflicting result. Few investors who lose large sums of money drive the aggregate results for the very rich. In fact, Panel B reveals that most individual investors have cumulative positive trading gains within the groups of wealth. Only the bottom quintile has more investors experiencing monetary trading losses than investors experiencing gains over this eight-year period. Panel B reports that when investors lose by trading, they lose larger amounts of money than the investors who win by trading, and this is reflected in the aggregate result. This finding again stresses how volume weighted results can differ from equally weighted ones when studying retail investors.

---

[36] As in ALS (2022), we chose the maximum wealth over a period after the account opening to classify new investors. Robustness exercises in section 4, show that older accounts that do not depend on this classification drive the monetary results. Nevertheless, the Appendix has a replication of Table 3 when ranking new accounts by their three-year net inflow to equity market (Table A1).

Table 3 – Cumulative trading gains and returns within retail investors over an eight-year period (2012-2019).

Investors / Market		**Cumulative trading gains and return**											
		***Top* 1%**		**5th Quintile**		**4th Quintile**		**3th Quintile**		**2nd Quintile**		***Bottom* Quintile**	
		Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)
Panel A													
All investors	All markets	-13482	-1.6%	-3154	-1.0%	-228	-0.6%	-105	-0.7%	-109	-1.1%	-92	-2.0%
Trade only in stock market	Stock Market	-6853	-1.5%	-2340	-1.1%	-369	-0.8%	-151	-0.9%	-205	-1.7%	-84	-2.5%
Trade in derivatives market	All markets	-6629	-1.6%	-813	-0.8%	141	-0.2%	46	-0.5%	95	0.2%	-8	-1.4%
	Derivatives	-1202	-0.4%	-1324	-1.1%	-431	-1.9%	-268	-2.2%	-145	-2.4%	-92	-3.4%
	Stock market	-5427	-1.3%	510	0.2%	572	1.8%	314	1.8%	240	2.7%	84	2.1%
Panel B													
% of investors w/ negative trading result		36%		33%		37%		42%		48%		54%	
Ave. gain | Positive result (Thousand R$)		1354		73		16.9		8.7		5.6		2.4	
Ave. gain | Negative result (Thousand R$)		-4550		-175		-30.9		-12.9		-6.8		-2.5	
Ave. Wealth 2012-2019 (Millions R$)		98978		48490		9254		4282		2295		778	
# Investors		17316		328823		346024		345733		346351		346405	

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Monetary cumulative trading gains are calculated from all trades between 2012 to 2019 considering excess returns. Stock market gains combine daily and intraday gains measures, together with derivative gains they constitute the “All markets” cumulative trading gain. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019. The opposite signals between the monetary and return measure happen because the money measure is dollar-weighted and that wealth in equity fluctuated during the eight years, with a sharp increase in the last year. A positive result in the end years of the sample, when investors had larger holdings, may shift the monetary result to a different sign of the return result (see section 4.1).

The large concentration of holding on only a few individual investors also happens on the trading volume data. Results reported on this section, which show a broad difference between investors with different wealth levels imply that results for individual investors obtained with aggregated and volume-weighted data deserve a caveat. Several important results on the individual investor literature were found using volume-weighted data, Kaniel, Saar and Titman (2008), Kelley and Tetlock (2013) and Boehmer et. al. (2021) are some influential examples. This study suggests that empirical results with volume-weighted data may be less representative of all individual investors and much more representative of a small group of wealthy investors than one would think[37]. Overall, disaggregated results show that quintiles of investors with lower wealth have a higher proportion of investors having a negative experience when trading.

An interesting pattern arises when disaggregating the results of retail investors who trade in the derivative markets. Excluding the Top 1% of investors, who drive the results of Table 2, investors who trade derivatives contracts have better overall results. This is in line with associating participation in the derivatives market with higher financial sophistication. Individual investors lose when trading derivatives contracts in all categories of wealth, with poorer investors doing significantly worse, over 2 p.p. worse than the 5th quintile. However, their result in the stock market is generally positive, which raises the question if the bad result on the derivatives market is because of the complexity of the instrument or if its due to hedging, taking losses on that market, while wining on the stock market.

In summary, Table 3 shows that more investors with lower wealth have worse results when trading than investors with higher wealth. Also, the bottom quintile of wealth had the worst overall return performance. This indicates a wealth transfer from poor to rich through trading. Yet, Table 2 showed that on aggregate institutions are the ones winning money when trading. It could be that investors classified with lower wealth by their direct holdings have proportionally more wealth in equity through indirect holdings (institutions). If that were the case, there would be no wealth transfer, because poorer investors would be profiting proportionally more through institutions.

---

[37] On a similar rationale Barber, Lin and Odean (2021) conciliate the evidence that retail order imbalance positively predicts returns, but retail investor trades lose money. The former evidence is obtained using equally weighted tests over the stocks, while the latter is obtained by using volume weighted tests. Just like results may differ when using equally or volume weighted over the stocks, results could differ when using equally or volume weighted over retail investors.

Two evidence are contrary to the former argument. First, 41% of institutions trading gains come from foreign funds. This reveals a clear wealth transfer from domestic investors to foreign investors. If any domestic investor has an indirect holding through international foreign funds, they would probably be at the top of the wealth distribution because of the fixed cost to do so. Second, aggregate data from the Brazilian IRS does not support the argument that investors with less direct holdings in equity would have proportionally more indirect holdings in equity through funds.

3.1.3 Retail investors had good market timing.

An interesting exercise is the decomposition of the results in the stock market into market timing ability or stock selection ability. With modifications in equation 2 and 3, it is possible to further decompose the stock market gain into market timing and stock selection. For market timing, we substitute stock excess return in equation 2 for $market\ excess\ return_{t:\tau}$. Then, to get the hypothetical gains and losses from market timing, i.e., gains from net flows in/out of equity market, equation 3 becomes:

$$market\ excess\ return_t = \ market\ return_t - risk\ free_t \ \ (5)$$

which does not depend on stock s anymore, all trades have the same benchmark market return. Thus, if an investor changes its position from stock X to stock Y in the same amount of money (with no net flow), both stocks will have the market return and the cumulative trading result from this trade will cancel out. The measure only changes due to net flows.

For stock selection, stock excess returns relative to the risk free in equation 3 is replaced by excess return relative to market return. To get the hypothetical gains from stock selection ability equation 3 becomes:

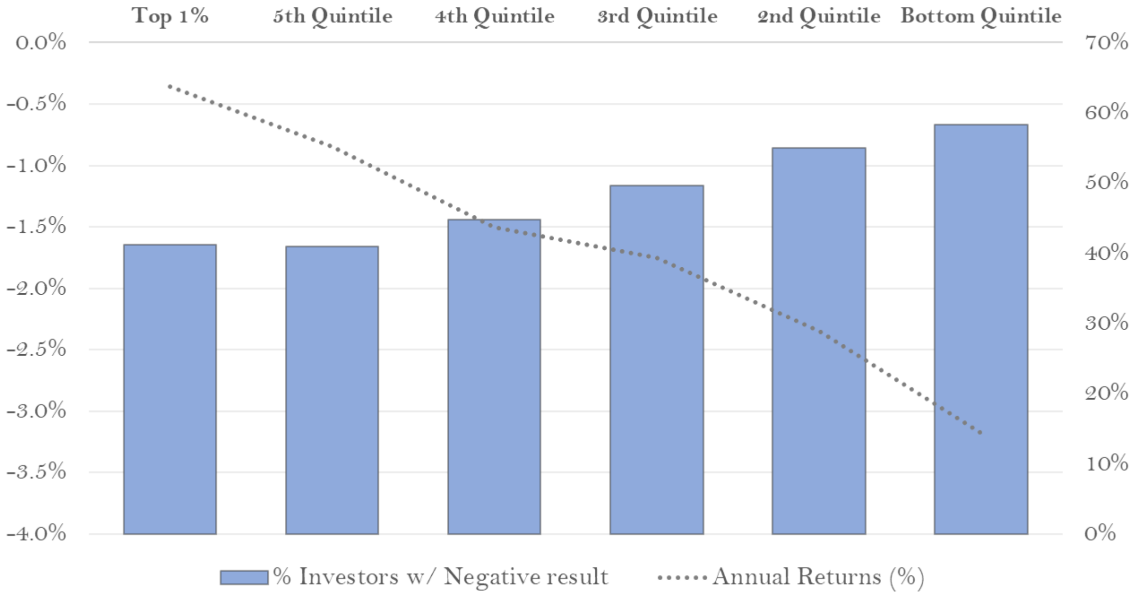
$$excess\ return_{s,t} = \ total\ return_{s,t} - market\ return_t \ (6)$$

which compares how well the new trade did against the market. Note that if an investor changes its position from stock X to stock Y in the same amount of money, the cumulative trading result from that trade is the same as when using equation (3).

Figure 2 shows for each retail wealth group the proportion of investors with negative results and the annual returns when excluding market timing gains and losses.

We calculate market timing using equation (5). Overall, retail investors had positive market timing, especially lower quintiles groups. Thus, when excluding market timing, and focusing on stock selection, we have more investors having negative results and even lower annual returns, with a clearer monotonicity between wealth groups.

Figure 2 – Proportion of investors with negative results and annual returns excluding market timing - within retail investors over an eight-year period (2012-2019).



Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Annual returns (%) and the percentage of investors with negative result comes from the cumulative monetary trading gains and losses, calculated from all trades between 2012 to 2019 considering excess returns equation (3), and discounting market timing equation (5). Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss, considering Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019. calculate market timing using equation 5 and get the stock selection as a residual

3.2 Monetary results under different measures.

On this section we address how much other measures of monetary outcome may leave on the table when not considering same day return, derivatives market and the actual holding period of the investor. We compare our measure with ALS (2022) measure and Barber, Lee, Liu and Odean (2009), BLLO from now on. Our measure is based on ALS (2022), with both considering the actual holding period of the investor. However, we also consider the actual execution price of the trade and derivatives contracts. Trading flows from ALS (2022) are calculated from the difference in shares from one day to another in the investor portfolio, considering closing prices. This means that the same day return is aways set to zero. Also, when a derivatives contract expires and shares are exchanged, ALS (2022) will calculate the flow value by the closing prices, while we consider premiums paid and the difference between closing prices and contract price to get monetary loss and profit for the derivative market.

Table 4 – Different measures outcomes of monetary gains, annualized returns, and proportion of investors with negative trading outcome - accumulated (2012-2019).

		Cumulative trading outcome		
		BPR	**ALS**	**BLLO**
All	Millions (R$)	-17170	-10951	-12272
	Annual (%)	-1.3%	-0.8%	-1.6%
	% of investors w/ negative trading result	42.7%	38.4%	30.5%
***Top* 1%**	Millions (R$)	-13482	-11520	-5110
	Annual (%)	-1.6%	-1.3%	-0.8%
	% of investors w/ negative trading result	36.1%	30.9%	30.0%
5th Quintile	Millions (R$)	-3154	-498	-4984
	Annual (%)	-1.0%	-0.3%	-1.5%
	% of investors w/ negative trading result	33.4%	29.4%	30.1%
4th Quintile	Millions (R$)	-228	570	-1165
	Annual (%)	-0.6%	0.5%	-1.9%
	% of investors w/ negative trading result	36.7%	32.7%	30.3%
3th Quintile	Millions (R$)	-105	323	-551
	Annual (%)	-0.7%	0.5%	-1.7%
	% of investors w/ negative trading result	41.6%	37.3%	31.0%
2nd Quintile	Millions (R$)	-109	133	-304
	Annual (%)	-1.1%	0.2%	-2.0%
	% of investors w/ negative trading result	47.7%	43.2%	31.3%
***Bottom* Quintile**	Millions (R$)	-92	41	-159
	Annual (%)	-2.0%	0.0%	-2.4%
	% of investors w/ negative trading result	53.9%	49.6%	29.6%

Note: Table 4 shows the cumulative monetary result, annualized returns, and proportion of investors with negative trading outcome for different measures. BPR is the benchmark measure of this paper (Table2 and 3), considering derivatives market, same day return and true holding period of the investor. ALS is An, Lou and Shi (2022) measure and BLLO is Barber, Lee, Liu and Odean (2009) measure, both applied to our data sample. ALS calculates flows from the difference in shares from one day to another in the investor portfolio, considering closing prices. BLLO measure is the difference in gains from a buy and sell portfolio, one that mimics the net daily purchases and one that mimics the net daily sales. Shares are included in the portfolio for 248 days. For BLLO returns are calculated considering the portfolio constructed only with new trades, not the actual portfolio which considers stock holdings the investor already had.

BLLO (2009) construct a buy and sell portfolio for each investor group, one that mimics the net daily purchases and one that mimics the net daily sales, the difference between buy minus sell portfolios gives the total monetary outcome. They consider the execution price and account for the same day return, but not for the derivatives market or the actual holding period. Shares are included in the portfolio for a fixed horizon, we use 248 days (results for other horizons are left to the appendix, Table A2). On the original paper, gains are compared with the market portfolio, but we adjust to compare with the risk free.

Table 4 compares the cumulative monetary result considering our measure (BPR), ALS and BLLO measures. Accounting for the same day return and derivatives led to losses that were 57% greater, compared to ALS measure. Also, the wealth transfer from rich to poor is not clear, investors with lower wealth have worse performance in the derivatives market and with same day return. When comparing to BLLO measure, Table 4 shows that monetary loss is also smaller when not considering derivatives and actual holding period,. Loss can be smaller because retail investors underperform in the derivatives market (Table 2), and because retail investors' true purchase and sale timing is worse than when considering a fixed period.

4. Trading gains dynamics.

Table 2 to 4 reported the cumulative trading result from 2012 to 2019 for all investors and for investors divided in groups of wealth. The cumulative result was relative to a benchmark of no trading that was set in the beginning of the period, 2012. On this section we add to this investigation by looking at the cumulative trading gains from different starting points, each year under study. We also break down results for old and new accounts to study the performance differences between the two groups and their time-series dynamics. Newer accounts trade more and have lower wealth in equity markets. It is interesting to study how different cohorts of new investors are doing in the equity market, and if there are any trends in their behavior[38].

Additionally, we focus on new investors and sort them by their first month inflow to the stock market. An endogeneity concern for our results is that two investors could have the same wealth, but direct different shares of its wealth to equity holdings based on their ability/experience in the stock market. Consequently, an investor with higher ability would also have higher wealth in equity holdings. This can partially drive the results of

---

[38] Unfortunately, we do not have information on the social demographics that could help us understand their social background. However, we know from B3 stock exchange reports that new investors are much younger on average.

section 3, but it is hard to argue that on average an investor with one thousand reais in stocks has the same total wealth as an investor with ten million reais in stocks. Regardless, using new investors' first month net inflow to classify investors should help alleviate the mentioned problem. It is expected that investors with higher wealth have higher first month investments on the stock market, while at the same time their ability is not yet revealed for themselves. We find that investors with smaller first month net inflows have *ex-post* worse results.

4.1 Old and new accounts through time.

The cost of directly participating in the stock market has been decreasing for decades because of technology. With the increase in popularity of commission-free brokerage apps in recent years, investors participation and activity are even higher. With that in mind, we take a closer look at new investors dynamics. New investors are expected to have inferior performance when compared to older investors. Seru, Shumway and Stoffman (2010) show that experience and learning make an investor get better at trading with time, and investors which have inferior performance leave the market. We investigate how new investors from recent years did on trading compared to new investors from older years.

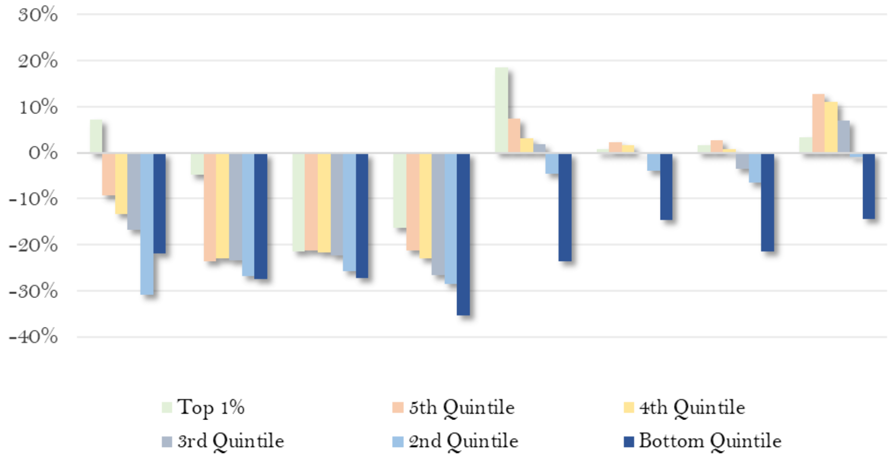
Figure 3 reports annualized returns from cumulative trading results when breaking down by each year, wealth group, and new vs. old investors. Each year we regroup individuals into wealth categories. We rank existing accounts in the end of the previous year by their total direct holdings value, and rank accounts opened within the same year by their maximum wealth in a three-year period[39]. Thus, the initial position of each year is the benchmark portfolio of no trade. Figure 3 shows that the bad monetary and return performance of retail was persistent, most retail investor groups had cumulative loss when trading in the eight years under study.

---

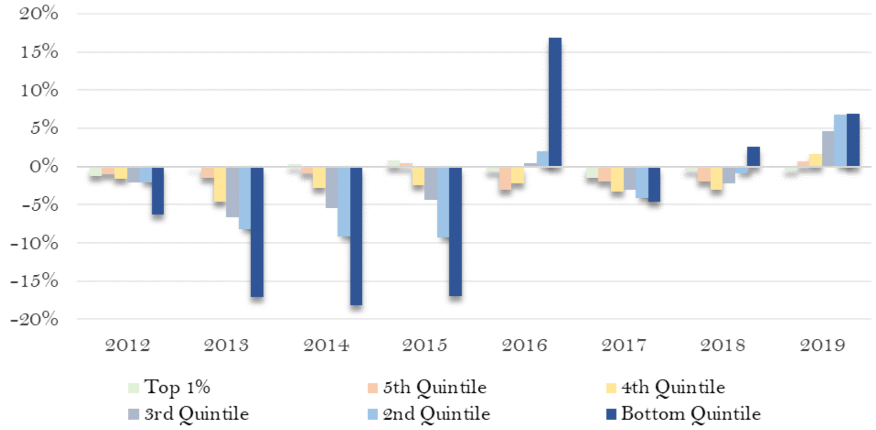
[39] For accounts opened in 2017, 2018 and 2019 we classify them by quarterly cohorts, given that they have less than three-year of data to accumulate the net inflow to capture maximum wealth in the period.

Figure 3 – Annualized returns from monetary cumulative trading gains by year (2012-2019) and group of investors.

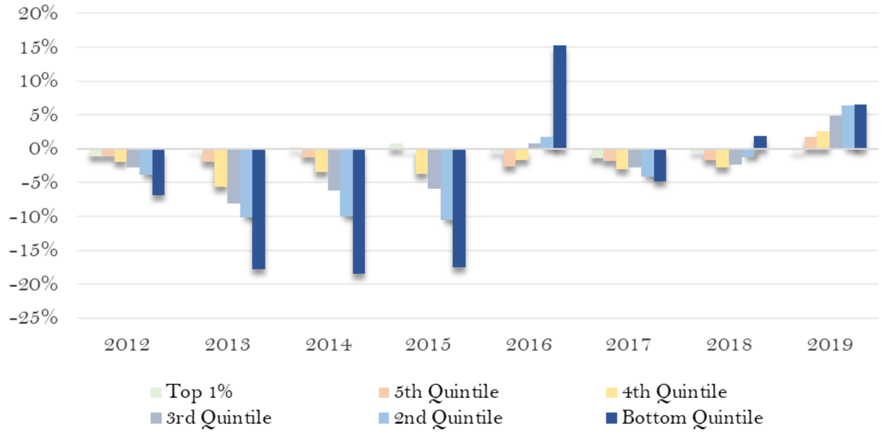
A – New Accounts



B – Old Accounts



C – All Accounts



Note: Figure 3 reports annualized returns calculated from monetary cumulative trading results when breaking down by each year, wealth group, and new vs. old investors. New investors have accounts opened within the same year of analysis. Monetary cumulative trading gains are calculated from all trades within a year considering excess returns. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. In the beginning of each year investors are regrouped into wealth categories. Existing accounts are ranked by their total direct holdings value in the end of the previous year, and accounts opened within the same year are ranked by their maximum wealth in a three-year period - for accounts opened in 2017, 2018 and 2019 the classification of new accounts is done by quarterly cohorts, given that there is less than three-years to estimate maximum wealth.

Figure 3A shows results for accounts opened within the same year of analysis. New investors had a worse performance on average than old investors (Figure 3B), with larger losses in percentage terms. Figure 3A reveals that new investors from more recent years had better performance than new investors from the first half of the sample. The equity market had higher returns in the second half of the sample, thus “market timing ability” of these new investors drove part of this result. When removing the gains from market timing, the difference between the performance of new investors on the first and second half of the sample reduces, but new investors after 2015 continue to show better performance (Figure A2 on the appendix). Figure 3A also reveals that the only two groups of new investors that had negative trading results in all years are the ones classified with low levels of wealth. Specifically, new investors in the bottom quintile lose on average 25% of their holdings in their first year of trading.

Figure 3B and 3C are almost identical, given our methodology of focusing on monetary gains, old investors which have larger direct holdings will drive the results when considering all accounts. Overall, Figure 3 reveals that retail investors lose by trading in most of the subgroups reported. The evidence that investors with lower wealth do worse and transfer wealth to the wealthy is very persistent for new investors, but less so when accounting for all accounts. The first half of the sample has a clear monotonic relation between wealth and performance, but this pattern does not repeat itself after 2015, with poorer investors doing better in three years even when discounting “market timing” due to higher market returns on recent years (Figure A2 on the appendix).

The remarkably high cumulative trading gain of the bottom quintile of wealth in 2016 is interesting. We attribute this outlier result to three factors. First, 80% of the trading gain of the bottom quintile of wealth came from market timing in 2016. That was the year with the highest equity market returns in our eight-year sample, when the Ibovespa Index rose 38.9%. Poorer investors had significant positive net flow in that year. Second, and related to the first point, investors with lower wealth have portfolios with higher betas in relation to the market. The bottom quintile had a portfolio beta of 1.11, while the top 1% in wealth had a beta of 0.9 (Table A3 in the Appendix). The third factor is the investment style of investors with lower wealth. Section 5 shows that investors with lower wealth have a contrarian investment style (Table 8), buying stocks which had relatively bad returns in the recent past. During our eight-year sample, 2016 was the only year when a contrarian investment strategy was on average[40] better than a momentum strategy, with an excess return in relation to Ibovespa Index of 24.3%.

4.2 Reducing endogeneity concerns.

To evade the endogeneity problem caused by investors adjusting their wealth share in equity based on their experience in trading in the stock market, we focus on new investors who have not yet known their ability when they enter the stock market.

Every month we sort new investors into groups of wealth based on their accumulated net inflow to the stock market from their first trading day until the 30 days after. Investors with higher inflows most likely have higher total wealth, while they still do not know how skillful they are in trading. We follow those groups for the next twelve months, calculating cumulative monetary trading gains and losses of each group formed in each month. Figure 4 shows the time-series averages of annualized returns and proportion of investors with negative returns for each group-month (January 2012 to December 2018). We have 508,296 new investors from 2012 to 2018. Having 6,015 new investors a month on average, with the maximum in October 2018 with 23,267 new investors, and the minimum in June 2014 with 1,430 investors.

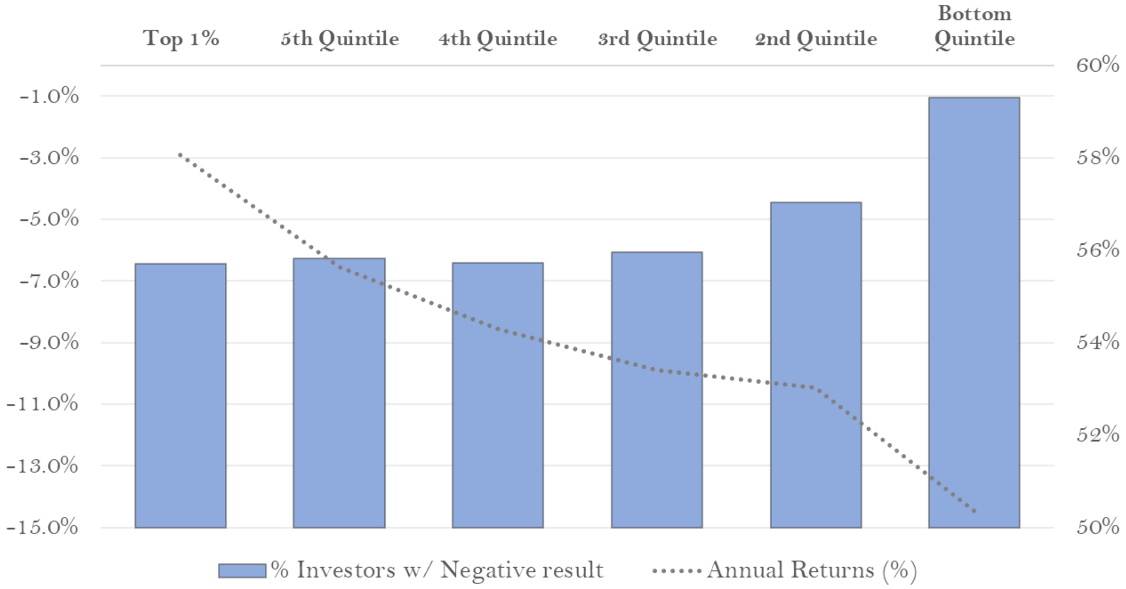
Figure 4 shows that new investors which have *ex-ante* higher net inflow to the stock market have *ex-post* better performance for the next twelve months. The monotonicity is very clear for the average annualized returns, with poorer investors, proxied by initial inflow, doing significantly worst in their first year. Monotonicity is less visible for the proportion of investors with bad experience, but we still have that the bottom quintile has more investors experiencing loss of money. Again, as in the previous section, new investors have the worst performance overall.

In summary, this section showed that section 3 findings are robust to different subsamples of time and group of investors. Next section studies the heterogeneity of the results, presenting relations between account size, different investment styles and wealth transfers.

---

[40] For each year in the sample, we back test all 40 variations of momentum and all 40 variations of contrarian strategies reported in Table 7 and get their simple average for contrarian and for momentum for each year.

Figure 4 – Time-series average of the proportion of investors with negative results and annual returns for new investors.



Note: Investors' group are defined monthly by their accumulated net inflow to the stock market after 30 days of their first trade. For each month only new investors are sorted into groups. For each group month-wealth we calculate annual returns (%) and the percentage of investors with negative results from the cumulative monetary trading gains and losses, calculated from all trades between the first month they traded and the next twelve months. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss, considering Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of the following twelve months. Figure 4 shows the time-series average from 2012-1 to 2018-12 of annual returns (%) and proportion of investors with negative returns.

5. Investment styles and wealth transfers.

This section investigates how two common empirical traits of individual investor behavior - high turnover, and under-diversification - relate to wealth transfer between and within groups. We also investigate whether a contrarian trading strategy can help explain the heterogeneity in retail losses.

5.1 Investor activity, diversification, and trading results.

Retail investors' high activity is arguably the most robust empirical finding about their behavior in the stock market. Not only that, but high activity is frequently considered the "villain" responsible for retail underperformance (Barber and Odean 2000, and Barber and et al 2009). Another common empirical finding is the under-diversification of individual investors. Goetzmann and Kumar (2008) and Campbell, Ramadorai and Ranish (2019) reveal that lack of diversification is associated with low levels of wealth and with underperformance. On the other hand, studies such as Ivković, Sialm and Weisbenner (2008) and Van Nieuwerburgh and Veldkamp (2010) argue that under-diversification can be a result of superior information, and investors with concentrated portfolio instead outperform.

Table 5 reports the proportion of investors classified based on their relative level of wealth, diversification (number of stocks in portfolio), and trading activity (turnover). Investors are divided into three categories (terciles) of diversification, trading activity and the same six wealth categories as before. Like in Goetzmann and Kumar (2008) and Anderson (2013), results show that low levels of wealth and under-diversification are positively correlated traits among investors, only 8% of investors in the bottom quintile are in the top tercile for diversification. However, different from the previous studies, our data does not show a negative pattern between activity and wealth. The richest Top 1% has more investors classified with high turnover than the bottom quintile, but, overall, investors terciles of activity are well distributed among groups of wealth and among groups with different diversification.

Panel C of Table 5 informs about the terciles thresholds and remind that the division of groups are in relative terms. An investor classified as having high diversification could have had as little as 4 stocks in her portfolio on average. An investor classified as low activity may have a portfolio turnover higher than 100% in a year. Further, our diversification measure, although popular in the literature, is naïve. It does not look at the weights given to stocks in the portfolio or the correlation among them.

On an attempt to understand how trading activity and under-diversification relate to our results within and between each group of wealth, the next exercise uses the same grouping as Table 5 and estimate returns and the proportion of investors losing money within each group based on cumulative monetary gain. The idea is to observe if among distinct groups of wealth, investors have similar performance if we account for their diversification and trading activity. It might be the case that investors with higher wealth, but lack diversification, do as poorly as investors with lower wealth in the same diversification group.

Table 5 - Proportion of investors classified by terciles of diversification and turnover.

Panel A - Proportion within each wealth group of the intersection of wealth, diversification and activity

	***Top* 1%**				***Top* Quintile**				**Quintile 4**			
	Less Diver.	Medium Diver.	High Diver.		Less Diver.	Medium Diver.	High Diver.		Less Diver.	Medium Diver.	High Diver.	
Less Activity	7%	7%	18%	32%	5%	6%	18%	29%	7%	8%	16%	32%
Medium Activity	4%	6%	21%	30%	4%	8%	24%	36%	6%	11%	18%	35%
High Activity	5%	12%	21%	38%	6%	12%	17%	35%	8%	14%	12%	34%
	16%	24%	60%		15%	26%	59%		21%	33%	46%	

	Quintile 3				**Quintile 2**				***Bottom* Quintile**			
Less Activity	10%	11%	12%	33%	15%	12%	8%	35%	25%	10%	4%	39%
Medium Activity	9%	13%	12%	33%	12%	13%	8%	32%	18%	10%	3%	31%
High Activity	10%	15%	8%	34%	14%	14%	5%	33%	20%	9%	2%	31%
	29%	38%	33%		41%	39%	20%		62%	30%	8%	

Panel B - Observations, investors in each intersection

***Top* 1%**				***Top* Quintile**				**Quintile 4**			
1149	1177	3174		16364	19753	57900		24223	28592	56308	
645	980	3614		14160	25315	78948		20570	37748	61473	
913	2026	3638		20160	40065	56158		27260	48904	40946	
			17316				328823				346024

Quintile 3				**Quintile 2**				***Bottom* Quintile**			
34659	37159	42382		52102	41068	26885		85131	35508	13292	
29519	43275	42587		40225	45432	26256		61266	34966	9847	
36104	51074	28974		48312	48551	17520		68674	31788	5933	
			345733				346351				346405

Panel C - Terciles threshold and median investor

	1st Tercile	2nd Tercile	3rd Tercile
Diversification Ave. number of stocks in portfolio	Below 1.4 stocks, w/ 1.0 as median	Between 1.4 and 3.3 Stocks, w/ 2.1 as median	Above 3.3 Stocks, w/ 5.7 as median
Activity Ave. monthly turnover (%)	Below 12% turnover, w/ 3.8% as median	Between 12% and 49% of turnover, w/ 25% as median	Above 49% of turnover, w/ 117% as median

Note: Table 5 reports the proportion of investors that are classified based on their relative level of diversification and trading activity within each wealth category. Investors are divided by the average number of stocks they had throughout the sample into three groups (terciles) of diversification. Investors are also divided by their average monthly turnover (%) into three groups (terciles) of activity. Turnover is measured by the monetary volume of buy and sell orders divided by two times the average monthly holding. Panel A presents the proportion of the intersection between diversification and turnover terciles within a wealth group. Panel B gives the total number of investors for each intersection and the total amount for each group of wealth. Panel C reports the threshold of each tercile of diversification and activity measure.

The intersection of investors with low diversification and high trading activity had massive losses for all wealth categories. Table 6 reveals that investors that have lower diversification groups have in general the worst results, regardless of wealth. In fact, all groups classified in the bottom tercile of diversification had negative returns from trading. Surprisingly, investors with lower wealth, but that are less active and more diversified have better results than any other group, which could be an indication that the under-performance of investors with lower wealth might be associated primarily to its lack of diversification. Observations are reported on Table 5, Panel B.

Panel B of Table 6 reports the proportion of investors losing money within each group. Again, the results suggest that groups with low diversification are the ones with the highest proportions of investors losing money. However, the proportion of investors with negative results in groups of investors with low activity are greater than the high activity group. This is different than the negative pattern found between activity and return in Panel A. Together, these results inform that some highly active investors lose considerable amounts of money, which drives panel A relation between activity and return. On the Appendix we evaluate the robustness of these results by first dividing investors into groups of similar wealth, and then double-sorting investors into various categories of trading activity and diversification, results remain qualitatively the same.

Table 6 – Annual returns and proportion of investors with negative monetary result by the intersection of wealth groups, terciles of turnover, and number of stocks in portfolio.

Panel A - Group returns (Year %)

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Low Activity	-1.6	-1.7	-1.1	-2.1	-2.0	0.5	-4.4	-1.6	1.8
Medium Activity	-0.8	-4.0	-1.3	-5.9	-2.9	0.8	-8.6	-1.3	1.3
High Activity	-11.7	-10.7	-3.8	-18.6	-5.9	-1.2	-18.6	-6.1	1.1

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Low Activity	-6.1	-1.1	2.5	-11.9	-0.3	3.1	-14.1	1.3	3.8
Medium Activity	-10.5	-0.1	2.4	-13.4	0.9	1.7	-10.2	0.7	2.6
High Activity	-22.0	-6.0	-0.4	-14.7	-5.5	1.4	-22.1	-5.6	-1.7

Panel B - Percentage of investors within a group with negative return

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Low Activity	56%	53%	46%	58%	53%	35%	62%	49%	30%
Medium Activity	37%	27%	15%	47%	30%	15%	49%	30%	18%
High Activity	53%	48%	29%	56%	44%	27%	54%	44%	30%

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Low Activity	64%	48%	30%	66%	48%	31%	69%	45%	38%
Medium Activity	50%	32%	22%	52%	36%	26%	49%	41%	35%
High Activity	54%	46%	35%	56%	50%	40%	57%	53%	48%

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors are divided by the average number of stocks they had throughout the sample into three groups (terciles) of diversification. Investors are also divided by their average monthly turnover (%) into three groups (terciles) of activity. Turnover is measured by the monetary volume of buy and sell orders divided by two times the average monthly holding. The total number of observations within each group is reported on Table 5, Panel B. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative result are investors with the cumulative trading measure below zero at the end of 2019.

5.2 Contrarian trading pattern and trading losses.

Now we investigate whether the contrarian trading pattern of individual investors can help explain the heterogeneity in retail losses. We take a popular empirical approach of labeling individual investors as contrarian (or momentum) based on their response to past returns[41]. In short, a momentum trading pattern would be buying shares of a company when their price is increasing, and a contrarian behavior would be buying when their price is falling. We study if this trading pattern correlates with investors wealth, and if trading losses of individuals concentrate among the investors who trade as contrarians.

Momentum strategies have been widely popular in academic papers and among investment professionals. The strategy has been shown to be profitable in different asset classes and markets (Asness, Moskowitz and Pedersen, 2013). However, the robust empirical evidence that momentum strategies are profitable does not translate to any investor having a momentum behavior will profit, and that any investor having a contrarian behavior will lose. Different from momentum strategies in academic papers, individual investors rarely short sell, do not buy diversified portfolios, and have inconsistent rebalancing periods, which leads to different holding periods for each time they buy a stock.

Table 7 presents the annual excess return for different momentum and contrarian strategies during the eight-years of the sample. We show strategies with different holding periods and using different past return intervals to rank the stocks into a portfolio. All strategies are long-only, which is closer to an individual investor reality. Table 7 reports excess returns relative to Ibovespa Index returns. On the left side of table 7 are the returns of strategies which buy stocks who had worse past returns, loser stocks, and we label them as contrarian strategies. In panel A, stocks are sorted into losers or winners in relative terms, stocks below the median in the cross-section are considered losers. In panel B, stocks are sorted by their own absolute past return, in the time-series, with stocks with return below zero considered as losers. Table 7 portfolios are equally weighed.

Table 7 shows that between 2012 and 2019 momentum strategies were more profitable than contrarian strategies. Using the cross-section to sort the stocks (panel A),

---

[41] Evidence of contrarian behavior by retail investors are found in Choe, Kho, and Stulz (1999) for Korean investors, in Grinblatt & Keloharju (2001) for Finish investors, in Kaniel *et. al*. (2008) for the US, and Chague *et. al*. (2018) for Brazil.

we find that 19 out of the 20 different combinations of momentum strategies had higher excess returns in comparison with its contrarian analogue. The difference between the simple average among all contrarian and the simple average of all momentum strategies was -6.7% a year. Panel B reports qualitatively equivalent results as Panel A. If investors have any sort of tilt towards either strategy, these results suggest that investors who tilted their strategy towards contrarian may have had lower returns. On unreported results, we observe the same qualitative results when using excess returns relative to the risk-free, and value weighting the portfolios.

Table 7 – Annual excess returns for different contrarian and momentum strategies between 2012 to 2019, long-only and equally weighted portfolios.

Past return interval / Rebalancing period		**Contrarian**				**Momentum**			
		Monthly	Quarterly	Bi-Annual	Yearly	Monthly	Quarterly	Bi-Annual	Yearly
Panel A									
Cross-Section	(t-21):t	-0.7%	2.4%	3.4%	2.6%	8.2%	4.8%	2.8%	3.3%
	(t-62):t	-1.2%	2.0%	1.2%	0.2%	7.6%	4.3%	4.9%	5.6%
	(t-125):t	-2.5%	-1.6%	-0.6%	-0.2%	9.0%	7.9%	6.6%	5.9%
	(t-251):t	-1.6%	-2.0%	-0.7%	0.4%	7.2%	7.5%	6.3%	4.6%
	(t-754):t	-1.0%	-0.3%	0.1%	1.5%	7.0%	6.2%	5.5%	4.4%
Panel B									
Time-Series	(t-21):t	-1.7%	1.9%	1.7%	1.0%	5.7%	4.3%	2.0%	3.4%
	(t-62):t	-2.2%	-1.1%	-1.5%	-0.4%	6.3%	4.5%	3.9%	6.4%
	(t-125):t	-3.0%	-4.4%	-2.7%	-3.4%	6.9%	4.1%	4.4%	3.5%
	(t-251):t	-2.9%	-3.4%	-3.2%	-2.3%	5.8%	5.1%	5.2%	3.6%
	(t-754):t	-5.9%	-5.9%	-4.2%	-2.3%	6.2%	6.1%	5.0%	3.8%

Note: Table 7 reports the annual excess return, relative to the equity market (Ibovespa Index), for different momentum and contrarian strategies between 2012 to 2019. All strategies are long-only, and portfolios equally weighted. Panel A sort stocks into losers or winners in relative terms, stocks below the median in the cross-section are considered losers. Panel B sort stocks by their own absolute past return, in the time-series, with stocks with return below zero considered as losers. Contrarian strategies select stocks that were considered losers in the respective past return period, while momentum strategies select stocks that were classified as winners. Portfolios are rebalanced on the first day of the month either monthly, quarterly (January, April, July, and October), bi-annually (January and July), and Yearly (January).

5.2.1 Retail investors as contrarians.

As noted by Barber and Odean (2013), the characterization of an investor as a contrarian is difficult and is related to the investor's belief, not only its actions. Online contrarian definitions associate contrarian investment as going against the market trend, purchasing, or selling in contrast to the contemporaneous sentiment. One feasible way to translate this into the investor action is to consider a contrarian investment when investors go in the opposite direction of a signal. For this exercise, we use recent changes in price as signals, but that is not the only way investors can have contrarian behavior[42]. For example, Luo *et al.* (2021) characterize investors as being contrarian to earnings announcement signals.

We present evidence that individuals are more contrarian than institutions, and that poorer investors are more likely to present contrarian behavior. We look at every single purchase made by each investor during 2012 to 2019. Our rationale is simple, we define for each purchase if that trade was a momentum, or a contrarian trade relative to a specific past return interval. Investors who have a higher proportion of contrarian trades (60%) are considered contrarian for the sample period. It is important that we rank investors based only on their purchases. Individuals are undiversified and rarely short sell. Therefore, the decision to sell focuses on the few (or even only) stocks they hold, while the buying decision is over the full set of available stocks. Furthermore, the disposition effect affects the selling decisions of individual investors.

We define three classifications for the investor trading behavior relative to past returns - contrarian, neutral, and momentum. Investors classified as contrarian have a higher proportion of its purchases in a loser portfolio. Specifically, the classification follows the steps:

1. Stocks ordered by their past X-months return for each day t.
2. Characterize stocks below the median as loser stocks, and above the median as winner stocks, using the cross-section or the time-series.
3. Rank investors by their proportion of loser stocks bought according to definition “2”.
4. Divide investors into three categories. Contrarian investors have at least 60% of their purchases buying a loser stock. Neutral investors have between 60% to 40% of their purchases buying a loser stock. While momentum investors have less than 40% of their purchases buying a loser stock.

---

[42] A different question that we do not try to address, is why an investor would behave as a contrarian or momentum investor. Contrarian behavior could be justified by the investor’s belief in some sort of mean reversion property of the market or stock. If the investor believes that exists overreaction in the market, expecting mean reversion and being a contrarian make sense from her point of view. On the other hand, extrapolation belief theories such as in Barberis et. al. (2018) could help explain momentum behavior by the investor. Extrapolation leads to the belief that an upward trend will tend to continue to go up.

Contrarian	Neutral	Momentum
Proportion of purchases in a loser portfolio higher than 60%.	Proportion of purchases in a loser portfolio between 40% to 60%.	Proportion of purchases in a loser portfolio lower than 40%.

Table 8 shows the proportion of investors for each investor group classified as contrarian, neutral, or momentum using the previous 12-months return. Panel A, which uses the cross-section to divide stocks into winners or losers, shows that close to 50% of the institutional investors had the majority of their purchases on past winner stocks, in contrast with less than 35% of retail investors. Only institutions and the retail investors of the highest quintile of wealth have more investors classified as momentum than as contrarians. The first column takes the difference between the proportion of momentum investors and contrarian investors to help understand the differences in trading pattern between the groups.

Table 8 – Investor's classification by its purchase pattern, proportion of investors by group.

Investor / Classification		**Difference Mom - Cont**	**Contrarian**	**Neutral**	**Momentum**	**Obs**
Panel A - Cross-Section						
Institutions		18.3%	27.9%	25.9%	46.2%	32378
Retail	***Top*** **1%**	12.1%	27.6%	32.7%	39.7%	15010
	Top **Quintile**	5.1%	30.8%	33.3%	35.9%	298565
	Quintile 4	-2.1%	35.9%	30.2%	33.9%	304549
	Quintile 3	-7.5%	40.4%	26.7%	32.9%	290633
	Quintile 2	-14.1%	46.0%	22.1%	31.9%	275232
	Bottom **Quintile**	-25.4%	55.4%	14.6%	30.0%	247573
Panel B - Time-Series						
Institutions		34.8%	23.2%	18.9%	58.0%	32378
Retail	***Top*** **1%**	55.5%	14.3%	16.0%	69.7%	15010
	Top **Quintile**	51.9%	15.9%	16.4%	67.7%	298565
	Quintile 4	47.7%	18.1%	16.1%	65.8%	304549
	Quintile 3	42.9%	20.8%	15.6%	63.6%	290633
	Quintile 2	35.9%	24.9%	14.2%	60.9%	275232
	Bottom **Quintile**	21.4%	33.8%	11.1%	55.2%	247573

Note: Investors are characterized by their purchases. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock in the cross-section are stocks with past 12-month return below the median, while in the time-series are stocks with past 12-month return below zero.

The first column reveals that investors with more wealth have more momentum than contrarian trades, and this relation is monotonic among the groups of wealth[43]. This monotonic relation is observed in the contrarian proportion (decreasing with wealth), and in the momentum proportion (increasing with wealth). This finding for the Brazilian data relates to Grinblatt and Keloharju (2000, 2001) finding for Finish investors. The authors also found evidence that individuals investors follow contrarian behavior with respect to past returns, with that behavior being stronger for smaller accounts. However, unlike our findings the authors did not find differences in performance among these investors or trading patterns.

The monotonic relation between investor wealth and contrarian behavior is present when considering other horizons of past returns such as 1 month, 6 months and 36 months, but reverses for daily and weekly past returns (Table A6 on the appendix). Overall, retail investors have more momentum trades for shorter horizons, which relates to the finding of extrapolation in shorter horizons discussed in Barberis *et al* (2018). Next, we relate investors' contrarian trading pattern to their performance.

5.2.2 Contrarian investors losses.

Did investors characterized as contrarians have higher losses? To answer this, we sort each group of wealth into the contrarian, neutral and momentum categories as in Table 8. Previously we showed that overall, purchasing past winner stocks was more profitable than purchasing past loser stocks for the period under study. However, after assorting an investor to a group, the investigation takes into consideration all his trades (purchases, sell, derivatives) over the entire period. Thus, trading results depend on their specific stock selection, holding period and other trades, and is not obvious that groups will have different results.

---

[43] We classify investors based on their purchases. When an investor has only one purchase for the entire eight-year period, she is classified as either momentum or contrarian based solely on that purchase. To alleviate concerns that the results are driven by investors who are being classified with few purchases (observations), we replicate Table 8 with investors who made at least 10 different purchases and get qualitatively similar results.

Table 9 – Annual returns and proportions of investors with negative monetary result by wealth groups and contrarian investment style.

		Top 1%	5th Quintile	4th Quintile	3th Quintile	2nd Quintile	*Bottom* Quintile
Panel A Cross-section classification (12-month past returns)							
Returns Annual (%)	Contrarian	-0.79	-1.66	-1.56	-1.89	-1.83	-2.9
	Neutral	-1.2	0.62	1.53	1.49	1.84	1.47
	Momentum	-1.75	0.28	0.72	1.3	0.72	1.33
% of Investors with negative result	Contrarian	34%	35%	36%	39%	43%	49%
	Neutral	26%	25%	27%	31%	36%	42%
	Momentum	30%	27%	30%	33%	37%	43%
Panel B Time-Series classification (12-month past returns)							
Returns Annual (%)	Contrarian	-1.61	-2.57	-3.04	-3.92	-3.17	-5.49
	Neutral	-0.04	0.02	0.82	1.03	0.81	1.18
	Momentum	-1.74	0.79	1.38	1.52	1.25	1.39
% of Investors with negative result	Contrarian	52%	55%	55%	55%	58%	62%
	Neutral	38%	37%	36%	39%	43%	47%
	Momentum	23%	21%	23%	27%	32%	37%

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock in the cross-section are stocks with past 12-month return below the median, while in the time-series are stocks with past 12-month return below zero. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations of each wealth group is reported on Table 8. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative result are investors with the cumulative trading measure below zero at the end of 2019.

Table 9 shows that except for the top 1%, results within each group of wealth in Panel A and B have the same pattern, with the average annual performance of contrarian investors being worse than momentum or neutral investors. Thus, within a group of similar wealth, investors who had more contrarian purchases had worst overall performance in all its trades. Because the proportion of contrarian investors is higher for groups in bottom quintiles, we can expect worst results for those groups. Each contrarian group holds between 24% to 27% of the total wealth of their respective wealth groups. This means that in the bottom quintile of wealth, where there is a higher proportion of contrarians, the contrarian investors are the ones with less wealth within that group. Further, Table 9 reveals that contrarians in the lower level of wealth have even lower returns, but there is no clear relation between returns and wealth for neutral and momentum investors.

The bottom rows of each panel of Table 9 show that more investors are losing money due to trading in the quintiles of lower wealth and in contrarian groups. The result that the proportion of investors losing money is larger for contrarian investors than for momentum investors holds for different past returns classification and all wealth groups (see Table A7 on the appendix). The contrasting results for the top 1% are surprising. The top 1% was the only group that had monetary loss for the investors classified as neutral and momentum. Table 9 presents results for the accumulated trading losses of an eight-year period. A closer look at the dynamics of this result shows that the pattern found in the other five wealth groups existed for the top 1% until the end of the year 2018. However, results changed in 2019, when the very rich momentum investors had a bad performance.

5.2.3 Contrarian investors and lack of diversification.

Section 5.1 revealed that the lack of diversification of retail investors was related to their underperformance. Investors that had worse diversification had higher losses in all groups of wealth. Table 11 and 12 investigate how these results hold when we sort investors into groups of wealth, diversification, and contrarian investment style.

Table 10 – Annual returns and proportions of investors with negative monetary result by the intersection of wealth groups, terciles of number of stocks in portfolio, and contrarian investment style - using cross-section classification with past 12-month returns.

Panel A - Group returns (Year %)

	Top **1%**			***Top*** **Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Contrarian	-2.0	-1.8	-0.3	-7.6	-5.5	0.2	-7.9	-4.3	1.0
Neutral	-3.5	-3.0	-0.9	-2.8	-2.3	1.2	-4.1	-1.3	2.6
Momentum	-0.8	-3.3	-1.7	0.9	-0.5	0.4	-1.2	0.1	1.3

	Quintile 3			**Quintile 2**			***Bottom*** **Quintile**		
Contrarian	-8.3	-3.3	1.1	-9.9	-2.7	1.2	-11.7	-3.1	1.0
Neutral	-3.0	-0.4	2.7	-4.7	0.2	3.2	-3.3	-0.7	3.1
Momentum	-1.6	0.5	2.5	-5.6	0.8	2.4	-2.6	2.6	1.3

Panel B - Percentage of investors with negative return

	Top **1%**			***Top*** **Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Contrarian	42%	39%	29%	49%	43%	25%	49%	41%	25%
Neutral	39%	37%	22%	47%	38%	19%	46%	36%	19%
Momentum	39%	36%	26%	39%	33%	22%	40%	33%	23%

	Quintile 3			**Quintile 2**			***Bottom*** **Quintile**		
Contrarian	49%	41%	27%	51%	43%	30%	53%	46%	35%
Neutral	47%	37%	22%	48%	38%	26%	50%	41%	30%
Momentum	40%	34%	25%	43%	36%	28%	45%	40%	34%

Panel C - Observations

Top **1%**			***Top*** **Quintile**			**Quintile 4**		
835	1185	2129	17354	26545	48125	24370	38046	47067
289	922	3691	5920	21153	72313	7689	28144	56136
726	1380	3853	13843	28996	64316	18374	37533	47190

Quintile 3			**Quintile 2**			***Bottom*** **Quintile**		
33408	45707	38315	48210	51418	26923	82769	42186	12259
9777	30810	36887	12024	28357	20451	11226	18298	6590
25482	39039	31208	32848	36623	18378	45472	22167	6606

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors are divided by the average number of stocks they had throughout the sample into three groups (terciles) of diversification. Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock are stocks with the past 12-month return below the median. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations within each group is reported on Panel C. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

Table 11 – Annual returns and proportions of investors with negative monetary result by the intersection of wealth groups, terciles of number of stocks in portfolio, and contrarian investment style - using time-series classification with past 12-month returns.

Panel A - Group returns (Year %)

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Contrarian	-1.3	-1.9	-1.7	-6.7	-6.1	-0.7	-7.8	-6.1	0.3
Neutral	-1.4	-1.2	0.0	-5.7	-3.2	0.8	-6.9	-2.1	2.2
Momentum	-1.4	-3.5	-1.4	0.7	-0.4	1.1	-0.9	0.4	2.0

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	-9.4	-5.6	0.3	-5.8	-5.2	1.7	-17.4	-5.1	-0.2
Neutral	-4.9	-1.3	2.6	-6.5	0.1	1.8	-4.3	-1.0	3.2
Momentum	-1.2	0.7	2.6	-6.0	0.8	2.9	-1.8	1.4	2.1

Panel B - Percentage of investors with negative return

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Contrarian	52%	56%	50%	59%	60%	49%	60%	58%	47%
Neutral	46%	44%	35%	53%	48%	31%	53%	44%	27%
Momentum	33%	30%	20%	34%	28%	16%	35%	28%	17%

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	59%	56%	45%	61%	56%	49%	63%	59%	52%
Neutral	52%	43%	29%	52%	44%	34%	54%	46%	36%
Momentum	36%	30%	20%	39%	33%	23%	40%	35%	28%

Panel C - Observations

***Top* 1%**			***Top* Quintile**			**Quintile 4**		
571	656	912	12734	15674	19023	17604	20508	17059
194	525	1688	4150	12609	32137	5478	16789	26639
1085	2306	7073	20233	48411	133594	27351	66426	106695

Quintile 3			**Quintile 2**			***Bottom* Quintile**		
23219	23549	13576	33816	26268	8574	57873	21297	4401
7343	19832	18151	9577	19444	10056	9841	13656	3937
38105	72175	74683	49689	70686	47122	71753	47698	17117

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors are divided by the average number of stocks they had throughout the sample into three groups (terciles) of diversification. Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock are stocks with past 12-month return below zero. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations within each group is reported on Panel C. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

Like in 5.1, lack of diversification is a characteristic that was highly associated with investor trading loss. Table 10 and 11 report that investors that followed a momentum strategy but had on average less than 1.4 stock in their portfolio, had worse returns and higher probability of having a loss for the cumulative measure. Contrarian investment style was still detrimental to investor wealth, even within a group with the same diversification and wealth. Panel B of Tabel 10 and 11 report that for all 36 separate groups of wealth and diversification, investors following a contrarian strategy were more likely to have losses than investors following a neutral or momentum strategy. On the appendix, we replicate Table 10 and 11 for level of activity instead of diversification. Higher activity also intensifies the losses of investors with contrarian investment style.

In summary, section 5 shows that contrarian trades, low diversification and high activity intensify retail losses, especially the first two. Newer accounts are on average less diversified and more active, having worse performance than older accounts. Commission-free brokerage apps make it easier for all investors to participate and trade in the stock market and derivatives market. Our evidence suggests that with higher trading we can expect an increase in monetary transfer from poor investors to richer investors in the stock market, especially for newer and under-diversified investors. On the next section we use every individual trade and holdings to further our investigation on monetary transfer and its consequences for distribution of wealth.

6 Inequality counterfactuals - how much of the increase in inequality comes from active trading?

The distribution of wealth in risky assets can change because of three distinct factors. First, even in a scenario when there is no trading, individuals' wealth will vary differently due to the heterogeneity in returns in their initial holdings in the stock market. Next, individuals' wealth in risky assets will also change because of net capital flows in/out of equity market. For instance, if only very wealthy individuals increase their position in risky assets, there will be a mechanical increase in inequality because of this inflow. Finally, new trading will also lead to different holdings and affect wealth distribution. This could happen because of market timing ability, with net flows in/out of equity market, or because of stock selection ability across individual stocks. On previous exercises our focus was on wealth transfers coming from new trading, the third factor, now we assess how much each of these factors contribute to changes in wealth inequality.

We calculate different counterfactuals of wealth and compare their Gini Indexes to estimate how new trading contributes to changes in inequality of wealth held in risky assets. Table 3 shows that investors with lower levels of wealth had relatively larger losses and had more investors losing money. These results initially suggest that trading should contribute to an increase in inequality of holdings in risky assets. However, other factors also affect the distribution of these holdings, such as heterogeneity of returns in their initial holdings, and net capital flows in/out of equity market holdings.

The last factor is mechanical, for example, if investors with lower initial wealth are the ones with higher positive net flows, then inequality should decline. On the other hand, the former factor says that inequality could change due to return heterogeneity from previous selected portfolios, old trading choices. If the initial portfolio of investors with high wealth has higher returns than the initial portfolio of investors with lower levels of wealth, the inequality would increase. Studying the Indian stock market, Campbell, Ramadorai and Ranish (2019) found that return heterogeneity was an important driver of the increase in inequality. However, their decomposition did not disentangle initial portfolio (past trading choices) from new trading, or either considered the actual holding period of each investor.

The decomposition is created upon a simple counterfactual logic. We calculate an inequality measure, e.g., Gini Index, for the account size both at the beginning of the period t and at the end of the period τ, accounting for all the trades of all investors in both, the stock and derivatives market. The difference between the Gini Index at the beginning of the period and the Gini Index after all trades take place is how much the inequality changed in the studied interval. The second step of the exercise is to calculate counterfactuals of how the account size would change if different trade channels were shutdown. We start shutting down all new trades and new flows, which shows how account size would change due to initial holdings return heterogeneity. On another counterfactual, we shutdown new trading in the derivatives market to understand how account size would change due to derivative trading. Differences between the Gini Index calculated under those counterfactuals will give us an estimate on how inequality evolves under different hypothesis.

Table 12 – Definitions of the counterfactual exercise.

Counterfactual	**Counterfactual exercise**	**Which redistribution factors are shut down?**
Baseline (A)	Investors hold the **same exact portfolio** in the beginning and **do not** trade	All.
Initial holdings return heterogeneity (B)	Investors hold their own portfolio in the beginning and **do not** trade	New trading across individual stocks, derivatives, and net flows.
Derivatives (C)	Investors hold their own portfolio in the beginning and trade only derivatives contracts.	New trading in the stock market.
Daily Imbalance Stock Market (D)	Investors hold their own portfolio in the beginning and trade freely across different individual stocks at its closing prices, with negative net flows being invested in bonds and with positive net flows coming from disinvestment in bonds.	Intraday trading in the stock market.
Benchmark (E)	Investors hold their own portfolio in the beginning and trade freely across different individual stocks (at their average intraday execution price), derivatives contracts and in/out stock market.	None.

Note: Table 12 define the counterfactual exercises for the decomposition of changes in the inequality measure assuming that positive net flows come from disinvesting in safe assets, while negative net flows are invested in safe assets. The difference between the Gini Index calculated from (E) and (A) gives the total variation in the Gini Index in the period. Difference from Gini calculated from (B) and (A) gives the contribution to the total variation from initial holdings return heterogeneity. Differences between the Gini calculated from (C) and (B) gives the contribution of new trades in the derivative market. Differences between the Gini calculated from (D) and (C) gives the contribution of new trades in the stock market, coming from daily imbalances. Differences between the Gini calculated from (E) and (D) gives the contribution of returns in the same day for new trades in the stock market. While differences between the Gini calculated from (E) and (B), gives the total contribution from new trading in the stock and derivatives market.

One important limitation of the data is that we cannot observe investor wealth outside her equity market holdings. This limitation makes it impossible to know from where came the money for a positive net flow into the equity market, and to where did the money went after a negative net flow from the equity market. Therefore, to build the counterfactuals we assume that positive net flows come from disinvesting in bonds, while negative net flows are invested in bonds. This hypothesis considers that positive net flows to equity market are accounted as investor`s wealth in the beginning, but as a safe asset. Also, any negative net flow continues to be investor`s wealth in a safe asset[44].

We chose this assumption because during the eight years of the sample, positive net flows are greater for individuals with low levels of equity wealth. There is ample evidence that investors with lower levels of wealth have also lower participations rates on equity market[45]. Thus, it is better to assume that those investors entering the stock market had previous investments in other safe assets, rather than think that the wealth was created at that instant. Also, this hypothesis avoids the interpretation that net flows are informing us about the dynamics of consumption and investment of these individuals. Finally, the assumption has a closer relation to our approach in the previous exercise, when estimating trading gains with excess returns.

Table 12 presents a summary of the counterfactuals[46]. The baseline (A) is to suppose that investors at the beginning of the period hold the exact same portfolio (market and bonds) and do not trade, so Gini Index in the end remains the same as in the beginning. The *benchmark* counterfactual (E) is having investors starting with their own portfolio and trading freely from day t to day τ with negative net flows invested in bonds and with positive net flows coming from disinvestment in bonds. Hence, the Gini Index calculated with the benchmark minus the Gini Index calculated with the baseline give us the total Gini Index variation. Estimation of the other Gini Index differences follows the counterfactual differences:

---

[44] Although we highlighted that net capital flows in/out of equity market are a factor that contributes to changes in distribution of holdings, our assumption to treat net flows modifies this factor. Because investors already have wealth in safe assets at the beginning (which will be invested in equity) and at the end (that were invested in equity), net flows will change the distribution because of the return heterogeneity between these two asset classes and can be captured by the market timing ability of investors. A different assumption is to consider that positive net flows come from wealth that was created at that instant, while negative net flows are straight away consumed and do not constitute wealth anymore, see Table A10 on the appendix for a description of the counterfactuals under this assumption. Figure A3 on the Appendix shows that on average net flows contributed to reduce inequality in direct holdings, with wealthier investors having negative net flow and investors with smaller holdings having positive net flow.

[45] Vissing-Jørgensen and Attanasio (2003), Calvet et. al., (2007), and Bach et. al., (2020) all find evidence of higher stock market participation rates for wealthier households. Guiso and Sodini (2013) report that limited participation in the stock market is a worldwide phenomenon and has inverse correlation with investor wealth.

[46] To accurately mimic what would be investor wealth in the beginning and end of the period in each counterfactual, we discount payout flows (dividends and others) that happened through time in the investor initial portfolio and on his new trading flows. Given the exercise assumptions on net flow, we assume these payout flows are invested in treasury bonds.

$$
\begin{gathered}
Gini\ (E) - Gini\ (A) = \Delta\ total \\
Gini\ (B) - Gini\ (A) = \Delta\ due\ to\ initial\ holdings \\
Gini\ (E) - Gini\ (B) = \Delta\ due\ to\ trading \\
\\
Decomposition\ of\ trading\ contribution: \\
\\
Gini\ (C) - Gini\ (B) = \Delta\ due\ to\ trading\ derivatives \\
Gini\ (D) - Gini\ (C) = \Delta\ due\ to\ daily\ trading\ in\ stock\ market \\
Gini\ (E) - Gini\ (D) = \Delta\ due\ to\ intraday\ trading\ in\ stock\ market
\end{gathered}
\tag{7}
$$

6.1 Inequality increases because of new trading.

Figure 5 shows the variation in the Gini Index when we do the counterfactual exercise for the beginning of each year. The initial portfolio of each year is the result of all past trades up to that point and the safe assets that will be invested in equity through the year. We compare the Gini Indexes obtained from the distribution of the initial portfolios with the Gini Index obtained after accounting for the returns of the portfolio and new trades in that year. An average year had an increase of 0.04 in the Gini Index, while the average beginning Gini Index was 0.916[47]. Thus, there was a high concentration of wealth in this market, which got higher because of differences in trading performance. That happened every year, except for 2016, when the difference is close to zero. This pattern in Figure 5 confirms that underperformance by poorer investors was very persistent.

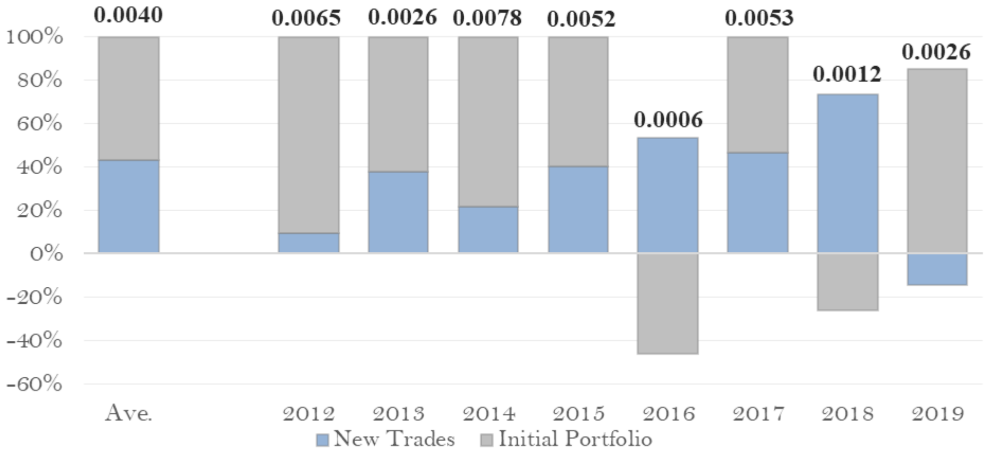
Each year, the variation in Gini could have come from return heterogeneity of past trades (initial portfolio of that year) or from new trades made in that same year. Figure 5A shows that in general both factors positively contribute to the increase in inequality through time. New trades account for 43% of the increase in the Gini Index on an average year. This indicates that poorer investors have worse short-term performance after they trade than wealthier investors. Further, the initial portfolio of each year was the largest contributor to rise in inequality, which suggest that previous choices of poorer investors continue to underperform the ones made by richer investors.

---

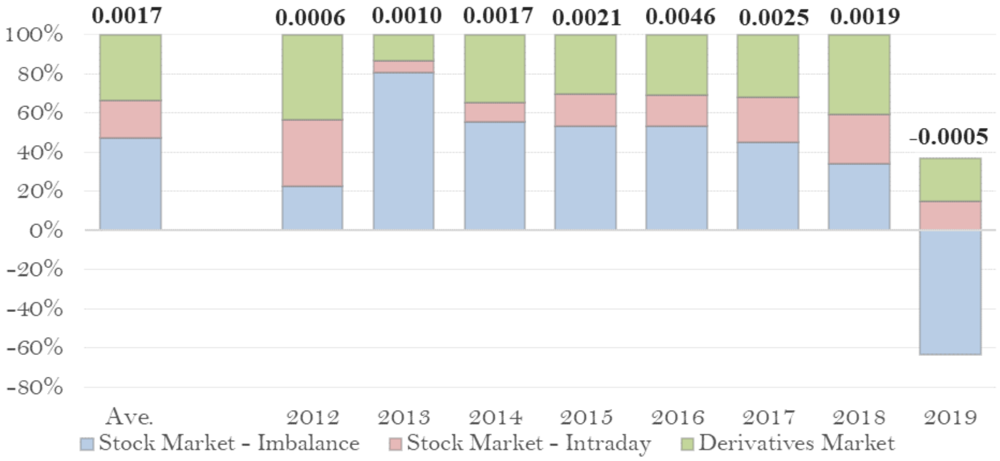
[47] Gini Index calculated in the beginning of each year are different not only because of returns and observed trading. For example, investors who participated in the market in 2014 and were considered in the Gini Index calculation of that year could be out of the market in 2015, and that changes the distribution when calculating the Gini Index at the end of 2014 and in the beginning of 2015. Also, other adjustments in position that we do not account during the year like follow-ons, inheritance, and donation could have marginal impact.

Figure 5 – Yearly decomposition of Gini Index variation through counterfactuals (2012-2019).

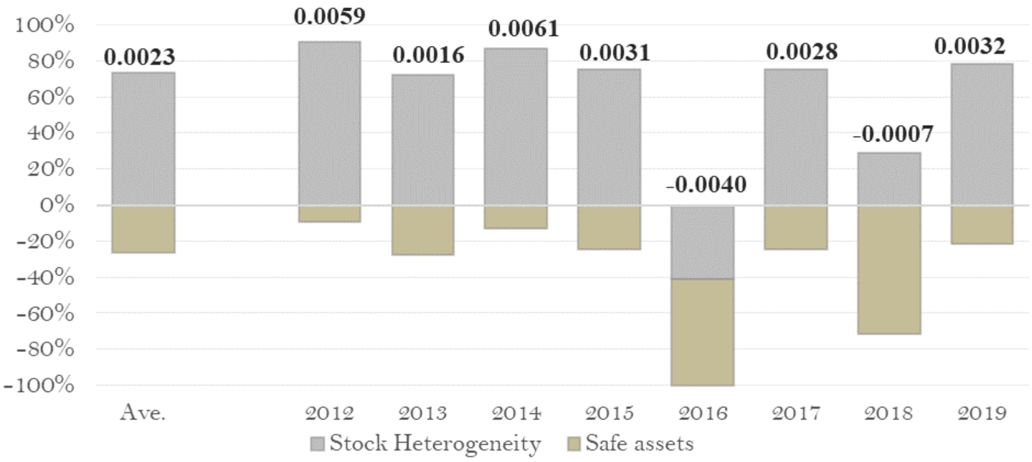
A – Initial Portfolio vs. New Trading



B – New Trading Decomposition



C – Initial Portfolio Decomposition



Note: Figure 5 reports the variation in the Gini Index from different counterfactuals described in Table 12, when performing the counterfactual exercise for the beginning of each year, until the end of the same year. The first column reports the time-series average of the year-by-year results. “New trades” and “initial portfolio” are further decomposed in the second (B) and third (C) figure.

Figure 5B further decomposes the new trading contribution to inequality in different components. Intraday results, which account for day trading and execution price, contribute to the increase in inequality, with poorer investors doing worse in this area. Although only 14% of investors engage in trading derivatives, this was the market with the largest differences in performance between investors with distinct levels of wealth, which led to derivatives trading being, on average, responsible for 34% of the increase in inequality through new trades. Daily market order imbalances, which is what is usually captured in retail trading data, accounted for less than 50% of the total increase in Gini Index due to new trading over the years.

The proportion of equity and safe assets that each investor starts affects the initial portfolios return heterogeneity. Figure 3C decomposes the initial portfolio results in contribution from stock return heterogeneity and from the safe asset allocation. The safe asset allocation between investors helped to reduce the inequality and all increase observed in Gini Index happened due to initial portfolio return heterogeneity across stocks. As robustness, Figure 4A in the appendix repeats the previous exercise using only accounts that already had equity holdings in the beginning of each year (old accounts). This helps to alleviate concerns about the effect of safe assets in the initial holdings' component and their contribution to inequality. Figure 4A reports no substantial change in results.

In summary, the counterfactual exercises support Table 3 evidence that active trading leads to an increase in inequality in risky assets. The counterfactual decomposition reveals new trades and portfolio rebalance within the same year as relevant contributors to inequality growth, about 43% on average. It is likely that a portion of the return heterogeneity contribution to inequality found by Campbell, Ramadorai and Ranish (2019) was coming from intramonth trade. We also find that same day returns, and trading derivatives were two relevant channels that contributed to the concentration of wealth in risky assets.

As pointed out by ALS (2022), wealth transfers reported here could be even larger in bubble periods. However, our results suggest that redistribution through trading is relevant even in calm periods, with no bubbles. Several factors explain this difference in outcomes. First, this study accounts for derivatives and intraday execution, two trading channels which increase wealth transfer from poor to rich. Second, instead of focusing on the ultra-rich in comparison to a larger number of investors with lower wealth, our exercises disaggregate on each investor. Table 3 reports disparity in trading results among the lower wealth groups, which also contributes to inequality growth. Third, different from ALS (2022) investigation, initial portfolios were different for investors with distinct levels of wealth, with the resulting return heterogeneity being the major contributor to wealth inequality growth.

7. Conclusion.

This paper uses high-quality administrative data on equity trading in Brazil to analyze how heterogeneity in trading outcome results in wealth transfer and impact the dynamics of wealth concentration in equity holdings. We document that individual investors as a group lose to institutions when they trade and that individuals with lower wealth have higher relative losses and higher propensity to experience a loss. Our analysis reveals that individual active trading led to a wealth transfer from poor to rich and increased concentration in equity holdings.

We show that investors with lower wealth lost on average 2% of their holdings in a year because of trading. This was two times higher than investors with higher wealth. Our methodology focuses on monetary (dollar-weighted) trading results, which is closer to what investors truly experience when trading. Under this perspective, we surprisingly find that most investors experience monetary gains when trading, but that the investors who lose, lose higher amounts, and drive aggregate results. This fact emphasizes that aggregate (volume-weighted) results for retail investors deserve a caveat and may be a deficient representation of all individual investors, and much more representative of a small group of wealthy and highly active investors.

Our findings that intraday execution and derivatives market increase individual trading losses suggest that results from exercises that do not take these factors into account could be seen as a lower bound for the individual trading loss. We found that losses were 57% greater when considering these other factors. Further, poorer investors also do relatively worse when trading derivatives contracts and in their same day return, which contributes to widening the wealth concentration in risky assets.

This investigation also revealed that individual investor investment style is associated with their wealth in the equity market. Investors with lower wealth are more contrarian, buying stocks with recent price drops, and less diversified than investors with higher wealth. By sorting investors by wealth and investment style, we provide evidence that lack of diversification and contrarian trading pattern by individual investors increase individual loss and wealth transfer from poor to rich.

Our results are robust to different definitions of account size, investors groups, and subsamples. The evidence found in this study is more relevant in nowadays context of the growing popularity of commission-free brokerage apps, which incentivize trading in the stock and derivatives market.

4. The relationship between investors trading activity and social media.

Abstract

This paper uses Brazilian administrative investor level data to investigate the effect of social media interruptions on investors trading activity. On days when social media platforms exogenously stop working, retail investors and domestic institutions reduce their trading activity. The result is observed for the money traded in stocks and for the number of investors trading stocks, with drops as high as -14.3% for retail traded monetary volume and -7.6% for the number of active retail investors. Interruptions which occur outside market hours (a sort of placebo) have no effect on trading activity. Foreign investors, which are less affected by the local outages, do not respond to the interruptions. Investors shunning away from the markets when they cannot access social media platforms supports the hypothesis that social media are vehicles of information about the market, with the potential to reduce informational frictions and improve market liquidity. In fact, as investors leave the market on outage days, market liquidity worsens. Additional results further corroborate the information hypothesis, such as social media outages affecting return responses to firms' announcements. On outage days, firms' announcements do not predict the next day returns like they do in regular trading days.

Keywords: investor activity; stock market; social media; information diffusion.

1. Introduction.

Retail and institutional investors face a different world than decades ago. There are differences in the methods available to trade stocks, from orders made by phone call to online orders in web browsers, and most recent in mobile apps. Differences in the way investors collect information, from physical newspapers and television broadcasts to an unlimited source on the internet just one click away on smartphones. Communication is also different, with social media platforms now being a big part of how everyone communicates, investors or not.

Social media, which seems to have become omnipresence in the life of many, serves several purposes. It is a direct communication tool, with direct messaging between users. It is an information tool, which publicizes and disseminates news from traditional and alternative sources. And it is also a tool for entertainment, where users can perform a wide variety of activities, even playing games. With that in mind, it is hard to imagine that social media does not play a role in how investment decisions and trading are realized nowadays. But which role does it play? Does it improve market quality? For example, if social media improves the dissemination of information among investors like conventional media, it can improve market efficiency (Peress, 2014). This study takes a step in the direction of understanding the connection between trading and social media by using Brazilian investor level data to investigate the effect of social media interruptions on investors trading activity and market liquidity.

We explore how a common shock to all investors that limits their interaction with a social media platform affects trading activity and market liquidity. We identify 62 social media platforms interruptions between 2012 and 2019, with 33 occurring during market hours. We focus on four widely popular social media in Brazil; Facebook, Instagram, Twitter, and Whatsapp[48]. Trading activity measures are calculated from the administrative data of *Comissão*

---

[48] The popularity of these social media platform was growing in Brazil during the eight-years under study. In Brazil, Twitter had between 8 to 15 million active users and over 50 million registered accounts, Instagram had between 20 to 70 million active users, Facebook had between 30 to 90 million active users and over 150 million registered accounts, while Whatsapp, the most popular platform, had between 40 to over 100 million active users. In 2013, Whatsapp app was already installed in 80% of all smartphones in Brazil. Brazil had a population of 205 million in 2015 (middle of the sample). These platforms continue to grow nowadays, a 2022 survey showed that 94% of Brazilians answered that they are active in at least one social media platform. 92% answered they use Whatsapp regularly, with Instagram being the second most popular in 2022.
https://www.uol.com.br/tilt/noticias/redacao/2012/07/31/twitter-passa-dos-500-milhoes-de-usuarios-mas-numeros-mostram-queda-de-microblog-no-brasil.htm;
https://g1.globo.com/tecnologia/noticia/2015/11/instagram-tem-29-milhoes-de-usuarios-ativos-por-mes-no-brasil.html; https://www.tecmundo.com.br/redes-sociais/139130-brasil-terceiro-pais-usuarios-facebook.htm

de Valores Mobiliários (CVM), equivalent to the SEC in the United States. This data comprises all investors (institutions and individuals) for the largest stock market in Latin America. Objectively, we study how monetary volume and number of active investors respond to days with social media outage, while controlling for seasonality, fixed effects, and other variables which may also affect trading activity. Our main result is that investors reduce their activity on days when social media platforms are exogenously[49] interrupted. On those same days market liquidity worsens. This result is robust to different specifications and measures of trading activity, with "placebo" outages (occurring outside the market trading hours) having no results.

Social media platforms could serve as communication and information diffusion channels (Blankespoor, Miller and White, 2014 and Hu et al., 2021). For example, social media could help firms' disclosures to reach more investors, reducing information asymmetry among investors, and improving news incorporation into stock prices. Under this view, interruptions would deteriorate investor collection of information, reduce their trading activity, and worsen market liquidity, having a negative effect on market efficiency (Peress, 2014 and Xu, Xuan and Zheng, 2021). On the other hand, social media platforms could also serve as a distraction to investors (Peress and Schmidt 2020, and Brown et al., 2022). If this is the case, distracted investors would increase their trading activity on outage days.

Our empirical exercise finds evidence that supports the former hypothesis. Investors reduce the money they trade in stocks and their overall participation (propensity to trade) in the stock market on days when their interaction with social media is impaired. On outage days we observe a drop of 5% in the number of active domestic funds, and a drop of 7.6% in the number of active retail investors. Foreign funds, which are less affected by local outages, have a drop of less than 1% which is not statistically different from zero. Market trading volume reduces by 9.2% on social media outage days. This number is close to Peress (2014) estimates for the trading volume reaction to newspaper strike days (-12%). Moreover, analysis of return predictability indicates that social media propagates corporate announcements. Social media outages change how announcements affect return autocorrelation suggesting it influences how news is incorporated into prices. Overall, results support the idea of social media as a source

---

https://g1.globo.com/tecnologia/tem-um-aplicativo/noticia/2013/11/presente-em-mais-celulares-no-brasil-whatsapp-bate-chat-do-facebook.html; https://www1.folha.uol.com.br/tec/2022/07/94-tem-conta-em-alguma-rede-social-whatsapp-ldera-com-92.shtml

[49] The exogeneity of the platform's interruption is from the perspective of an investor, not from the perspective of the own platform undergoing the interruption.

of information and communication, increasing market activity and improving dissemination of information and liquidity and probably market efficiency.

Other results also corroborate that social media platforms are sources of diffusion of market information for retail and for domestic funds. The reduction in participation is stronger among smaller firms, which usually have less coverage from traditional media and could find in social media a way for circulating information. Retail investors reduce their activity relatively more on stocks that recently had higher returns. Social media could also be a way to spread the word that these stocks had higher returns. Without social media, like in outage days, these stocks would receive less attention from retail investors. Results are also stronger for retail investors and funds with less wealth in stocks. These investors may be less sophisticated, with no access to specialized platforms that deliver markets news and analysis, and therefore could rely more on social media. However, we find that herding behavior is not affected by social media outages, which is puzzling. Investors reduce their trading, but there is no effect on their correlated trading.

The paper which most relates to our study is Mohr (2021). Like in our study, the author uses social media outages to understand how investors trading activity may be influenced by the platforms. The study for the United States finds that retail investors increase in 4% their selling volume and increase in 3.5% their closing of positions. The study finds no effect for the number of active retail investors or any effect for non-retail investors. This evidence, contrary to ours, suggest that retail investors in the United States might be distracted by social media platforms. Are Brazilians and Americans retail investors that much different regarding their use of social media and its impact on trading? We have not yet found a common ground to understand these differences in results. Mohr (2021) study has two additional results that go in the same direction as what we found in our investigation. For lottery-like stocks and for stocks with recent high past returns, retail investors in the US reduce their trading activity. Thus, for at least a subset of stocks Americans retail investors appear to have a similar reaction as Brazilians investors when social media platforms stop working.

This paper contributes to a broad literature that investigates how general media (conventional or social) can help information transmission in finance and its consequences for the stock market (Fang and Peress, 2009; Engelberg and Parsons, 2011; Peress 2014, Blankespoor, Miller and White, 2014; Chen at. al. 2014; Tetlock, 2014). We also contribute to the literature that studies how social media affects investors and investment decisions.

Antweiler and Frank (2004) show that online post on social media predict volatility, while Chen et al. (2014) shows evidence that online posts on social media can also predict future stock returns and earnings surprises. Pedersen (2022) presents a novel model about how naïve investors and influencers on social media could lead to higher trading volume and bubbles on the stock market.

This paper also relates to the literature about how investors deal or respond to innovative technologies that affect their investment decisions, such as in Barber and Odean (2002) studying online brokerages, Xu, Xuan, and Zheng (2021) studying internet searching, and Barber et al (2021) studying commission-free brokerage apps. Lastly, the paper relates to a recent literature that uses platform outages as exogenous shocks to study investors activity (Mohr, 2021, and Eaton et al., 2022).

The rest of the paper is organized as follows. Next section discusses different hypotheses that could explain how investors interact with social media and how this would affect their trading activity. Section 3 describes the outages and trading data. Section 4 presents the main results of the paper, the response of trading activity to social media interruptions. Section 5 explores sources of heterogeneity on the response, at firm and investor level. Section 6 looks at social media outages impact on market liquidity. Section 7 shows that social media outages change how corporate announcements affect return predictability. Section 8 and 9, investigates investor herding and investor level characteristics, respectively, and the last section concludes.

2. Hypothesis development.

Does the impossibility of interacting with social media affect investors trading activity and market liquidity? If so, in what direction?

2.1 Dissemination of information.

Social media platforms are tools for communication and information dissemination. If social media improves corporate news coverage it can reduce informational frictions (Fang and Peress, 2009). Blankespoor, Miller and White (2014) find evidence in that direction, that firms use of Twitter spread corporate news, reducing information asymmetry among investors, and improving market liquidity, especially for smaller firms which are less “visible” in conventional media. In general, improving information dissemination improves market quality, increasing market liquidity, reducing crash risk, and facilitating the incorporation of firm-specific information into stock price (Peress, 2014; Ding, Zhou and Li 2020; and Xu, Xuan and Zheng, 2021). An increase in information dissemination can also lead to more trading. Investors could be reluctant to trade if they feel they have an information disadvantage (Kyle, 1985), and better dissemination increases the degree of investor recognition of a stock, increasing disagreement and leading to more trading (Harris and Raviv, 1993). These hypotheses suggest that social media outages impair investor information and have the potential to reduce their trading activity and worsen market liquidity.

Recent studies have results which corroborate this hypothesis. Hu et al. (2021) shows that increased site traffic of Reddit is correlated with increased retail order flow, while Cookson, Engelberg and Mullins (2022) show evidence that investors follow peers in social media that have similar sentiment as them, creating an echo chamber of information and increasing trading. However, we note that the ease of communicating and obtaining information through social media, does not necessarily mean that investors will be able to distinguish what is noise and what is signal from this source, as suggested by Antweiler and Frank (2004), and more recently by Ammann and Schaub (2021).

2.2 Distraction.

There is a large economic and finance literature studying the effects of inattention under different scenarios. Investors have limited attention and could be distracted from trading because of other activities. DellaVigna and Pollet (2009) show that investors are more distracted on Fridays, and this influences their respond to corporate announcements on this day of the week. Hirshleifer, Lim and Teoh (2009) report that investors also have weaker responses when several earnings announcements are disclosed on the same day. Peress and Schmidt (2020) find that on distraction days, proxied by news pressure on television broadcasts, there is a decrease in trading activity and market liquidity worsens. Social media platforms can be a constant source of distraction to investors, decreasing their trading activity on average. Under this scenario, trading would increase when investors cannot access social media. Brown et al. (2022) and Mohr (2021) report results that corroborate the distraction hypothesis. The first shows that internet outages on blackberry mobile devices lead to higher trading volume and frequency, while the latter reports that outages of Twitter and Facebook in the United States are correlated with the increase of retail investors selling volume and closing their positions.

3. Data

3.1 Social media outages.

We study how a shock that limits the ability of an investor to interact with social media affect its trading. For a social media platform interruption to be considered as an outage, it must meet several criteria. First, it must impact the ability of users to interact with the website or phone app. This could be an inability to post or read content or an inability to log on. Second, the events must have a duration of at least 30 minutes. Interruptions shorter than 30 minutes have less coverage from the news (sometimes with only one news coverage) limiting our understanding about the interruption. Third, events must occur during regular trading hours. Although we collect outages outside of trading hours, we only use them as a sort of placebo testing. Last, interruption must have affected Brazil and Brazilian users.

Our final sample has 62 hand collected interruptions occurring between 2012 and 2019. Out of these 62 interruptions, 33 happened during regular trading hours and 29 happened outside trading hours (early morning, late afternoon or night, weekends, and holydays). We study four popular social media platforms in Brazil; Twitter, Facebook, Instagram, and Whatsapp. Information on the interruptions is collected from news articles, which often include the time the outage began, the duration, the typical problem users are experiencing, regions reporting the problem, and if this is a generalized interruption or if it only affects part of users[50].

---

[50] We collect information from various sources of news articles to gather more accurate information on the interruptions and use the common factors among the different sources to classify interruptions as outages and to characterize their details.

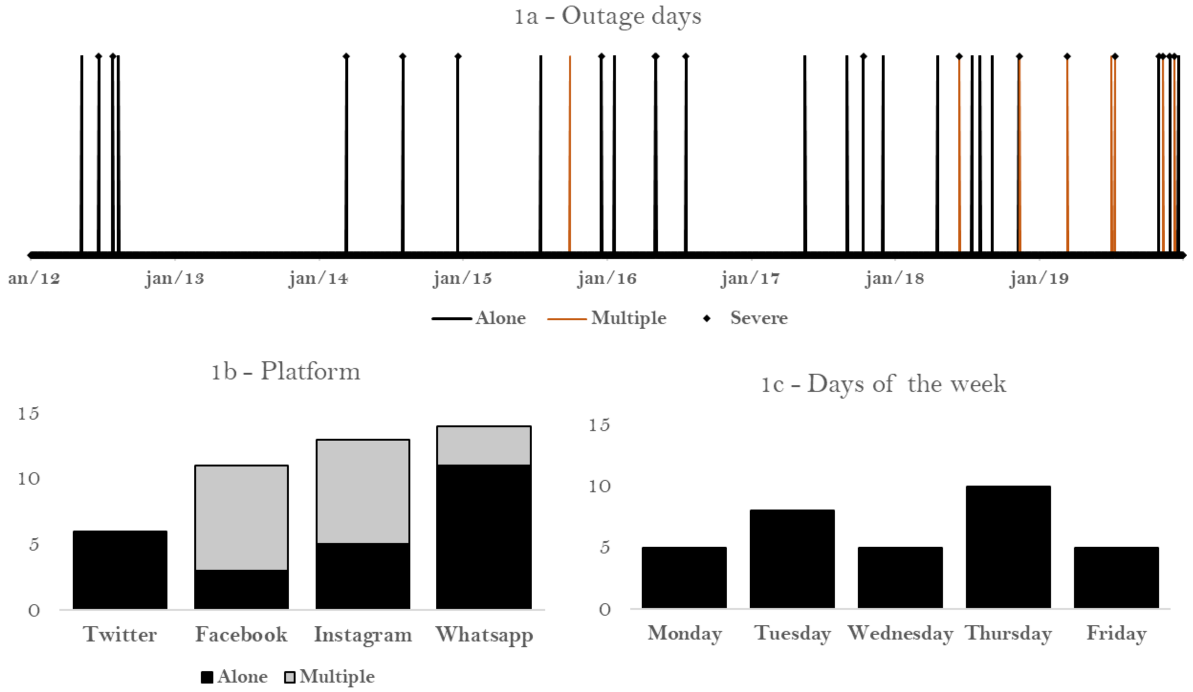
Table 1 – Interruptions of social media platforms during market hours.

Date	Platform	Start	End	Time Span (Min)	Market Hour Time Span (Min)	Weekday	Type of problem	Multiple Platform	Severe
08/05/2012	instagram	11:30	13:30	120	120	Tuesday	feature instability	No	No
21/06/2012	twitter	13:00	15:00	120	120	Thursday	outage	No	Yes
26/07/2012	twitter	12:00	14:30	150	150	Thursday	access instability	No	Yes
10/08/2012	whatsapp	15:30	18:30	180	90	Friday	outage	No	No
11/03/2014	twitter	15:00	16:00	60	60	Tuesday	outage	No	Yes
01/08/2014	facebook	12:30	13:45	75	60	Friday	outage	No	Yes
18/12/2014	whatsapp	10:00	17:00	420	420	Thursday	feature instability	No	Yes
16/07/2015	instagram	16:00	17:00	60	60	Thursday	access instability	No	No
28/09/2015	facebook	16:00	17:00	60	60	Monday	access instability	No	No
17/12/2015	whatsapp	0:01	14:00	840	210	Thursday	outage	No	Yes
19/01/2016	twitter	6:30	12:00	330	120	Tuesday	feature instability	No	No
02/05/2016	whatsapp	14:00	23:59	600	180	Monday	outage	No	Yes
03/05/2016	whatsapp	0:01	16:00	960	360	Tuesday	outage	No	Yes
19/07/2016	whatsapp	14:00	18:00	240	180	Tuesday	outage	No	Yes
17/05/2017	whatsapp	13:30	14:30	60	60	Wednesday	access instability	No	No
31/08/2017	whatsapp	12:00	13:30	90	90	Thursday	feature instability	No	No
11/10/2017	facebook; instagram	12:00	14:00	120	120	Wednesday	access instability	Yes	Yes
30/11/2017	whatsapp	16:20	17:20	60	40	Thursday	outage	No	No
17/04/2018	twitter	11:00	11:45	45	45	Tuesday	outage	No	No
12/06/2018	instagram	12:30	15:30	180	180	Tuesday	feature instability	No	Yes
13/07/2018	instagram	15:30	16:20	50	50	Friday	outage	No	No
03/08/2018	facebook	12:30	13:30	60	60	Friday	access instability	No	No
03/09/2018	facebook; instagram	16:30	18:30	120	30	Monday	outage	Yes	No
09/11/2018	instagram	16:15	17:00	60	45	Friday	feature instability	No	No
12/11/2018	facebook; instagram	15:40	17:40	120	80	Monday	outage	Yes	Yes
13/03/2019	facebook; instagram; whatsapp	13:00	18:00	300	240	Wednesday	feature instability	Yes	Yes
03/07/2019	facebook; instagram	10:00	13:00	180	120	Wednesday	feature instability	Yes	No
11/07/2019	twitter	15:30	17:15	105	90	Thursday	outage	No	Yes
30/10/2019	facebook; instagram; whatsapp	12:00	16:00	240	240	Wednesday	feature instability	Yes	Yes
11/11/2019	whatsapp	11:30	14:00	150	150	Monday	outage	No	Yes
28/11/2019	facebook; instagram	11:00	12:30	90	90	Thursday	access instability	Yes	Yes
10/12/2019	whatsapp	15:45	17:00	75	75	Tuesday	feature instability	No	Yes
19/12/2019	facebook; instagram; whatsapp	10:30	11:45	75	75	Thursday	feature instability	Yes	No

Note: Information on the interruptions is collected from news articles. The common factors of different news articles are used to define the time the outage began, the duration, the typical problem users are experiencing, and if this is a generalized interruption or if it only affects part of users. ‘Feature instability’ is when users have difficulties posting or reading on social media, ‘access instability’ is when users have intermittent difficulties logging on, while ‘outage’ is when the app and website are completely down. Severe interruptions last at least 1 hour, are reported in the news as a generalized outage (not only for some users) and are officially acknowledged by the company representatives.

Table 1 brings a list of the 33 interruptions that happened during the market hours (the list with the other 29 interruptions outside market hours is relegated to the appendix). We classify them as three distinct types of problems reported by users. ‘Feature instability’ is when users have difficulties posting or reading on social media, ‘access instability’ is when users have intermittent difficulties logging on, while ‘outage’ is when the app and website are completely down. Table 1 also classifies if the outages were more significant or not. Severe interruptions last at least 1 hour, are reported in the news as a generalized outage (not only for some users) and are usually officially acknowledged by the company representatives. Other interruptions are either shorter, or news articles report it as affecting only part of users.

Figure 1 – Timeline, social media platforms and days of the week of interruptions occurring during market hours.



Note: Information on the interruptions is collected from news articles. Alone outages are when there is only one platform that had interruptions problems on that day. Multiple outages are when at least two platforms had interruptions on that day. Severe interruptions last at least 1 hour, are reported in the news as a generalized outage (not only for some users) and are officially acknowledged by the company representatives.

Figure 1 summarizes information on the events. Interruptions appear more common in the second half of the sample, in more recent years. We conjecture three reasons why this may happen. First, this could be due to bias on finding news articles from further years past, if that is so, only the more relevant interruptions of past years would still be available on the ‘history’ of news articles. Second, it could be that social media was becoming increasingly more popular, and news editors and writers consider it more relevant to report the outages on the recent days than on the past days. Third, it could be simply that platforms had more problems with their servers in recent years. We note that twitter has less interruptions than other platforms, especially in recent years. During our sample, Facebook purchased Instagram (in 2012) and Whatsapp (in 2014), which contributed to the chance of multiple platform outages in the later years of the sample. Interruptions are well distributed among weekdays. All interruptions are driven by site testing, server problems, and some are even judicially imposed. This would make them exogenous to other market features that could drive the results[51].

3.2 Trading data at investor level.

To investigate how investors react to social media outages we rely on daily investor-level administrative data from the *Comissão de Valores Mobiliários* (CVM), the Brazilian equivalent to the Securities and Exchange Commission (SEC). The data is highly reliable and contain all stock transactions between January 2012 and December 2019 on the B3, the only Brazilian stock exchange. The data is at the investor-stock-day level, and for each triple we observe shares bought and sold, number of transactions, and monetary volume for all investors, including institutions and individuals[52]. We complement the data with general company information as well as daily prices, volume, outstanding shares, and historical events from the Economatica database.

Brazilian stock market is the largest in Latin America with a market capitalization that fluctuated around U$ 1 trillion between 2012 and 2019. The final sample has 485 stocks, which were traded by 1,776,953 individual investors and 58,336 institutions, with over 3.9 billion trades and 133 million daily observations.

Table 2 reports descriptive statistics of variables studied in our empirical investigation. On an average stock-day, we have 167 retail investors, each trading an average volume of 52 thousand reais on a stock, and 61 institutions trading an average volume of 827 thousand reais each. Thus, on aggregate, institutions trade five times more monetary volume than retail investors on a stock-day. These variables are highly right skewed, with large standard deviations. Conditioned on trading, a median individual investor would trade 6.5 thousand reais

---

[51] During the period studied, Brazilian stock exchange had an irrelevant number of listed tech companies. Thus, it seems very unlikely that outages of social medias could impact the perception and perspectives of companies from different industries than tech and be driving the results.

[52] Data are de-identified, with investors assigned permanent reference numbers that allow us to follow them over time. This data also provides quarterly snapshots of all investor’s holdings in the stock market starting in December 2016. Additional information about the data is found on Birru *et. al.* (2022), and on Bonomo, Paiva and Ribeiro, (2022).

on a stock-day, way less than the average. When studying how exogenous social media outages may impact trading activity, we remodel our data to a stock-day panel structure. The second part of table 2 reports simple descriptive statistics for variables used in our regressions under this format.

Table 2 – Descriptive statistics of main variables.

Variables	**Investor Category**	**Mean**	**Median**	**Standard Deviation**	**Obs.**
Number of investors by Stock-Day	Retail	166.9	34	515.6	101140333
	Institutions	61.4	28	78.7	32056893
Monetary Volume by Investor-Stock-Day (Thousands R$)	Retail	52.0	6.5	453.0	101140333
	Institutions	827	82	5745	32056893
Monetary Volume aggreg. by Stock-Day (Thousands R$)	Retail	8674	540	43655	606109
	Institutions	50739	3107	168480	522458
Variables in Panel Structure (Stock-Day)					
Log (Volume + 1)	Retail	9.1	10.8	5.9	807700
	Institutions	8.7	10.3	7.2	807700
ΔLog (Volume + 1)	Retail	0.00	0	2.8	807700
	Institutions	0.00	0	2.7	807700
Log (# Investor + 1)	Retail	2.7	2.6	2.2	807700
	Institutions	2.0	1.6	2.0	807700
ΔLog (# Investor + 1)	Retail	0.00	0	0.7	807700
	Institutions	0.00	0	0.4	807700
Return		0.09	0	3.9	807700
Market Capitalization (Millions)		10229	1219.8	33853	807700
Log (Market Capitalization)		20.70	20.9	2.4	807700

Note: First three variables are conditioned on trading. That is, observation exists only if there is trading activity on that stock and day. Variables in panel structure complete the missing observations (days which the stock was not traded) with zeros. Monetary volume and number of investors comes from CVM trading data, while return and market capitalization comes from Economatica data.

4. Empirical exercise.

Our empirical exercise explores how a common shock to all investors at specific days affects their trading activity. Objectively, we study how daily measures of trading activity respond to days with social media outage, while controlling for other variables that may also affect trading activity. Given that the shock is common to all investors trading any stock, we propose a simple time-series regression for the beginning of our investigation:

$$\Delta Y_t = \beta * Outage_t + Controls_{t-1} + WeekDay_t + MonthDay_t + \ Monthly\ FE_t \qquad (1)$$

Where Y_t is a measure of trading activity, $Log(Monetary\ Volume + 1)_t$ in our initial specification. $Outage_t$ is a dummy variable equal to one on days identified as an outage day, and zero otherwise. Our coefficient of interest is β, which for small changes can be interpreted as the percentage change of trading volume on social media outage days. We use lagged values of return and market capitalization as controls and a set of weekday dummies and 'monthday'[53] dummies to control for seasonality. We also include Monthly FE, so our analysis is within the same month. That is, β is estimated using the within month variation and can be interpreted as the response of trading activity to outage days in relation to other days of the same month. Including Monthly FE is a very general form to control non-linear trends that may happen in eight years of sample. All standard errors from time-series regressions in this study were calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method.

Our data is at the investor-stock-day level; thus, we first need to mold the data to one time-series for each investor type. The first step is to construct a stock-day panel for each investor category summing the volume of all investors of the same category over each stock-day. The second step is to log transform our main dependent variable and take the first difference for each stock. The last step is to collapse the stock-day panel to a time-series. This is done in three fashions: equally weighing the stocks for each day, value weighting the stocks (by the last day of market capitalization of each stock), and volume weighting the stocks (by the average daily monetary volume of the last year). The different averages over the cross-section of the stocks are performed on the dependent variable, as on the control variables, return and market capitalization.

4.1 Outage days affecting trading monetary volume.

Table 3 reports results on how monetary volume responds to social media outages, an exogenous shock that make retail and institutional investors not being able to interact with these popular platforms. Panel A considers as outages all 33 interruptions described on table 1. All estimates of outage coefficient are negative for retail and for institutional investors. The negative coefficient suggests that both types of investors reduce their monetary volume on days

---

[53] 'Monthday' dummies are a set of twenty individual dummies for each trading day of the month, from -10 to +10, with -10 being the trading day ten days before the end of the month, - 1 being the last trading day of the month, +1 being the first trading day of the month, and so on and so forth.

in which their ability to use the platforms was impaired. This first result favors the hypothesis that, on average, investors use these platforms to communicate and obtain information, and when the platforms are not available, they shun away from trading. The competing hypothesis that investors would be distracted by social media and, therefore, trade more when they cannot access them, is not supported by the results reported on table 3.

Panel A of table 3 shows that the effect happens with both investor type and appears to be stronger for institutional investors. This result seems puzzling, institutional investors are considered more sophisticated, many times having access to internal research teams and other payable data and research information. Furthermore, for most professional funds, using these platforms may be discouraged during working hours or simply prohibited. At the same time, institutions are at the end of the day made by people who interact with social media platforms, even in their work time. Panel A also gives further insights about which type of stocks results are stronger. For both retail and institutions, point estimates are larger for the model equally weighting the stocks in each day. Not only that, but the statistical significance of the results is also stronger for the EW model. This suggests that the effects of outage in trading could be stronger for smaller and more illiquid stocks, given the weaker results for Value Weight and Volume Weight models. We explore these heterogeneities in section 5.

As discussed in the Data section, not all interruptions have the same impact on the ability of users to interact with the social media platforms. Some events are reported as a more generalized interruption that affects all users. Panel B shows how trading volume responds to the 18 interruptions that are classified as more severe. For those events, point estimates are more negative for retail and institutions, with also larger t-statistics. For severe events, traded volume falls 14.3% for retail investors and 12% for institutions for the EW model. This result is surprisingly similar to Peress (2014) estimates for trading volume response to newspaper strike days, which was -12%.

Table 3 – Social media outages affecting trading monetary volume.

	Dependent Variable: ΔLog(Monetary Volume + 1)					
Investor Category	**Institutions**			**Retail**		
Model	EW	ValueW	VolumeW	EW	ValueW	VolumeW
Panel A - All events						
Outage$_t$	-0.097** (-2.56)	-0.083* (-1.86)	-0.064 (-1.26)	-0.076* (-1.74)	-0.045 (-1.31)	-0.047 (-1.63)
Log(Market Capitalization)$_{t-1}$	-0.565*** (-2.92)	-0.354** (-2.14)	-0.438*** (-4.03)	-0.694*** (-3.31)	-0.528*** (-3.40)	-0.237*** (-4.30)
Return$_{t-1}$	-0.007 (-0.681)	-0.010 (-1.60)	-0.003 (-0.705)	0.011 (1.07)	-0.003 (-0.652)	0.005** (1.97)
Day of the month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Day of the week dummies	Yes	Yes	Yes	Yes	Yes	Yes
Monthly FE	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,973	1,954	1,974	1,973	1,954
R2	0.22	0.18	0.20	0.08	0.08	0.10
Panel B - Severe events						
Outage$_t$	-0.120** (-2.41)	-0.110** (-1.98)	-0.099* (-1.79)	-0.143** (-2.37)	-0.063 (-1.31)	-0.079** (-2.20)
Log(Market Capitalization)$_{t-1}$	-0.571*** (-2.99)	-0.353** (-2.16)	-0.440*** (-4.05)	-0.698*** (-3.37)	-0.527*** (-3.40)	-0.238*** (-4.33)
Return$_{t-1}$	-0.007 (-0.681)	-0.009 (-1.58)	-0.003 (-0.688)	0.011 (1.08)	-0.003 (-0.639)	0.005** (1.99)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,973	1,954	1,974	1,973	1,954
R2	0.22	0.18	0.20	0.08	0.08	0.10
Panel C - Weak events						
Outage$_t$	-0.068 (-1.21)	-0.049 (-0.736)	-0.022 (-0.266)	0.004 (0.067)	-0.023 (-0.481)	-0.008 (-0.190)
Log(Market Capitalization)$_{t-1}$	-0.569*** (-2.95)	-0.361** (-2.20)	-0.442*** (-4.05)	-0.701*** (-3.37)	-0.531*** (-3.55)	-0.239*** (-4.33)
Return$_{t-1}$	-0.007 (-0.699)	-0.010 (-1.59)	-0.003 (-0.690)	0.011 (1.05)	-0.003 (-0.649)	0.005** (1.98)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,973	1,954	1,974	1,973	1,954
R2	0.22	0.18	0.20	0.08	0.08	0.10
Panel D - Out of market hours events						
Outage$_t$	0.010 (0.297)	0.027 (0.702)	0.008 (0.231)	-0.018 (-0.480)	0.046 (1.48)	0.003 (0.109)
Log(Market Capitalization)$_{t-1}$	-0.573*** (-3.01)	-0.362** (-2.22)	-0.442*** (-4.05)	-0.701*** (-3.37)	-0.532*** (-3.56)	-0.239*** (-4.31)
Return$_{t-1}$	-0.007 (-0.700)	-0.009 (-1.57)	-0.003 (-0.681)	0.011 (1.05)	-0.003 (-0.636)	0.005** (1.98)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,973	1,954	1,974	1,973	1,954
R2	0.22	0.18	0.20	0.08	0.08	0.10

Note: All models are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method. Day of the month dummies are 20 dummies for different trading days of the month, from -10 to +10. Day of the week dummies are five weekday dummies. Monthly FE are dummies for every month of the sample (96 dummies). FE$^{\#}$ in panels B to D indicates the use of the same dummies as in panel A. EW model equally weights the stocks for each day of the stock-day panel. ValueW model weights the stocks by the last day of market capitalization of each stock, and VolumeW model weights the stocks by the average daily monetary volume of the last year. Panel A considers as outages all 33 interruptions described on table 1. Panel B considers the 18 severe outages described in table 1, while Panel C considers the 15 non-severe outages described in table 1. Panel D considers the 29 outages described in table A1. Outages outside market hours are shifted to the next available trading day.

Panel C shows the results of equation (1) when focusing only on the events that affected partial users or were less than one hour long. The outage coefficient is not statistically significant for this regression. As a sort of a placebo test, we run equation (1) considering the outage dummies equal to one on the 29 days of which there were interruption on the social media service outside market trading hours[54]. Panel D reports the coefficients of this regression. Estimated coefficients for the outage dummy are close to zero for both investors in all models, with no statistical significance.

Furthermore, on unreported results, we explored if the decrease in monetary volume was coming from purchases or sells, or if the participation of each investor type was increasing or decreasing on outage days (ratio of investors volume to total volume), none of these investigations had statistically significant results. Thus, the effect is similar on purchases and sell and among investors type. Another investigation was to check results independently for each social media. Although this investigation leaves fewer outages' observations for the regression, some significance was still present. Results were stronger for Facebook and Whatsapp (Table A4 on the appendix).

As a robustness of table 3 results, we also make modifications on equation (1) time fixed effects and change the dependent variable. On the appendix we show that using separately year dummies and month dummies together with a time trend, a less restricted model than Monthly FE, does not change the results for any of the models. Using weekly dummies, focusing in within week variation, also does not affect results. The appendix also reports results When the dependent variable is not in the first difference or using the ratio of monetary volume by market capitalization as dependent. Results with these two different variables remain qualitatively the same but are weaker.

4.2 Outage days affecting number of active investors.

Previous results showed that investors trading volume respond negatively to social media interruptions. One shortcoming of using trading volume is that results are driven by investors that trade higher monetary volume. That is, our results came from equally weighting

---

[54] If the interruption occurred during the early morning, late afternoon, or night, we consider the outage dummies equal to one on the same day of the interruption. For interruptions on Saturday's, we consider the previous day, Friday, as the outage day (dummy equal to one). For interruptions on Sunday's, we consider the next day, Monday. Finally, for interruptions that happened on holydays, we consider the outage dummy equal to one for the next following trading day.

the stocks, but volume weighting the investors within each investor type. Another shortcoming is that our data considers all investors and all trades. Thus, it could be that when one type of investor is affected by social media interruption and is trading less money volume, the other type of investor will be only a reactive counterpart. This would not be a problem if each type of investor traded only within your own group, but this is not realistic.

Our data is at the investor-day level, and we take the first step in exploring this information by studying how social media outages affect the number of active investors within each investor type. The measure $Log(\ \#Investors + 1)_t$, equal weights the investors and will not have previously described shortcomings, capturing the decision of whether to trade (i.e., the extensive margin).

Panel A of Table 4 reports how the first difference of $Log(\ \#Investors + 1)_t$ respond to social media outages. From now on we focus on EW models for conciseness[55]. Results tell us that investors not only trade lower amount on outage days (Table 3), but that some decide not to trade at all (Table 4), the propensity to trade is reduced. Like what happened with trade volume, effects are stronger in severe outages, with more negative coefficients and higher t-statistics. Point estimates are now more negative for retail investors, there is a contraction of 7.6 percent of investors being active on days when social media access is interrupted during trading hours. For institutions, the number of investors are the number of active funds trading on that day. Although in smaller magnitude, it is interesting that there is this negative correlation between social media outage and funds being active. Last, we want to call attention to the result of the coefficient of lagged return. The coefficient is positive and highly significant for retail investors, but not to institutions. Higher last day returns can be a proxy for attention grabbing which is more likely to affect retail investors (Barber and Odean, 2008).

Table 4 also reports results splitting institutional investors into domestic institutions and foreign institutions. Our sample of social media interruption is focused on events that happened in Brazil. Some of the outages are worldwide, but most are local or have stronger effects in Brazil. Foreign institutions are larger (trade higher volumes and have larger holdings) and are located outside Brazil. We expect that the results found for institutions would be stronger for domestic institutions, which were certainly affected by the interruption. Panel B shows that this is the case. The number of foreign funds participating in trading does not change during the

---

[55] Value Weight and Volume Weight models for number of investors as the regressand followed a qualitatively similar pattern as presented in Table 3 for monetary volume. As in Panel D of table 3, regressions using events occurring out of market hours were not statistically significant.

outage days, but it does so for domestic institutions. Panel C reports results for our first variable of interest, monetary volume. Although in smaller magnitude, monetary volume of foreign institutions is affected by social media interruption. However, foreign institutions are responsible for about half of total monetary volume. They certainly trade with domestic institutions and investors, and this result could be just a reaction to having less counterparts to trade with.

Table 4 – Social media outages affecting number of active investors and monetary volume for different investors categories.

Panel A

Dependent	**ΔLog(#Investors + 1)**					
Investor Category	**Institutions**			**Retail**		
Events	All	Severe	Weak	All	Severe	Weak
Outage$_t$	-0.021** (-2.26)	-0.036*** (-2.85)	-0.004 (-0.290)	-0.031 (-1.62)	-0.076** (-2.58)	0.021 (0.998)
Log(Market Capitalization)$_{t-1}$	-0.135*** (-2.62)	-0.136*** (-2.69)	-0.137*** (-2.60)	-0.291*** (-2.87)	-0.293*** (-2.93)	-0.296*** (-2.95)
Return$_{t-1}$	-0.001 (-0.666)	-0.001 (-0.656)	-0.001 (-0.689)	0.016*** (3.23)	0.016*** (3.26)	0.016*** (3.23)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974	1,974	1,974
R2	0.30	0.30	0.30	0.11	0.11	0.11

Panel B

Dependent	**ΔLog(#Investors + 1)**					
Investor Category	**Domestic Institutions**			**Foreign Institutions**		
Events	All	Severe	Weak	All	Severe	Weak
Outage$_t$	-0.026** (-2.35)	-0.049*** (-2.97)	0.002 (0.134)	-0.009 (-0.892)	-0.009 (-0.651)	-0.010 (-0.615)
Log(Market Capitalization)$_{t-1}$	-0.084 (-1.40)	-0.085 (-1.43)	-0.086 (-1.11)	-0.152*** (-3.07)	-0.152*** (-3.09)	-0.152*** (-3.00)
Return$_{t-1}$	-0.003 (-1.33)	-0.003 (-1.32)	-0.003 (-1.32)	0.003 (1.36)	0.003 (1.36)	0.003 (1.36)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974	1,974	1,974
R2	0.21	0.21	0.21	0.33	0.33	0.33

Panel C

Dependent	**ΔLog(Monetary Volume + 1)**					
Investor Category	**Domestic Institutions**			**Foreign Institutions**		
Events	All	Severe	Weak	All	Severe	Weak
Outage$_t$	-0.109*** (-2.96)	-0.144** (-2.56)	-0.068 (-1.48)	-0.054 (-1.55)	-0.064* (-1.67)	-0.040 (-0.679)
Log(Market Capitalization)$_{t-1}$	-0.503** (-2.45)	-0.510** (-2.55)	-0.508** (-2.49)	-0.479*** (-3.40)	-0.483*** (-3.51)	-0.481*** (-3.41)
Return$_{t-1}$	-0.011 (-0.996)	-0.011 (-0.995)	-0.012 (-1.02)	0.006 (0.842)	0.006 (0.847)	0.006 (0.826)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974	1,974	1,974
R2	0.19	0.19	0.19	0.24	0.24	0.24

Note: All models are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method. Fixed-Effects include 20 day of the month dummies, 5 day of the week dummies, and 96 monthly dummies. All models equally weigh the stocks for each day of the stock-day panel. "All events" consider as outages all 33 interruptions described on table 1. "Severe events" consider the 18 severe outages described in table 1, while weak events consider the 15 non-severe outages described in table 1.

5. Heterogeneity effects of social media outages.

On this section we explore how social media outages might affect the trading activity on different firms and on different investors within an investor category. That is, we explore two different dimensions of sources of heterogeneity in our results.

5.1 Firm level heterogeneity.

Table 3 different weighting models gave a hint that there might be heterogeneity of the results between stocks. Results were stronger for Equally weighted model than Value and Volume weighted models, which suggests that social media affected more the trading volume of small and illiquid stocks. To better explore this and other firm level heterogeneity we now move on from the time-series model to a stock-day panel model. Our new equation of interest becomes:

$$\Delta Y_{f,t} = \beta * Outage_t + \rho * Outage_t X Firm\ Characteristic_{f,t-1} +$$

$$Controls_{f,t-1} + WeekDay_t + MonthDay_t + MonthlyFirm\ FE_{f,t} \qquad (2)$$

Where $\Delta Y_{f,t}$ is the first difference of the log transformation of each stock trading volume. Outage dummies remain the same as in equation (1), but now we also interact different stock level characteristics with the outage dummy to capture the reaction on different stocks. Our control variables, return and market capitalization, are also at the stock-day level. Monthly fixed effects are now at the stock level, which allows for each stock to have a very general form of non-linear trend, and coefficients of equation (2) are estimated with within stock-month variation. The other time dummies to capture seasonality remain the same. Standard errors are now clustered by day and stock, given that we have both cross-section dependence and time-series dependence in our panel.

Table 5 – Firm level heterogeneity of social media outages affecting trading monetary volume.

Dependent Variable: ΔLog(Monetary Volume + 1)

Panel A - Retail	**All Events**						**Severe Events**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outage_t	-0.075* (-1.8)	-0.075* (-1.8)	-0.074* (-1.7)	-0.075* (-1.8)	-0.075* (-1.8)	-0.075* (-1.8)	-0.141** (-2.5)	-0.141** (-2.5)	-0.138** (-2.4)	-0.140** (-2.5)	-0.132** (-2.3)	-0.129** (-2.3)
$\text{Return}_{f,t-1}$	0.001 (0.81)	0.001 (0.82)	0.001 (0.9)	0.0010 (0.788)	0.0009 (0.74)	0.001 (0.82)	0.001 (0.82)	0.001 (0.82)	0.001 (0.93)	0.0010 (0.79)	0.0009 (0.75)	0.001 (0.84)
$\text{Log(Market Capitalization)}_{f,t-1}$	-0.273*** (-5.2)	-0.274*** (-5.3)	-0.273*** (-5.3)	-0.270*** (-5.2)	-0.270*** (-5.2)	-0.268*** (-5.2)	-0.274*** (-5.3)	-0.274*** (-5.3)	-0.274*** (-5.3)	-0.270*** (-5.2)	-0.270*** (-5.2)	-0.268*** (-5.2)
Outage_t * $\text{Log(Market Capitalization)}_{f,t-1}$		0.046* (1.8)				0.045* (1.8)		0.061* (1.7)				0.060* (1.7)
Outage_t * $\text{Return}_{f,t-1}$			-0.007 (-0.91)			-0.007 (-0.87)			-0.016* (-1.7)			-0.015 (-1.6)
$\text{Return}_{f,t-2:t-5}$				0.006*** (3.1)		0.005*** (2.9)				0.006*** (3.1)		0.005*** (2.8)
Outage_t * $\text{Return}_{f,t-2:t-5}$				-0.003 (-0.25)		-0.002 (-0.2)				-0.007 (-0.57)		-0.003 (-0.26)
$\text{Return}_{f,t-6:t-22}$					0.009** (2.2)	0.007* (1.9)					0.009* (2.30)	0.008** (2.0)
Outage_t * $\text{Return}_{f,t-6:t-22}$					-0.004 (-0.16)	-0.0008 (-0.03)					-0.054* (-1.83)	-0.051* (-1.9)
Day of the month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of the week dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Monthly FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	807,215	807,215	807,215	807,215	807,214	807,214	807,215	807,215	807,215	807,215	807,214	807,214
R2	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004

Panel B - Domestic Institutions	**All Events**						**Severe Events**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outage_t	-0.109*** (-2.9)	-0.109*** (-2.9)	-0.107*** (-2.9)	-0.109*** (-2.9)	-0.109*** (-3.0)	-0.108*** (-2.9)	-0.143*** (-2.7)	-0.143*** (-2.7)	-0.138*** (-2.6)	-0.141*** (-2.6)	-0.144*** (-2.8)	-0.138*** (-2.7)
$\text{Return}_{f,t-1}$	-0.008*** (-4.9)	-0.008*** (-4.9)	-0.007*** (-4.7)	-0.008*** (-4.9)	-0.008*** (-4.9)	-0.008*** (-4.8)	-0.008*** (-4.9)	-0.008*** (-4.9)	-0.007*** (-4.8)	-0.008*** (-4.9)	-0.008*** (-4.9)	-0.008*** (-4.8)
$\text{Log(Market Capitalization)}_{f,t-1}$	-0.302*** (-7.2)	-0.302*** (-7.2)	-0.301*** (-7.2)	-0.299*** (-7.2)	-0.298*** (-7.2)	-0.297*** (-7.2)	-0.302*** (-7.2)	-0.302*** (-7.2)	-0.302*** (-7.2)	-0.300*** (-7.2)	-0.299*** (-7.2)	-0.298*** (-7.2)
Outage_t * $\text{Log(Market Capitalization)}_{f,t-1}$		0.015 (0.57)				0.013 (0.51)		0.010 (0.36)				0.007 (0.25)
Outage_t * $\text{Return}_{f,t-1}$			-0.015 (-1.6)			-0.015 (-1.6)			-0.027** (-2.0)			-0.027* (-2.1)
$\text{Return}_{f,t-2:t-5}$				0.004* (2.0)		0.003* (1.8)				0.004** (2.1)		0.003* (1.9)
Outage_t * $\text{Return}_{f,t-2:t-5}$				0.0002 (0.02)		0.0004 (0.05)				-0.014 (-1.2)		-0.014 (-1.3)
$\text{Return}_{f,t-6:t-22}$					0.007* (1.8)	0.006 1.61)					0.007* (1.8)	0.006 (1.6)
Outage_t * $\text{Return}_{f,t-6:t-22}$					-0.002 (-0.10)	0.002 (0.10)					-0.0008 (-0.04)	0.009 (0.45)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	807,215	807,215	807,215	807,215	807,214	807,214	807,215	807,215	807,215	807,215	807,214	807,214
R2	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
Within R2	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003

Note: All models are estimated with standard errors clustered by day and stock. Day of the month dummies are 20 dummies for different trading days of the month, from -10 to +10. Day of the week dummies are five weekday dummies. Firm-Monthly FE are a set of dummies for every month of the sample (96 dummies) for each firm (485), a total of 46,560 dummies. Fixed-Effects in panel B indicate the use of the same dummies as in panel A. "All events" consider as outages all 33 interruptions described on table 1, while "severe events" consider the 18 severe outages described in table 1.

Table 5 reports twelve different models for retail investors (Panel A) and for domestic institutions (Panel B)[56] using either all interruptions or only severe interruptions of social media outages. Our first observation on the results is that column (1) and (7) for both panels, which are models without interaction with stock-level characteristics, are similar in point estimates and significance to what is reported on Table 3 for the equally weighted model. This is expected given our full set of stock-month fixed effects and standard error clusterization method.

Columns (2) and (8) of Panel A shows that retail investors respond less negatively to social media outages on stocks of larger firms. If retail investors use social media to gather information about stocks they follow and invest, it is reasonable to think that this effect is stronger for stocks that have less media and news attention, leaving social media platforms as one of the main sources for information for these investors. With the interruption, the small stocks would be the most affected, and thus its volume.

For severe interruption, firms with recent higher past returns appear to be more affected (columns 9, 10, and 11), with their trading volume reducing even more on outage days. We conjecture that this might be associated with the empirical pattern that large past returns are attention grabbing events for stocks, which increase retail volume. Without social media platforms to spread which were those stocks with strong past returns, their volume will drop disproportionately more when compared to stocks with strong past returns on non-outage days. Our reasoning relates to social media as a broadcast of information like newspaper. For example, Peress (2014) showed evidence that newspaper was important in spreading-out last day market news.

Panel B shows that institutional investors do not react differently to different firm sizes on outage days. The coefficients of the interaction, although positive, are small and not statistically significant. Stocks with strong one day past returns are also more affected for institutions. It is harder to justify this as due to attention grabbing, institutions in general would have the information about which were the stocks with higher previous day return regardless of social media. In general, retail investors' trading activity (panel A) is more affected by recent past returns than institution trading activity (panel B), with higher coefficients and t-statistics on the past returns' variables.

---

[56] The rest of the results of the papers focus on domestic investors (institutions or retail). Section 4 suggests that foreigners funds are less affected by social media outages. We confirm this outcome remains for the different models of section 5. These results are available upon request.

Table A5 on the appendix reports the same regression results as in table 6 with number of investors as the dependent variable. Overall, results for firm heterogeneity are less strong for the number of investors as the dependent than for monetary volume. On unreported results, we also investigated whether the effect was different for lottery-like[57] stocks as found in Mohr (2022). Although the coefficient is negative for all models, it is not statistically different from zero, thus the outage effect does not appear to be stronger on lottery-like stocks.

5.2 Investor level heterogeneity.

Investor heterogeneity within categories is also explored using the stock-day panel model, but now investors of the same type are split into new categories, allowing different response to outages for each one of them. Our estimated equation is:

$$\Delta Y_{i,f,t} = \beta_i * Outage_t \; X \; Investor\ Category_i + \ Investor\ Category_i \ + \\ Controls_{f,t-1} + \ WeekDay_t + MonthDay_t + MonthlyFirm\ FE_{f,t} \qquad (3)$$

Where $\Delta Y_{i.f,t}$ is the first difference of the log transformation of each stock f and investor category i trading volume or number of investors. The only difference is that we include a dummy for each investor category i and interact it with the outage dummy. All the other variables remain the same as in equation (2)

Our main investigation in this section is how social media outage affects investors with different account sizes (total value of holdings in stocks), for retail and institutions. We split each investor type into four additional groups by account sizes for every day of the sample. We split investors into four groups that hold the same amount of wealth in stocks. That is, these groups have different numbers of investors, but the same wealth (25% each). There is a high inequality of portfolio sizes among retail and institutions (see Bonomo, Paiva and Ribeiro, 2022), so the wealthier group will have few investors with large holdings, while the poorer group will have many investors with smaller holdings. The data for institutions is at the fund level, so it could be that one fund of an institution will be classified as small, while a different fund from the same institution will be classified as large. Because the data is anonymized, we cannot recover which funds have the same ‘parent’ institution.

---

[57] We define lottery-like stocks at a monthly frequency based on three criteria defined in Kumar (2009): i) above median idiosyncratic volatility, ii) above median idiosyncratic skewness, and iii) below median price.

Table 6 – Investor level heterogeneity of social media outages affecting trading monetary volume and number of active investors.

Panel A - Retail

Dependent	**ΔLog(Monetary Volume + 1)**		**ΔLog(# Investors + 1)**	
Events	All	Severe	All	Severe
	(1)	(2)	(3)	(4)
Outage$_t$ * Wealth Group 1	-0.058 (-1.58)	-0.108** (-2.06)	-0.026 (-1.57)	-0.063*** (-2.66)
Outage$_t$ * Wealth Group 2	-0.082** (-2.01)	-0.163*** (-3.06)	-0.028 (-1.62)	-0.069*** (-2.81)
Outage$_t$ * Wealth Group 3	-0.035 (-0.771)	-0.093 (-1.39)	-0.013 (-1.12)	-0.036** (-2.05)
Outage$_t$ * Wealth Group 4	-0.061 (-0.909)	-0.190** (-2.04)	-0.008 (-0.948)	-0.025** (-2.19)
Return$_{f,t-1}$	0.004*** (4.45)	0.004*** (4.45)	0.002*** (7.23)	0.002*** (7.23)
Log(Market Capitalization)$_{f,t-1}$	-0.212*** (-5.96)	-0.212*** (-5.97)	-0.062*** (-6.49)	-0.062*** (-6.50)
Wealth Group 2	0.0003 (0.166)	0.0004 (0.227)	-0.0004 (-0.606)	-0.0004 (-0.585)
Wealth Group 3	-0.0005 (-0.148)	-0.0002 (-0.074)	-0.0009 (-0.719)	-0.0009 (-0.748)
Wealth Group 4	0.0002 (0.039)	0.0009 (0.171)	-0.001 (-0.671)	-0.001 (-0.704)
Fixed-Effects	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	3,228,860	3,228,860	3,228,860	3,228,860
R2	0.002	0.002	0.004	0.004

Panel B - Domestic Institutions

Dependent	**ΔLog(Monetary Volume + 1)**		**ΔLog(# Investors + 1)**	
Events	All	Severe	All	Severe
	(1)	(2)	(3)	(4)
Outage$_t$ * AUM Group 1	-0.066* (-1.67)	-0.130** (-2.21)	-0.022* (-1.86)	-0.053*** (-3.07)
Outage$_t$ * AUM Group 2	-0.086** (-2.13)	-0.137** (-2.56)	-0.021** (-2.06)	-0.039*** (-2.87)
Outage$_t$ * AUM Group 3	-0.066 (-0.998)	-0.121 (-1.24)	0.0007 (0.075)	-0.001 (-0.094)
Outage$_t$ * AUM Group 4	0.009 (0.091)	-0.062 (-0.632)	-0.007 (-0.704)	-0.010 (-0.764)
Return$_{f,t-1}$	-0.003*** (-3.75)	-0.003*** (-3.75)	-0.0002** (-2.44)	-0.0002** (-2.44)
Log(Market Capitalization)$_{f,t-1}$	-0.149*** (-5.78)	-0.149*** (-5.79)	-0.020*** (-4.86)	-0.020*** (-4.87)
AUM Group 2	-0.0002 (-0.053)	-0.0004 (-0.146)	-0.0001 (-0.118)	-0.0002 (-0.215)
AUM Group 3	-0.0004 (-0.094)	-0.0005 (-0.114)	-0.0004 (-0.294)	-0.0005 (-0.352)
AUM Group 4	-0.002 (-0.168)	-0.002 (-0.123)	-0.0004 (-0.218)	-0.0005 (-0.297)
Fixed-Effects	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	3,228,860	3,228,860	3,228,860	3,228,860
R2	0.002	0.002	0.004	0.004

Note: All models are estimated with standard errors clustered by day and stock. Fixed-Effects include twenty days of the month dummies, five days of the week dummies, and 46,560 firm-monthly dummies. "All events" consider as outages all 33 interruptions described on table 1, while "severe events" consider the 18 severe outages described in table 1. Within each broad investor category, retail and domestic institutions, investors are split into four additional groups by account sizes, leaving the groups holding the same wealth or asset under management in stocks, for retail and domestic funds, respectively. Groups with investors with less wealth are represented by lower numbers. Domestic institutions are at the fund level, we cannot recover which funds have the same 'parent' institution.

Panel A of table 6 reports the results for retail investors with different wealth, with groups with investors with less wealth represented by lower numbers. Groups are sorted leaving the total wealth on stocks the same. Point estimates and statistical significance are generally larger (absolute terms) for investors with lower wealth. That is, groups with many investors with lower wealth appear to be more affected. This is especially true for the model which uses number of investors as the dependent (that equal weights the investor), columns (3) and (4). For the model that volume weights the investor there is a strong response in column (2), which could be biased by few investors that trade large amounts among the rich investors. These results suggest that the very wealthy retail investors appear to react less to social media outages.

Panel B of table 6 show results for domestic institutions, separating funds in groups with the same total asset under management (in stocks). Smaller funds are the only ones to react to the social media outages, when accounting for monetary volume and number of funds being active. The few and large funds (groups 3 and 4) do not react to the outages in any model (1 to 4). Together with results in section 4.2, that foreign funds (which are larger) do not react to social media outages, this result suggest that among the investors that were initially though as more sophisticated, the ones that are likely to be the most sophisticated ones do not react to the outages.

We also explored if the reaction to outage was concentrated in more active investors (Table A6 on the appendix). Results for investors with different activity patterns were mixed. But overall, results seem stronger for highly active retail investors and, in contrast, stronger for least active investors for domestic institutions. On unreported results we also explored two other possible sources of heterogeneity at the investor level; (i) we grouped investors by their trade size, and (ii) split investors into investors with new accounts (opened after 2016) and older accounts. None of these groups appear to have differences in their response to days with social media outages. Unfortunately, we do not have social demographic information on the investors. Therefore, any heterogeneity analysis comes from information that we can extract from holdings and transactions, such as account size, turnover, or other trading patterns.

6. Social media outages effect on market liquidity.

Does social media help to alleviate market frictions and improve market liquidity? On this section we try to answer this question by looking how liquidity measures are affected when social media is exogenously down. Understanding if social media improves market liquidity is important given that lack of liquidity can reduce market efficiency (Chordia, Roll, and Subrahmanyam, 2008). The increase in trading volume for different investors categories is already a sign of improved liquidity (Chordia, Roll, and Subrahmanyam, 2001). The following investigation is at the stock-day level, it measures the reaction of the entire market, a broad market response, not of a specific investor group.

To test how social media affects market liquidity, we rely on three measures. First, we repeat our monetary volume measure, but aggregating all investors for each stock. As noted by Chordia, Roll, and Subrahmanyam (2001), monetary volume can also be viewed as a liquidity variable. Second, we use turnover, shares traded as a proportion of shares outstanding. Third, we use the Amihud (2002) measure of illiquidity, the absolute value of return by the monetary volume traded (all scaled by 10^6), a measure of price reaction to volume. Higher measures of volume and turnover indicate more active markets, as do lower measures of Amihud proxy.

Table 7 – Social media outages effect on market liquidity

	Mon. Volume	**Turnover**	**Amihud Illiquidity**
Panel A - All Events	(1)	(2)	(3)
Outage$_t$	-0.062** (-2.37)	-0.011*** (-2.60)	0.042** (2.22)
Return$_{f,t-1}$	0.002*** (3.89)	0.0005*** (4.11)	-0.018*** (-5.16)
Log(Market Capitalization)$_{f,t-1}$	-0.223*** (-8.91)	-0.044*** (-6.55)	0.173*** (4.12)
Fixed-Effects	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day
Observations	807,702	807,702	796,484
R2	0.007	0.01	0.004
Panel B - Severe Events	(5)	(6)	(7)
Outage$_t$	-0.092*** (-2.85)	-0.016** (-2.56)	0.042* (1.84)
Return$_{f,t-1}$	0.002*** (3.90)	0.0005*** (4.11)	-0.018*** (-5.16)
Log(Market Capitalization)$_{f,t-1}$	-0.223*** (-8.92)	-0.044*** (-6.56)	0.173*** (4.13)
Fixed-Effects	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day
Observations	807,702	807,702	796,484
R2	0.0087	0.01	0.004

Note: All models equally weight the stocks for each day and are estimated with standard errors clustered by day and stock. The three dependent variables are in their logarithm differences - first we take the logarithm for each day, and then the difference between days. Monetary volume aggregates all investors for each stock. Turnover is shares traded as a proportion of shares outstanding. Amihud Illiquidity is the absolute value of return by the monetary volume traded (all scaled by 10^6). Fixed-Effects include twenty days of the month dummies, five days of the week dummies, and 46,560 firm-monthly dummies. “All events” consider as outages all 33 interruptions described on table 1, while “severe events” consider the 18 severe outages described in table 1.

In Table 7 the three dependent variables are in their logarithm differences. Panel A and B show that market liquidity seems to improve with social media, and worsen in social media outages, with reduction of volume and turnover, and increase in Amihud illiquidity proxy. On unreported results we explore i) if the liquidity outcomes change in stocks with high retail ownership, and ii) if liquidity outcomes change due to firm market capitalization. On both exercises there was no difference in outcomes.

Retail traders are generally considered more noisy traders than institutions, and trade smaller stocks. From previous sections we know that retail traders (noise traders) leave the market. Under the adverse selection hypothesis (Section 2), liquidity should worsen in that scenario, because adverse selection risk increases. Results do not corroborate this hypothesis. Neither measure had a stronger effect (increase in illiquidity) for smaller stocks or for stocks with high retail ownership.

Overall, results support the idea of social media as a source of information and communication, increasing market activity and improving liquidity and likely market efficiency. Next section we provide evidence that social media does spread company information to investors, looking at the effect of social media outages in the return response to firms' announcements.

7. Companies' announcements on outage days.

Previous results suggest that investors use social media as an information tool, with investors reducing their participation on days when social media is down. We further corroborate this hypothesis by showing that social media outages affect how companies' announcements is incorporated into firm prices. If investors use social media to learn about companies' news, information diffusion will be impaired on social media outage days, affecting how news is incorporated into prices. Peress (2014) showed that newspapers strike affect return autocorrelation, which indicated that newspapers help prices incorporate news. We take a similar approach studying how social media outages affect return autocorrelation, focusing on days when there is news about firms, companies' announcements days.

We collect firms' announcements from the public CVM database of "Consultation of Listed Company Documents" (*Consulta de Documentos de Companhias Abertas*). All listed firms in B3 Brazilian stock exchange are required to publicize all its documents and announcements in the CVM database at the same time (or before) of making it public in their website or in investors relation calls. For our sample of 485 companies, we have 52,581 firm-day announcements during the eight-year period under study. On average 32 different firms have an announcement on a given day, with 83% of the trading days having some announcement of any firm. These firms' documents include earnings announcements, but also any other relevant information, such as mergers, spin-off, dividends, general operation guidance, and others[58].

Table 8 report the autocorrelation of returns under different circumstances. We regress next day return $(t+1)$ on today's return (t) using the same set of fixed effects as in other firm-day panel regressions, weekday, monthday and firm-year-month fixed effects. We also interact today's return with different terms, such as an outage indicator variable, and an announcement indicator variable. We set the indicator variable $announcement_{f,t}$ equal to one if an announcement occurs on day t for stock f, and zero otherwise. Here, the dependent indicator variables and returns are measured contemporaneously. Panel B study the autocorrelation between todays return and next week return, accumulated $t+1$ to $t+5$.

Model (1) shows that there is a negative autocorrelation in returns, suggesting the existence of a one-day reversal. This reversal could be a mean reversal effect after investors overreaction (maybe concentrated in extreme returns). Model (2) shows that returns backed by news do not have reversals. This result is found on both panels, for the next day return and next week return, meaning that a positive return on an announcement day do not reverse the day after, for example. Models (3) and (5) indicate that on outage days the reversal is also smaller, which might suggest that overreaction is less significant when social media does not help to spread information. Another possible explanation is if investors follow a short-term contrarian strategy and use social media to gather the signal for their strategy. When social media is down, investors would not trade as contrarians as much and reduce the reversal pattern.

---

[58] Specifically, we select five different types of announcements on the CVM documents database: "*DFP - Demonstrações Financeiras Padronizadas*", "*ITR - Informações Trimestrais*", "*Fato Relevante*", "*Aviso aos Acionistas*", "*Comunicado ao Mercado*". We use only the announcement date of the first document version available, that is, the first time the information became public. For documents made public after trading hours, weekends, and holydays, we shift their announcement date to the next trading day.

Table 8 – Outages and companies announcements effect on return autocorrelation.

Return$_{f,t+1}$

Panel A - Return$_{f,t+1}$			**All Events**		**Severe Events**	
	(1)	(2)	(3)	(4)	(5)	(6)
Return$_{f,t}$	-0.102*** (-4.57)	-0.116*** (-5.00)	-0.103*** (-4.59)	-0.118*** (-5.04)	-0.103*** (-4.60)	-0.117*** (-5.03)
Ret$_{f,t}$ * Announcement$_{f,t}$		0.129*** (5.21)		0.132*** (5.37)		0.130*** (5.24)
Ret$_{f,t}$ * Outage$_{t}$			0.059 (1.44)	0.083* (1.88)	0.107** (2.29)	0.120** (2.47)
Ret$_{f,t}$ * Announcement$_{f,t}$ * Outage$_{t}$				-0.217** (-2.04)		-0.101 (-1.14)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	807,159	807,159	807,159	807,159	807,159	807,159
R2	0.06	0.06	0.06	0.06	0.06	0.06

Return$_{f,t+1:t+5}$

Panel B - Return$_{f,t+1:t+5}$			**All Events**		**Severe Events**	
	(1)	(2)	(3)	(4)	(5)	(6)
Return$_{f,t}$	-0.158*** (-5.89)	-0.176*** (-6.54)	-0.158*** (-5.86)	-0.177*** (-6.54)	-0.158*** (-5.89)	-0.176*** (-6.54)
Ret$_{f,t}$ * Announcement$_{f,t}$		0.166*** (4.67)		0.171*** (4.79)		0.166*** (4.66)
Ret$_{f,t}$ * Outage$_{t}$			0.027 (0.343)	0.059 (0.779)	0.045 (1.08)	0.051 (1.30)
Ret$_{f,t}$ * Announcement$_{f,t}$ * Outage$_{t}$				-0.288* (-1.76)		0.008 (0.068)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	805,259	805,259	805,259	805,259	805,259	805,259
R2	0.23	0.23	0.23	0.23	0.23	0.23

Note: All models are estimated with standard errors clustered by day and stock. Panel A dependent variable is the next day return ($t+1$), while Panel B dependent variable is next week return, accumulated $t+1$ to $t+5$. $Announcement_{f,t}$ is a dummy variable equal to one if an announcement occurs on day t for stock f, and zero otherwise. Firms' announcements days come from the public CVM database of "Consultation of Listed Company Documents" (*Consulta de Documentos de Companhias Abertas*). Fixed-Effects include twenty day of the month dummies, five days of the week dummies, and 46,560 firm-monthly dummies. "All events" consider as outages all 33 interruptions described on table 1, while "severe events" consider the 18 severe outages described in table 1.

Models (4) and (6) investigate if the effect of announcement changes on outage days. If social media helps to spread the information of the announcement, this means that the announcement effect on prices will be smaller on days when there is no social media. That is what we find in column (4), the positive effect that announcements and outages days have on its own - reducing the reversal pattern - is weakened, a negative coefficient for the triple interaction. That is, without social media, announcements could have no impact on the return autocorrelation. Overall, the results support the social media news information diffusion hypothesis.

8. Investor Herding.

On days when investors are unable to interact with social media, we observe fewer investors trading stocks. Another investigation is to explore if these days their trading activity is less correlated with one another. Social media could be a vehicle by which investors are exposed to the same information and trading analysis, and therefore trade in similar direction. Herding is when a group of investors follow each other into the same stocks over a period. That is, investors from the same categories will end up on the same side of a trade (purchasing or selling) in a higher proportion that would be observed if these choices were random. Empirical evidence shows that institutions and retail investors herd on their trading activity[59].

The most popular measure to detect herding among investors was proposed by Lakonishok et al. (1992) (LSV from now on), which assesses correlated trading behavior of investors groups within a period. The LSV measure is as follows:

$$LSV_{i,t} = \frac{1}{\sum_{t=1}^{T} N_t} 1 \sum_t \sum_f \left(\left| br_{i,f,t} - \overline{br_{\imath,t}} \right| - E_t\left[\left| br_{i,f,t} - \overline{br_{\imath,t}} \right| \right] \right), \qquad (4)$$

Where N_t is the number of stocks traded during period *t*, $br_{i,f,t}$ is the number of investors of group *i* there are net buyers of stock *f* during period *t* divided by the number of active investors of group *i* on stock *f* during *t*, a buyer's ratio. The period-average buyers' ratio, $\overline{br_{\imath,t}}$, is the number of net buyers aggregated across all stocks for investors of group *i* during *t* divided by the number of all active investors of group *i* during *t*. The average subtraction controls for aggregate shifts of investors of a group in and out of market. The expectation term subtracted

---

[59] See Lakonishok, Shleifer and Vishny (1992), Grinblatt, Titman and Wermers (1995), and Sias (2004) for institutional investors and Dorn, Huberman and Sengmueller (2008) and Barber, Odean and Zhu (2009) for retail investors.

inside the summation is a correction proposed by LSV to get a test statistic that is zero under the null hypothesis of no correlated trading[60]. Note that the measure is calculated at the stock-investor level, and then averaged across stocks. The greater the value of $LSV_{i,t}$ the higher is the intensity of herding for that investor group. For example, a value of 10% suggests that on average 60% of the investors are on one side of the market and 40% are on the other side, assuming that the average position change is expected to be a 50%-50% split for this example.

We also explore investor herding with a second measure, from Sias (2004). Sias measure exams the cross-sectional correlation between a group of investors from one period to the next period, which is different from the LSV measure that looks at the cross-sectional temporal dependence within a period. Sias herding measure is the coefficient ($\beta_{i,t}$) calculated from the cross-sectional regressions of the standardized buyer's ratio of group *i* on stock *f* at period *t* on the standardized buyer's ratio of group *i* on stock *f* at period *t-1*:

$$\widetilde{br_{i,f,t}} = \beta_{i,t} * \widetilde{br_{i,f,t-1}} + \varepsilon_{i,f,t}\,, \qquad where$$

$$\widetilde{br_{i,f,t}} = \frac{br_{i,f,t} - \overline{br_{i,t}}}{\sigma\left(br_{i,f,t}\right)} \tag{5}$$

Where $\overline{br_{i,t}}$ is the cross-sectional average of the buyer's ratio and $\sigma\left(br_{i,f,t}\right)$ is the cross-sectional standard deviation, both over the *f* stocks in period *t*. The measure is calculated for each period *t*, and then averaged across the time-series, using the t-statistics computed from the time-series standard error.

Table 9 reports the LSV and Sias herd measures for three different investors categories and how these measures respond to social media interruptions. The measures are obtained considering a day as the period in equation (4) and (5). Panel A focuses on LSV herding measure, while Panel B focuses on Sias herding measure. Herding behavior does not seem to be affected by social media outages. Coefficients for the outage dummy are not statistically significant for any of the investor's category or herding measures. On unreported results, we explored the LSV herding measure as a dependent variable for the equations of section 5. No statistically significant heterogeneity was found. Due to the calculation procedure of the Sias herding measure, we are unable to use this measure in a stock-day panel structure.

---

[60] The expectation term accounts for the fact that we expect to observe more variation in the buyers ration in stocks with few investors trading. Please refer to Lakonishok et al. (1992) for details on the correction.

Table 9 – Social media outages affecting the correlated trading of investors.

Panel A - LSV herding measure

Dependent			LSV_t			
Investor Category	**Retail**		**Domestic Inst.**		**Foreign Inst.**	
Events	All	Severe	All	Severe	All	Severe
$Outage_t$	-0.019 (-0.072)	-0.127 (-0.983)	0.323 (1.60)	-0.261 (-0.847)	-0.065 (-0.412)	0.450 (1.52)
Log(Market Capitalization)$_{t-1}$	-3.58 (-1.23)	-2.02 (-1.41)	-1.74 (-1.08)	-3.64 (-1.14)	-2.00 (-1.40)	-1.73 (-1.12)
$Return_{t-1}$	-0.156*** (-2.61)	0.037 (1.19)	0.026 (0.760)	-0.155** (-2.48)	0.036 (1.18)	0.026 (0.768)
$Return_{t-2:t-5}$	-0.308** (-2.37)	0.037 (0.495)	0.180** (2.23)	-0.306** (-2.26)	0.036 (0.474)	0.181** (2.26)
$Return_{t-6:t-22}$	-0.112 (-0.249)	-0.130 (-0.686)	0.485** (2.35)	-0.107 (-0.226)	-0.136 (-0.719)	0.492** (2.40)
Ave. Monet. Volume$_{t-1:t-22}$	0.233 (0.470)	0.401 (1.04)	0.617 (1.19)	0.245 (0.451)	0.399 (1.04)	0.600 (1.19)
Ave. $LSV_{t-1:t-22}$	-0.651*** (-3.22)	-0.955*** (-6.68)	-0.422*** (-3.00)	-0.653*** (-3.07)	-0.956*** (-6.85)	-0.429*** (-3.22)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,953	1,953	1,953	1,953	1,953	1,953
R2	0.31	0.17	0.29	0.31	0.17	0.29
LSV - mean	0.1***		0.059***		0.068***	
LSV - SD	0.148		0.122		0.117	
Obs.	604426		509663		418164	

Panel B - Sias herding measure

Dependent			$Sias_t$			
Investor Category	**Retail**		**Domestic Inst.**		**Foreign Inst.**	
Events	All	Severe	All	Severe	All	Severe
$Outage_t$	-0.0002 (-0.019)	0.015 (1.23)	-0.005 (-0.274)	-0.012 (-0.969)	0.010 (0.656)	-0.004 (-0.168)
Log(Market Capitalization)$_{t-1}$	0.024 (0.175)	-0.030 (-0.174)	0.110 (0.572)	0.022 (0.156)	-0.032 (-0.186)	0.111 (0.576)
$Return_{t-1}$	-0.009*** (-3.20)	-0.002 (-0.477)	-0.006 (-1.61)	-0.009*** (-3.19)	-0.002 (-0.464)	-0.006 (-1.62)
$Return_{t-2:t-5}$	-0.002 (-0.255)	-0.005 (-0.572)	-0.006 (-0.573)	-0.002 (-0.243)	-0.005 (-0.553)	-0.006 (-0.575)
$Return_{t-6:t-22}$	-0.028 (-1.36)	-0.004 (-0.154)	0.009 (0.317)	-0.028 (-1.35)	-0.003 (-0.124)	0.009 (0.309)
Ave. Monet. Volume$_{t-1:t-22}$	0.069*** (2.86)	-0.040 (-1.06)	-0.082 (-1.28)	0.069*** (2.89)	-0.040 (-1.06)	-0.082 (-1.28)
Ave. $Sias_{t-1:t-22}$	-0.929*** (-6.17)	-0.963*** (-6.51)	-0.673*** (-4.64)	-0.934*** (-6.22)	-0.964*** (-6.52)	-0.673*** (-4.67)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW
Observations	1,952	1,952	1,952	1,952	1,952	1,952
R2	0.45	0.25	0.38	0.45	0.25	0.38
Sias - mean	0.322***		0.295***		0.317***	
Sias - SD	0.089		0.094		0.13	
Obs.	1975		1975		1975	

Note: All models are in time-series and are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method. Fixed-Effects include twenty days of the month dummies, five days of the week dummies, and 96 monthly dummies. “All events” consider as outages all 33 interruptions described on table 1, while “severe events” consider the 18 severe outages described in table 1. LSV measures are calculated for each stock as in Lakonishok et al. (1992) but using daily intervals. LSV measure is collapsed to time-series by equally weighting the measures on the cross-section of the stocks for each day. Sias herding measure is calculated as in Sias (2004) considering a daily interval. *D*ependent variables are in level, and we include the term $\rho * \overline{Y_{t-1:t-22}}$ in equation (1) to account for the persistency of the dependent variable. We also include other control variables that have been discussed in the literature as important to explain correlated trading.

Although the focus of this study is not on estimating the herding behavior of Brazilian investors, this is the first study that calculates such measure. The last row of each panel shows the unconditional mean for each herding measure and investor. For both measures retail investors appear to have larger herding behavior than institutions, which is in line with the international literature. LSV measure for retail investors in Brazil is the double estimated by Dorn, Huberman and Sengmueller (2008) for the United States. However, this comparison is not adequate. DHS study uses brokerage data between 98 and 2000, when investment activity was quite different from more recent years, our sample period. We also have a much larger mean for the Sias herding measure than for the LSV measure, this in line with what is observed in Sias (2004) investigation for funds using quarterly intervals to estimate herding.

9. Different reactions to outage at the investor level.

The evidence so far is based on an average result for a group of investors. Previous results showed that, on average, individual investors are less likely to trade on social media outage days. However, each investor may react differently, with some trading more on outage days. On this section we take a step on the direction of evaluating each retail investor response to the outages and exploring if investor with different responses have different trading characteristics.

To investigate the reaction at the investor level we will need to select a sample of very active investors (which traded almost every day), so we can have enough observation points to estimate their response to the social media outage days. For each year, we select retail investors which were active on at least 75% of the available trading days. This leaves us with a sample of 7116 individual investors. For each investor we run equation (1) time-series model and collect the estimated β and t-statistics, the coefficient of response to social media outages. Investors with absolute p-values above 0.1 are considered to have neutral response, a total of 6286 investors. Another 498 investors had negative and significant responses to the outages, while the other 332 had positive and significant responses. Table 10 reports some characteristics of these groups of investors. We acknowledge that the characteristics are already from a specific group of extremely active investors.

The first part of table 10 shows averages and median characteristics of each group. We first take the time-series average of the variable of interest for each investor in the group, and then take the average or median in the cross-section of the individuals. This sample of very active investors holds a much higher monetary value in stock holdings and is more diversified than the average retail investor in the sample. However, between the three groups there is little difference in the holdings value of the median investor of each group. The group of investors that reacted negatively to social media outages appears to be more diversified, with 1 extra stock on its portfolio (for the median investor).

Table 10 – Retail investor characteristics by reaction to social media outages.

	Reaction to Outage		
	Negative	**None**	**Positive**
Ave. Holdings (R$ Thousand)	839.0	1403.9	961.2
Median Holdings (R$ Thousand)	140.7	120.6	117.3
Ave. Number of Stocks	11.8	10.2	8.6
Median Number of Stocks	6.7	5.6	5.7
Median % of Portfolio in lottery stocks	2.8%	3.1%	3.4%
Annualized excess returns, EW on investors			
Portfolio (2012-2019)	2.2%	1.1%	1.7%
		(0.78)	(0.21)
Next day return from purchases in t	12.7%	-1.3%	13.0%
		(3.9)	(-0.06)
Next day return from purchases in t - On outages	-1.3%	-0.4%	0.9%
		(-1.14)	(-2.11)
Cumulative t+21 days returns from purchases in t	0.4%	-0.1%	0.1%
		(0.75)	(0.41)
Cumulative t+21 days returns from purchases in t - On outages	13.1%	1.9%	-1.9%
		(1.51)	(1.77)
Cumulative t+62 days returns from purchases in t	1.9%	0.6%	0.5%
		(4.05)	(3.02)
Cumulative t+62 days returns from purchases in t - On outages	8.9%	4.3%	2.1%
		(1.42)	(1.6)
Cumulative t+124 days returns from purchases in t	1.4%	0.4%	1.1%
		(4.16)	(0.68)
Cumulative t+124 days returns from purchases in t - On outages	5.0%	3.4%	1.2%
		(0.94)	(1.51)

Note: table 10 shows the characteristics and returns from three groups of retail investors formed from 7116 individual investors which were active at least 75% of the available trading days within a year. Investors are grouped based on their response to a social media outage estimated through equation (1). The first column comprehends 498 investors which had a negative (β) reaction with absolute *p*-value below 0.1. The second column comprehends 6286 investors which had either positive or negative (β) reaction, but with absolute *p*-values above 0.1. The last column comprehends 332 investors which had a positive (β) reaction with absolute *p*-value below 0.1. Average and median values for the groups are first obtained in the time-series for each investor in the group, and then in the cross-section of the individuals. For columns 2 and 3 we present the *t*-statistics for a test of means between column 2 and column 1, and column 3 and column 1 returns, respectively. This test of means assumes daily returns are independent.

The second part of table 10 shows annualized excess returns under different circumstances for each group, with column 2 and 3 being compared with column 1 through a test of means. Investors with negative responses had better overall returns for the sample period, but this was not statistically different from the two other groups. In general, investors with negative responses appear to have better returns on their purchases for different horizons, one day, month, quarter, and semester. More interesting, although investors with negative responses appear to do better on regular days, their one-day return on outage days purchases is statistically worse than investors with positive response to social media outage. From our initial hypothesis, investors from the first column would use social media as an information and communication tool for trading, while the last column would be distracted by it. The result could them come from investors from the first column performing worse because they have less information available on outage days, or because investors from the last columns, are doing relatively better when they are less distracted by social media.

10. Conclusions.

This study uses investor-day level data to investigate the effect of social media interruptions on investors trading activity. Our main result is that investors reduce their activity on days when social media platforms stop working. This result is observed for the money traded and for the number of investors trading stocks, happening for retail investors and domestic institutional investors, but not for foreign investors. This result is robust to different specifications and measures of trading activity, with "placebo" outages (occurring outside the market trading hours) presenting no results. As investors leave the market on outage days, market liquidity worsens, evidence that supports social media as improving market liquidity and efficiency in general.

Investors shunning away from the markets when they cannot access social media platforms supports the hypothesis that social media are vehicles of information about the market with the potential to reduce informational frictions and improve market liquidity, more so than to distract investors (the competing hypothesis). Other evidence also corroborates the informational hypothesis. First, results are stronger for smaller and illiquid firms, which are more likely to have social media as an important source for information dissemination. Second, results are stronger for stocks which had higher past returns, and most likely were grabbing attention, so when social media is out, the attention to these stocks is disproportionately affected. Third, results are driven by retail investors and funds with smaller accounts, which are the ones that do not have access to specialized platforms and research houses, which delivers market news, data, and analysis, but rely on payable subscriptions.

Another important evidence suggests social media spreads company information to investors. Analysis shows the existence of a one-day and one-week reversal in returns, with this reversal weakened when accompanied by a firm announcement. However, these results shift when social media is down. First, the reversal is smaller on social media outage days. Second, social media outages affect the return responses to firms' announcements. On days without social media, announcements seem to have no impact on the return autocorrelation. We believe this is because a channel (social media) which helps to spread the announcement information is closed.

Nonetheless, not all results supported this hypothesis. If social media platforms are a source of news dispersion, we expect that when days that this channel is blocked, investors would reduce their correlated trading. However, herding behavior is not affected by social media outages. Also, another puzzling result which its interpretation is not clear is the similar magnitudes and effects for retail investors and domestic institutional investors. Individual investor literature shows institutional investors as more sophisticated, and this premise suggests that they should be less affected by social media outages, gathering their information from other sources.

Overall, results support the idea of social media being used by investors as a source of information and communication, increasing market activity and improving market liquidity.

Bibliography

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31-56.

Ammann, M., & Schaub, N. (2021). Do individual investors trade on investment-related internet postings?. Management science, 67(9), 5679-5702.

An, L., Lou, D., & Shi, D. (2022). Wealth redistribution in bubbles and crashes. Journal of Monetary Economics.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. The Journal of finance, 59(3), 1259-1294.

Anderson, A. (2013). Trading and under-diversification. *Review of Finance*, *17*(5), 1699-1741.

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. The Journal of Finance, 68(3), 929-985.

Bach, L., Calvet, L. E., & Sodini, P. (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, *110*(9), 2703-47.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2021). Attention induced trading and returns: Evidence from robinhood users. *Journal of Finance, forthcoming*.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2009). Just how much do individual investors lose by trading?. *The Review of Financial Studies*, *22*(2), 609-632.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2014). The cross-section of speculator skill: Evidence from day trading. *Journal of Financial Markets*, *18*, 1-24.

Barber, B. M., Lee, Y. T., Liu, Y. J., Odean, T., & Zhang, K. (2020). Learning, fast or slow. *The Review of Asset Pricing Studies*, *10*(1), 61-93.

Barber, B. M., Lin, S., & Odean, T. (2019). Mediating Investor Attention. Working Paper, University of California, Berkeley.

Barber, B. M., Lin, S., & Odean, T. (2021). Resolving a paradox: Retail trades positively predict returns but are not profitable. *Available at SSRN*.

Barber, B. M., & Loeffler, D. (1993). The "dartboard" column: Second-hand information and price pressure. *Journal of Financial and Quantitative Analysis*, *28*(2), 273-284.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, *55*(2), 773-806.

Barber M., & Odean, T. (2002). Online investors: do the slow die first?. The Review of financial studies, 15(2), 455-488.

Barber, B. M., & Odean, T. (2008). 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), 785-818.

Barber, B. M., & Odean, T. (2013). The behavior of individual investors. In *Handbook of the Economics of Finance* (Vol. 2, pp. 1533-1570). Elsevier.

Barber, B. M., Odean, T., & Zhu, N. (2008). Do retail trades move markets?. *The Review of Financial Studies*, *22*(1), 151-186.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2018). Extrapolation and bubbles. *Journal of Financial Economics*, *129*(2), 203-227.

Baron, M., Brogaard, J., Hagströmer, B., & Kirilenko, A. (2019). Risk and return in high-frequency trading. *Journal of Financial and Quantitative Analysis*, *54*(3), 993-1024.

Barrot, J. N., Kaniel, R., & Sraer, D. (2016). Are retail traders compensated for providing liquidity?. *Journal of Financial Economics*, *120*(1), 146-168.

Bauer, R., Cosemans, M., & Eichholtz, P. (2009). Option trading and individual investor performance. *Journal of Banking & Finance*, *33*(4), 731-746.

Bernstein, J., & Bernstein, J. (1995). *The Compleat Day-trader: Trading Systems, Strategies, Timing Indicators, and Analytical Methods*. McGraw-Hill.

Blankespoor, E., Miller, G. S., & White, H. D. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *The accounting review*, *89*(1), 79-112.

Bianchi, M. (2018). Financial literacy and portfolio dynamics. *The Journal of Finance*, *73*(2), 831-859.

Birru, J., Chague, F., De-Losso, R., & Giovannetti, B. (2021). Attention and biases: evidence from tax-inattentive investors. *Fisher College of Business Working Paper*, (2019-03), 022.

Bodnaruk, A. (2009). Proximity always matters: Local bias when the set of local companies changes. *Review of finance*, *13*(4), 629-656.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Tracking retail investor activity. The Journal of Finance, 76(5), 2249-2305.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, *127*(3), 1243-1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803-843.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2017). Memory, attention, and choice. *The Quarterly journal of economics*.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022) Salience. *Annual Review of Economics*, *14.*, *forthcoming*.

Brown, N. C., Elliott, W. B., Wermers, R., & White, R. M. (2022). News or noise: mobile internet technology and stock market activity. Available at SSRN 3585128.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, *115*(5), 707-747.

Campbell, J. Y., Ramadorai, T., & Ranish, B. (2019). Do the rich get richer in the stock market? Evidence from India. *American Economic Review: Insights*, *1*(2), 225-40.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets*, *16*(4), 680-711.

Chague, F. D., Bueno, R. D. L. D. S., & Giovannetti, B. C. (2018). Individuals neglect the informational role of prices: Evidence from the stock market.

Chague, F., De-Losso, R., & Giovannetti, B. (2020). Day trading for a living?. *Available at SSRN 3423101*.

Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, *27*(5), 1367-1403.

Chen, G., Kim, K. A., Nofsinger, J. R., & Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of behavioral decision making*, *20*(4), 425-451.

Choe, H., Kho, B. C., & Stulz, R. M. (1999). Do foreign investors destabilize stock markets? The Korean experience in 1997. *Journal of Financial economics*, *54*(2), 227-264.

Choe, H., & Eom, Y. (2009). The disposition effect and investment performance in the futures market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *29*(6), 496-522.

Choi, J. J., Haisley, E., Kurkoski, J., & Massey, C. (2012). *Small cues change savings choices* (No. w17843). National Bureau of Economic Research.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. The journal of finance, 56(2), 501-530.

Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of financial Economics*, *87*(2), 249-268.

Cookson, J. A., Engelberg, J., & Mullins, W. (2022). Echo chambers. Available at SSRN 3603107.

Coval, J. D., Hirshleifer, D. A., & Shumway, T. (2005). Can individual investors beat the market?.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The journal of finance*, *66*(5), 1461-1499.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. The journal of finance, 64(2), 709-749.

Dichev, I. D. (2007). What are stock investors' actual historical returns? Evidence from dollar-weighted returns. *American Economic Review*, *97*(1), 386-401.

Ding, R., Zhou, H., & Li, Y. (2020). Social media, financial reporting opacity, and return comovement: Evidence from Seeking Alpha. Journal of Financial Markets, 50, 100511.

Dorn, D., Huberman, G., & Sengmueller, P. (2008). Correlated trading and returns. *The Journal of Finance*, *63*(2), 885-920.

Døskeland, T. M., & Hvide, H. K. (2011). Do individual investors have asymmetric information based on work experience?. *The Journal of Finance*, *66*(3), 1011-1041.

Eaton, G. W., Green, T. C., Roseman, B., & Wu, Y. (2021). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Available at SSRN 3776874*.

Engelberg, J., Sasseville, C., & Williams, J. (2012). Market madness? The case of mad money. *Management Science*, *58*(2), 351-364.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. The Journal of Finance, 64(5), 2023-2052.

Frydman, C., & Wang, B. (2020). The impact of salience on investor behavior: Evidence from a natural experiment. *The Journal of Finance*, *75*(1), 229-276.

Gerken, W. C., & Painter, M. (2019). The Value of Differing Points of View: Evidence from Financial Analysts' Geographic Diversity. *Available at SSRN 3479352*.

Goetzmann, W. N., & Kumar, A. (2008). Equity portfolio diversification. *Review of Finance*, *12*(3), 433-463.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland's unique data set. *Journal of financial economics*, *55*(1), 43-67.

Grinblatt, M., & Keloharju, M. (2001). How distance, language, and culture influence stockholdings and trades. *The Journal of Finance*, *56*(3), 1053-1073.

Grinblatt, M., & Keloharju, M. (2001). What makes investors trade?. *The journal of Finance*, *56*(2), 589-616.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. (2011). IQ and stock market participation. *The Journal of Finance*, *66*(6), 2121-2164.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics*, *104*(2), 339-362.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American economic review*, 1088-1105.

Guiso, L., & Sodini, P. (2013). Household finance: An emerging field. In *Handbook of the Economics of Finance* (Vol. 2, pp. 1397-1532). Elsevier.

Haliassos, M., & Bertaut, C. C. (1995). Why do so few hold stocks?. *the economic Journal*, *105*(432), 1110-1129.

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. The Review of Financial Studies, 6(3), 473-506.

Hartzmark, S. 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.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. The Journal of Finance, 64(5), 2289-2325.

Hu, D., Jones, C. M., Zhang, V., & Zhang, X. (2021). The rise of reddit: How social media affects retail investors and short-sellers' roles in price discovery. *Available at SSRN 3807655*.

Huberman, G. (2001). Familiarity breeds investment. *The Review of Financial Studies*, *14*(3), 659-680.

Huberman, G., & Sengmueller, P. (2008). Correlated trading and returns. *The Journal of Finance*, *63*(2), 885-920.

Ivković, Z., Sialm, C., & Weisbenner, S. (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis*, *43*(3), 613-655.

Ivković, Z., & Weisbenner, S. (2005). Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance*, *60*(1), 267-306.

Jones, C. M., Shi, D., Zhang, X., & Zhang, X. (2021). Understanding Retail Investors: Evidence from China. *SSRN working paper*.

Jordan, D. J., & Diltz, J. D. (2003). The profitability of day traders. *Financial Analysts Journal*, *59*(6), 85-94.

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The Journal of finance*, *63*(1), 273-310.

Kaniel, R., & Parham, R. (2017). WSJ Category Kings–The impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics*, *123*(2), 337-356.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? Insights from retail orders and stock returns. *The Journal of Finance*, *68*(3), 1229-1265.

Keloharju, M., Knüpfer, S., & Linnainmaa, J. (2012). Do investors buy what they know? Product market choices and investment decisions. *The Review of Financial Studies*, *25*(10), 2921-2958.

Kumar, A. (2009). Who gambles in the stock market?. *The Journal of Finance, 64*(4), 1889-1933.

Kuo, W. Y., & Lin, T. C. (2013). Overconfident individual day traders: Evidence from the Taiwan futures market. *Journal of banking & Finance*, *37*(9), 3548-3561.

Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, 1315-1335.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of financial economics*, *32*(1), 23-43.

Liang, B. (1999). Price pressure: Evidence from the “dartboard” column. *the Journal of Business*, *72*(1), 119-134.

Linnainmaa, J. T. (2003). The anatomy of day traders. *Available at SSRN 472182*.

Li, X., Geng, Z., Subrahmanyam, A., & Yu, H. (2017). Do wealthy investors have an informational advantage? Evidence based on account classifications of individual investors. *Journal of Empirical Finance*, *44*, 1-18.

Luo, C., Ravina, E., Sammon, M., & Viceira, L. M. (2020). Retail investors’ contrarian behavior around news and the momentum effect. *Available at SSRN 3544949*.

Massa, M., & Simonov, A. (2006). Hedging, familiarity and portfolio choice. *The Review of Financial Studies*, *19*(2), 633-685.

Meghir, C., Narita, R., & Robin, J. M. (2015). Wages and informality in developing countries. *American Economic Review, 105*(4), 1509-46.

Menezes-Filho, N. A., Muendler, M. A., & Ramey, G. (2008). The structure of worker compensation in Brazil, with a comparison to France and the United States. *The Review of Economics and Statistics*, *90*(2), 324-346.

Mohr, J. (2021). Pulling the Plug: Retail Traders and Social Media. *Available at SSRN 3917950*.

Odean, T. (1999). Do investors trade too much?. *American economic review, 89*(5), 1279-1298.

Ozik, G., Sadka, R., & Shen, S. (2021). Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. *Journal of Financial and Quantitative Analysis, 56*(7), 2356-2388.

Pedersen, L. H. (2022). Game on: Social networks and markets. Journal of Financial Economics.

Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *the Journal of Finance*, *69*(5), 2007-2043.

Peress, J., & Schmidt, D. (2020). Glued to the TV: Distracted noise traders and stock market liquidity. *The Journal of Finance*, *75*(2), 1083-1133.

Piketty, T. (2018). Capital in the twenty-first century. In *Capital in the twenty-first century*. Harvard University Press.

Ryu, D. (2012). The profitability of day trading: An empirical study using high-quality data. *Investment Analysts Journal*, *41*(75), 43-54.

Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, *131*(2), 519-578.

Seasholes, M. S., & Zhu, N. (2010). Individual investors and local bias. *The Journal of Finance*, *65*(5), 1987-2010.

Seru, A., Shumway, T., & Stoffman, N. (2010). Learning by trading. *The Review of Financial Studies*, *23*(2), 705-739.

Sias, R. W. (2004). Institutional herding. *The Review of Financial Studies*, *17*(1), 165-206.

Tetlock, P. C. (2014). Information transmission in finance. *Annu. Rev. Financ. Econ.*, *6*(1), 365-384.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, *108*(8), 2015-47.

Van Nieuwerburgh, S., & Veldkamp, L. (2010). Information acquisition and under-diversification. *The Review of Economic Studies*, *77*(2), 779-805.

Vissing-Jørgensen, A., & Attanasio, O. P. (2003). Stock-market participation, intertemporal substitution, and risk-aversion. *American Economic Review*, *93*(2), 383-391.

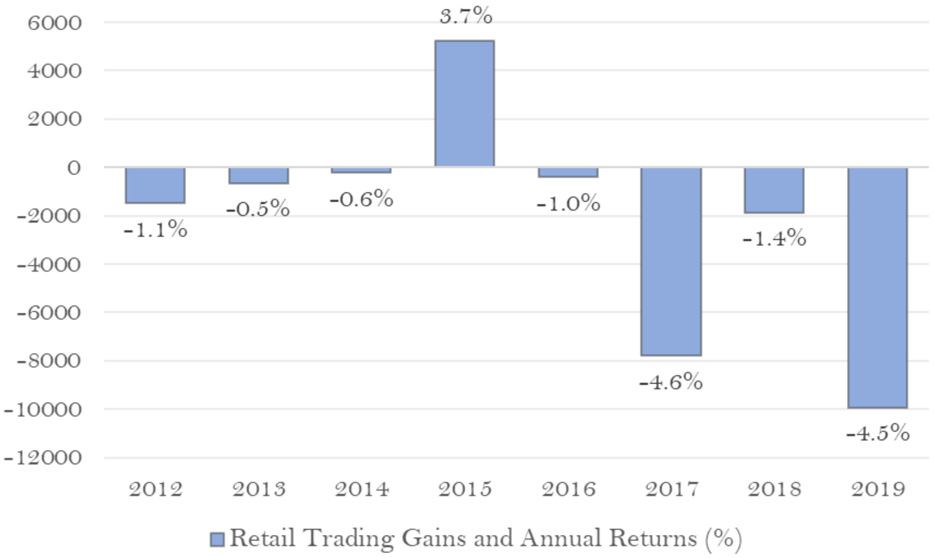
Wang, B. (2017). Ranking and salience. *Available at SSRN 2922350*.

Xu, Y., Xuan, Y., & Zheng, G. (2021). Internet searching and stock price crash risk: Evidence from a quasi-natural experiment. *Journal of Financial Economics*, *141*(1), 255-275.

Zhang, W., Wang, P., & Li, Y. (2021). Do messages on online stock forums spur firm productivity?. *Pacific-Basin Finance Journal*, *68*, 101609.

APPENDIX A – Chapter 3

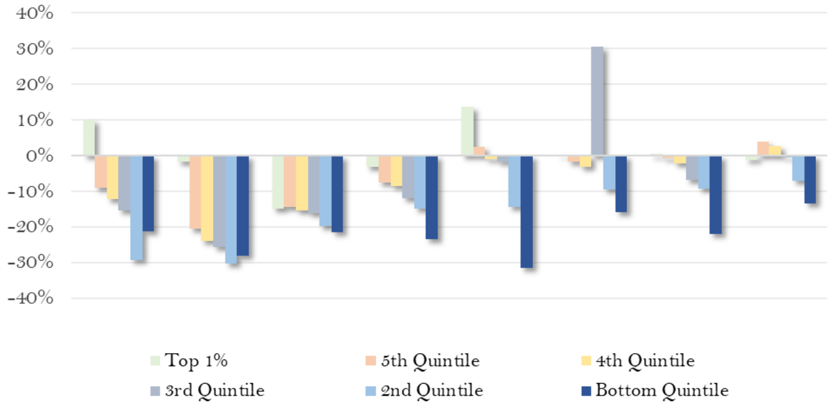
Figure A1 – Year over year result of the cumulative retail trading gains and returns for an eight-year period (2012-2019).



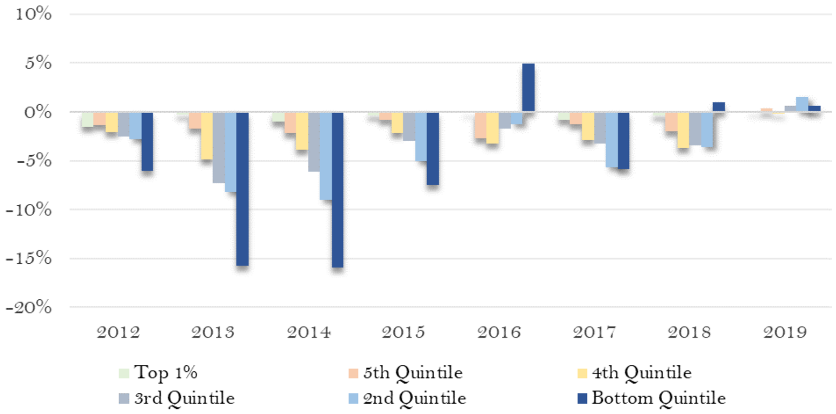
Note: Monetary cumulative trading gains are calculated from all trades between 2012 to 2019 considering excess returns relative to risk-free rate. Figure A1 breaks down the dynamics of the cumulative gain from 2012 to 2019, showing the contribution of each year to the total cumulative measure. Trading gains combine daily and intraday stock market gains, plus derivative gains. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss for each year. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). The Figure shows that the last year has the largest contribution in monetary terms (R$ 10 billion), but not in percentage (-4.5%), with the worst year for retail cumulative trading return being 2017.

Figure A2 – Annualized returns from monetary cumulative trading gains by year (2012-2019) and group of investors - discounting returns from market timing.

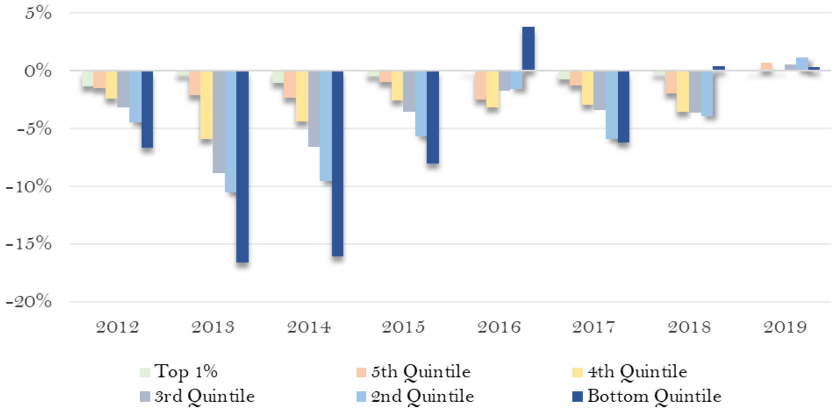
A – New Accounts



B – Old Accounts



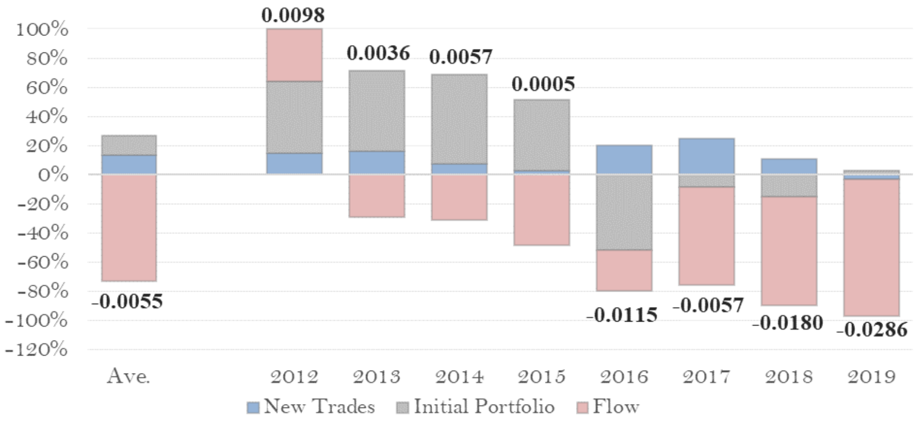
C – All Accounts



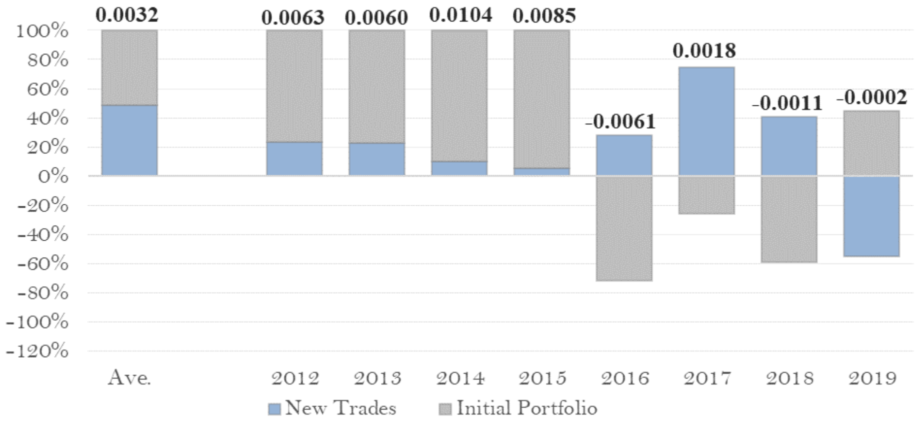
Note: This figure reports annualized returns calculated from monetary cumulative trading results when breaking down by each year, wealth group, and new vs. old investors. New investors have accounts opened within the same year of analysis. Monetary cumulative trading gains are calculated from all trades within a year considering excess returns. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. In the beginning of each year investors are regrouped into wealth categories. Existing accounts are ranked by their total direct holdings value in the end of previous year, and accounts opened within the same year are ranked by their maximum wealth in a three-year period - for accounts opened in 2017, 2018 and 2019 the classification of new accounts is done by quarterly cohorts, given that there is less than three-years to estimate maximum wealth. Return contribution from market timing is calculated from the contribution of the respective marginal daily gain and loss when considering Eq. (5) to calculate market timing.

Figure A3 - Yearly decomposition of Gini Index variation through counterfactuals (2012-2019) – considering net flows.

A – Initial Portfolio, New Trading, and Net Flows.



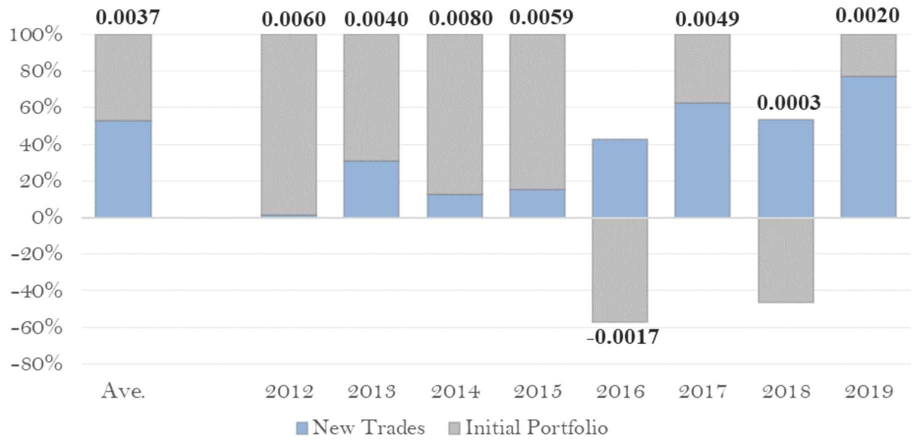
B – Initial Portfolio vs. New Trading.



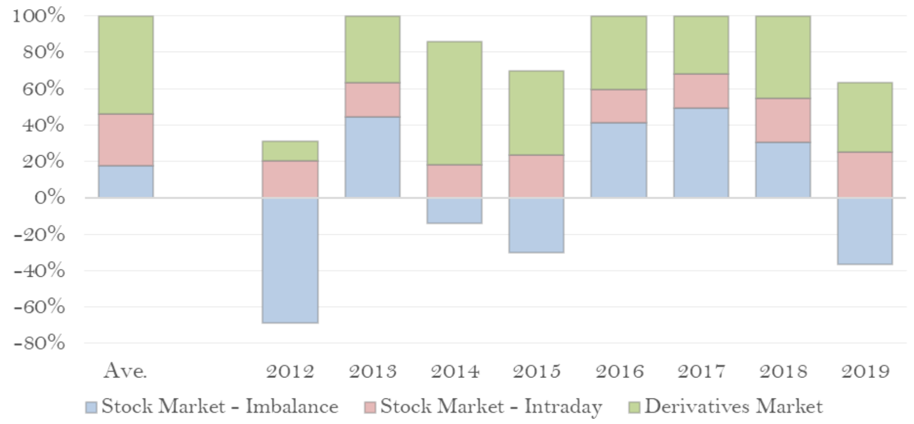
Note: Figure A3 reports the variation in the Gini Index from different counterfactuals described in Table A10, when performing the counterfactual exercise for the beginning of each year, until the end of the same year. The first column reports the time-series average of the year-by-year results. Figure (A) and (C) consider the effect of the net flows into the Gini Index variations. Figure (A) assumes that an investor with no direct holdings in the equity market has zero wealth for the Gini Index calculations. Figure (C) assumes that an investor with no direct holdings in the equity market is a missing observation and is not considered in the Gini Index calculation. For Figure (C), Gini Index in the beginning and end of the year are calculated with different number of investors. Figure (B) reports the results of figure (A) disregarding net flows.

Figure A4 - Yearly decomposition of Gini Index variation through counterfactuals (2012-2019) – old accounts.

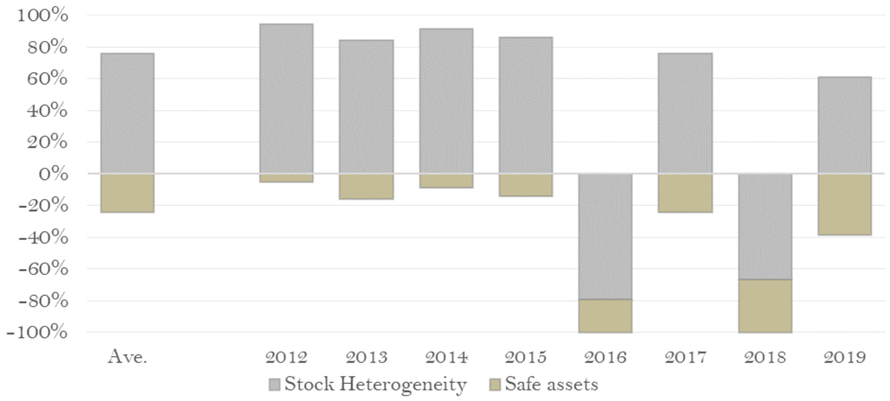
A – Initial Portfolio vs. New Trading



B – New Trading Decomposition



C – Initial Portfolio Decomposition



Note: Old accounts had positive direct holdings at the end of each previous year of the counterfactual exercise. Gini Index variation is calculated from different counterfactuals described in Table 12, calculating in the beginning of each year, until the end of the same year. The first column reports the time-series average of the year-by-year results. New trading and initial portfolio are further decomposed in the second (B) and third (C) figure.

Table A1 – Cumulative trading gains and returns within retail investors over an eight-year period (2012-2019), with new accounts ranked by maximum positive net flow.

		Cumulative trading gains / return											
Investors / Market		***Top* 1%**		**5th Quintile**		**4th Quintile**		**3th Quintile**		**2nd Quintile**		***Bottom* Quintile**	
		Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)	Millions (R$)	Annual (%)
Panel A													
All investors	All markets	-13972	-1.6%	-2781	-0.9%	-175	-0.5%	-46	-0.5%	-93	-0.9%	-103	-2.0%
Trade only in stock market	Stock Market	-7187	-1.6%	-2060	-1.0%	-353	-0.8%	-108	-0.6%	-200	-1.5%	-94	-2.3%
Trade in derivatives market	All markets	-6785	-1.7%	-720	-0.8%	178	0.0%	62	-0.4%	107	0.4%	-9	-1.4%
	Derivatives	-1223	-0.4%	-1334	-1.0%	-430	-1.8%	-255	-2.1%	-130	-2.1%	-88	-3.2%
	Stock market	-5562	-1.3%	614	0.3%	608	1.9%	317	1.8%	238	2.6%	79	1.9%
Panel B													
% of investors w/ negative trading result		39%		35%		38%		42%		47%		52%	
Ave. gain | Positive result (Thousand R$)		1376		76		17.7		9.0		5.7		2.4	
Ave. gain | Negative result (Thousand R$)		-4209		-166		-30.7		-12.9		-7.0		-2.7	
Ave. Wealth 2012-2019 (Millions R$)		98861		48887		9578		4527		2415		860	
# Investors		17316		328823		346024		345733		346351		346405	

Note: Investors’ wealth is defined by her wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum positive net flow within three years for new accounts (opened between 2012-2019). Monetary cumulative trading gains are calculated from all trades between 2012 to 2019 considering excess returns. Stock market gains combine daily and intraday gains measures, together with derivative gains they constitute the “All markets” cumulative trading gain. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative result are investors with the cumulative trading measure below zero at the end of 2019.

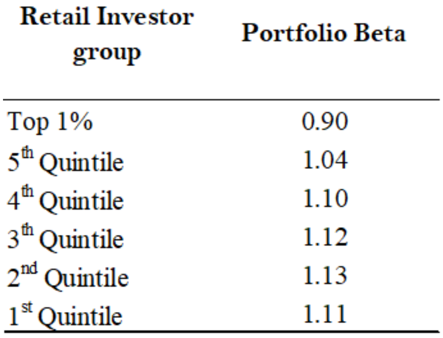
Table A2 – BLLO measure outcomes for monetary results, annualized returns, and proportion of investors with negative trading result.

Mean daily monetary profit from trade - accumulated for eight years (2012-2019)

		$Ret_{t:t+5}$	$Ret_{t:t+21}$	$Ret_{t:t+62}$	$Ret_{t:t+124}$	$Ret_{t:t+248}$
All	Millions (R$)	-12435	-14557	-12687	-10210	-12272
	Annual (%)	-5.5%	-5.8%	-1.4%	-0.7%	-1.6%
	% of investors w/ negative trading result	52.8%	49.3%	42.6%	33.7%	30.5%
***Top* 1%**	Millions (R$)	-5493	-6071	-5846	-4781	-5110
	Annual (%)	-2.2%	-3.3%	-0.5%	-0.2%	-0.8%
	% of investors w/ negative trading result	53.9%	48.6%	40.9%	34.6%	30.0%
5th Quintile	Millions (R$)	-5109	-6172	-4916	-3881	-4984
	Annual (%)	-5.5%	-5.7%	-1.5%	-0.7%	-1.5%
	% of investors w/ negative trading result	53.1%	47.8%	40.5%	33.3%	30.1%
4th Quintile	Millions (R$)	-1040	-1300	-1105	-864	-1165
	Annual (%)	-8.9%	-7.1%	-2.3%	-1.2%	-1.9%
	% of investors w/ negative trading result	52.8%	48.7%	42.1%	33.7%	30.3%
3th Quintile	Millions (R$)	-460	-595	-479	-368	-551
	Annual (%)	-8.8%	-7.0%	-2.1%	-0.9%	-1.7%
	% of investors w/ negative trading result	52.9%	49.7%	43.6%	34.8%	31.0%
2nd Quintile	Millions (R$)	-227	-291	-219	-205	-304
	Annual (%)	-8.0%	-7.2%	-2.2%	-1.4%	-2.0%
	% of investors w/ negative trading result	52.9%	50.3%	44.1%	34.8%	31.3%
***Bottom* Quintile**	Millions (R$)	-107	-129	-122	-110	-159
	Annual (%)	-5.5%	-5.8%	-3.2%	-1.9%	-2.4%
	% of investors w/ negative trading result	52.4%	50.1%	42.6%	31.8%	29.6%

Note: Table A2 shows the cumulative monetary result, annualized returns, and proportion of investors with negative trading result when calculating daily profits using the Barber, Lee, Liu and Odean (2009) measure with different fixed horizons for holding the portfolio. First, we construct a buy and sell portfolio for each investor group, one that mimics the net daily purchases and one that mimics the net daily sales, the difference between them gives the total monetary outcome. Shares are included in the portfolio for a fixed horizon, 5, 21, 62, 124, and 248 days. Portfolio profits are calculated by comparing as alternative a risk-free bond.

Table A3 – Portfolio betas with relation to the equity market for distinct groups of wealth.



Retail Investor group	Portfolio Beta
Top 1%	0.90
5th Quintile	1.04
4th Quintile	1.10
3th Quintile	1.12
2nd Quintile	1.13
1st Quintile	1.11

Note: Table A3 reports estimated portfolio betas for different wealth groups in relation to the equity market, proxied by the returns of the Ibovespa Index. For each wealth category we calculate their daily return based on their portfolio's holdings of the end of the previous day and make a simple time-series OLS regression of the obtained return on the Ibovespa Index return for the eight-year sample interval. Table A3 reports the beta coefficients of these regressions.

Table A4 – Return decomposition by groups of investor wealth – yearly averages.

Retail Investor group		**Return (Year Average)**	**Return Decomposition**			
			Stock Market			**Derivatives Market**
			Stock Selection	**Market Timing**	**Intraday**	
Top 1%	Trade only in stock market	-0.19	172%	-104%	32%	0%
5th Quintile		-0.49	103%	-84%	82%	0%
4th Quintile		-1.33	92%	-51%	59%	0%
3th Quintile		-1.22	121%	-87%	66%	0%
2nd Quintile		-1.65	98%	-69%	71%	0%
Bottom Quintile		-2.70	47%	0%	53%	0%
Top 1%	Trade in derivatives market	-0.86	19%	-15%	15%	81%
5th Quintile		-2.38	20%	-15%	25%	70%
4th Quintile		-5.25	21%	-10%	16%	72%
3th Quintile		-7.50	17%	-9%	9%	83%
2nd Quintile		-10.32	11%	-6%	8%	87%
Bottom Quintile		-10.79	3%	-18%	11%	104%

Note: Table A4 reports the time-series averages of annualized returns calculated from monetary cumulative trading results and the time-series averages of return decomposition when breaking down by each year and wealth group. Monetary cumulative trading gains are calculated from all trades within a year considering excess returns. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Return contribution of each category is calculated from the contribution of the respective marginal daily gain and loss of each division, while stock selection is obtained as the residual daily stock market gain after considering market timing as in Eq (5). In the beginning of each year investors are regrouped into wealth categories. Existing accounts are ranked by their total direct holdings value in the end of previous year, and accounts opened within the same year are ranked by their maximum wealth in a three-year period - for accounts opened in 2017, 2018 and 2019 the classification of new accounts is done by quarterly cohorts, given that there is less than three-years to estimate maximum wealth. Investors who traded at least one derivative contract are separated and showed in the “trade in derivatives market” group. Overall, retail investors have positive market timing. For investors who do not trade derivatives contracts, poor stock selection ability is the main contributor to their losses. The exception are the investors in the bottom quintile, which are the only group that do not have a positive market timing ability and have higher losses on intraday execution than on stock selection ability. Table A4 also reveals that most of the loss taken by retailers who trade derivatives contracts comes from trading these contracts. Unlike Table 3, investors who trade derivatives do significantly worse when we consider shorter intervals to estimate the gains and losses from trading.

Table A5 – Annual returns and proportion of investors with negative monetary result by wealth groups double sorted into terciles of turnover and number of stocks in portfolio.

Panel A - Group returns (Year %)

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Low Activity	-1.4	-1.6	-0.6	-2.1	-1.1	1.3	-3.7	-0.3	2.2
Medium Activity	-0.4	-4.2	0.2	-4.8	-1.3	1.2	-5.8	0.3	1.3
High Activity	-12.3	-6.1	-0.9	-10.0	-3.2	-0.1	-13.1	-2.6	1.4

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Low Activity	-5.7	-1.1	2.5	-13.2	-1.4	2.3	-14.0	-1.2	3.0
Medium Activity	-7.9	0.0	2.5	-14.5	-0.1	1.7	-18.8	-1.1	2.2
High Activity	-18.3	-5.6	-0.5	-11.9	-7.8	0.3	-35.5	-12.2	-3.5

Panel B - Percentage of investors within the group with negative return

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.	Low Diver.	Medium Diver.	High Diver.
Low Activity	54%	51%	41%	54%	44%	26%	58%	43%	26%
Medium Activity	32%	20%	12%	39%	21%	11%	42%	26%	16%
High Activity	53%	38%	24%	51%	33%	22%	52%	39%	28%

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Low Activity	63%	48%	30%	68%	51%	35%	73%	51%	44%
Medium Activity	48%	31%	22%	54%	39%	29%	52%	50%	40%
High Activity	54%	46%	35%	56%	53%	42%	54%	58%	50%

Panel C - Observations, Number of investors within group

***Top* 1%**			***Top* Quintile**			**Quintile 4**		
2131	1692	1949	33458	32220	43919	37463	31736	46131
1482	1906	2383	30105	37515	41976	34396	38697	42236
2159	2181	1433	46034	39986	23610	48416	41358	25591

Quintile 3			**Quintile 2**			***Bottom* Quintile**		
37826	34356	43051	45446	31066	38927	68286	9578	37593
34237	38128	42867	33522	37959	43957	57364	14193	43899
43623	42308	29337	36516	46691	32267	64588	16911	33993

Note: Investors are first classified by their wealth and then double-sorted into levels of diversification and activity. Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors, within wealth group, are divided by the average number of stocks they had throughout the sample into three groups (terciles) of diversification. Investors, within wealth group, are also divided by their average monthly turnover (%) into three groups (terciles) of activity. Turnover is measured by the monetary volume of buy and sell orders divided by two times the average monthly holding. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

Table A6 – Investor's classification by its purchase pattern, difference in proportion between momentum and contrarian classification for each group.

			%(Inverstors classified as Contrarians) - %(Inverstors classified as Momentum)					
Past return Interval		**Institutions**	**Retail**					
			***Top* 1%**	***Top* Quintile**	**Quintile 4**	**Quintile 3**	**Quintile 2**	***Bottom* Quintile**
Panel A								
Cross-Section	(t-1):t	3%	-8%	-5%	-3%	-1%	0%	2%
	(t-5):t	1%	-13%	-8%	-6%	-4%	-2%	0%
	(t-21):t	-1%	-16%	-9%	-6%	-4%	-4%	-5%
	(t-62):t	2%	-10%	-7%	-6%	-7%	-8%	-11%
	(t-125):t	7%	-5%	-3%	-3%	-5%	-8%	-15%
	(t-251):t	18%	12%	5%	-2%	-8%	-14%	-25%
	(t-754):t	35%	46%	32%	23%	16%	8%	-7%
Panel B								
Time-Series, relative to zero	(t-1):t	5%	2%	6%	8%	10%	11%	12%
	(t-5):t	7%	8%	13%	14%	15%	15%	14%
	(t-21):t	14%	16%	21%	21%	20%	17%	11%
	(t-62):t	29%	33%	36%	35%	33%	28%	19%
	(t-125):t	35%	48%	48%	46%	42%	36%	22%
	(t-251):t	35%	56%	52%	48%	43%	36%	21%
	(t-754):t	61%	68%	64%	59%	52%	42%	24%
Panel C								
Time-Series, relative to market returns	(t-1):t	-74%	-74%	-68%	-62%	-55%	-49%	-38%
	(t-5):t	1%	-11%	-7%	-5%	-3%	-1%	1%
	(t-21):t	0%	-15%	-8%	-4%	-3%	-3%	-4%
	(t-62):t	3%	-7%	-3%	-1%	-1%	-3%	-6%
	(t-125):t	10%	6%	8%	7%	5%	1%	-7%
	(t-251):t	22%	22%	17%	11%	5%	0%	-10%
	(t-754):t	45%	46%	30%	19%	12%	4%	-9%

Note: Table A6 reports the difference between the proportion of investors classified as momentum investors and the proportion of investors classified as contrarian investors for different wealth groups and definitions of loser or winner portfolios. Investors are characterized by their purchases. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Panel A classifies loser stocks in the cross-section, when their past 12-month return is below the median. Panel B classify in the time-series, with stocks with past 12-month return below zero being considered losers. Panel C classifies stocks as losers in the time-series, but relative to the market, with stocks with past 12-month return below the past 12-month return of the Ibovespa Index being classified as losers.

Table A7 - Difference between the proportion of momentum and contrarian investors with monetary loss for distinct groups of wealth and investment style.

Past return interval		**%(Momentum investors w/ loss) - %(Contrarian investors w/ loss)**					
		***Top* 1%**	***Top* Quintile**	**Quintile 4**	**Quintile 3**	**Quintile 2**	***Bottom* Quintile**
Panel A							
Cross-Section	(t-1):t	-1%	-3%	0%	1%	2%	2%
	(t-5):t	0%	-5%	-3%	-1%	0%	1%
	(t-21):t	-1%	-6%	-5%	-4%	-3%	-4%
	(t-62):t	0%	-6%	-5%	-4%	-4%	-4%
	(t-125):t	2%	-4%	-4%	-4%	-5%	-6%
	(t-251):t	-4%	-8%	-6%	-6%	-6%	-6%
	(t-754):t	-26%	-28%	-24%	-21%	-21%	-21%
Panel B							
Time-Series, relative to zero	(t-1):t	-6%	-5%	-2%	0%	2%	2%
	(t-5):t	-13%	-13%	-9%	-6%	-3%	-1%
	(t-21):t	-13%	-15%	-12%	-8%	-7%	-5%
	(t-62):t	-15%	-20%	-18%	-14%	-13%	-11%
	(t-125):t	-21%	-27%	-24%	-21%	-20%	-18%
	(t-251):t	-29%	-35%	-32%	-28%	-26%	-25%
	(t-754):t	-29%	-37%	-34%	-30%	-28%	-27%
Panel C							
Time-Series, relative to market returns	(t-1):t	2%	4%	6%	8%	8%	7%
	(t-5):t	-4%	-7%	-5%	-3%	-2%	-1%
	(t-21):t	-1%	-6%	-5%	-4%	-3%	-3%
	(t-62):t	-1%	-7%	-6%	-5%	-6%	-6%
	(t-125):t	-4%	-9%	-9%	-9%	-10%	-10%
	(t-251):t	-8%	-11%	-10%	-11%	-12%	-14%
	(t-754):t	-15%	-14%	-13%	-14%	-16%	-19%

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Panel A classifies loser stocks in the cross-section, when their past 12-month return is below the median. Panel B classify in the time-series, with stocks with past 12-month return below zero being considered losers. Panel C classifies stocks as losers in the time-series, but relative to the market, with stocks with past 12-month return below the past 12-month return of the Ibovespa Index being classified as losers. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations of each wealth group is reported on Table 8. Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

Table A8 – Annual returns and proportions of investors with negative monetary result by the intersection of wealth groups, terciles of monthly turnover (%), and contrarian investment style - using cross-section classification with past 12-month returns.

Panel A - Group returns (Year %)

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity
Contrarian	-0.4	-1.3	-12.2	0.4	-3.1	-10.5	0.7	-1.6	-8.1
Neutral	-1.1	-0.5	-5.1	1.4	0.6	-2.6	2.8	1.5	-1.8
Momentum	-1.6	-2.8	-8.7	0.4	-0.1	-0.1	1.2	0.5	-0.5

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	0.9	-1.7	-10.4	1.3	-1.9	-9.4	0.3	-1.6	-11.8
Neutral	2.9	1.6	-3.1	3.0	1.7	-1.2	3.8	1.0	-3.2
Momentum	2.0	1.4	-2.2	0.6	2.1	-1.7	3.3	0.6	-2.0

Panel B - Percentage of investors with negative return

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity
Contrarian	44%	21%	41%	39%	24%	44%	37%	28%	44%
Neutral	36%	15%	31%	25%	17%	32%	22%	21%	35%
Momentum	44%	17%	31%	33%	20%	29%	29%	25%	34%

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	37%	33%	46%	41%	39%	50%	49%	46%	53%
Neutral	23%	26%	40%	26%	32%	44%	33%	39%	48%
Momentum	28%	29%	39%	31%	35%	44%	34%	42%	48%

Panel C - Observations

***Top* 1%**			***Top* Quintile**			**Quintile 4**		
1372	1577	1200	32150	32374	27500	35149	39471	34863
1299	1882	1721	28046	36270	35070	24429	32803	34737
1802	2081	2076	30252	36723	40180	28328	35387	39382

Quintile 3			**Quintile 2**			***Bottom* Quintile**		
33760	43330	40340	31688	47015	47848	26367	52927	57920
18547	27772	31155	13775	21697	25360	6840	13268	16006
26074	32834	36821	23208	30760	33881	15943	27520	30782

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors are divided by their average monthly turnover (%) into three groups (terciles) of activity. Turnover is measured by the monetary volume of buy and sell orders divided by two times the average monthly holding. Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock are stocks with the past 12-month return below the median. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations within each group is reported on Panel C. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

Table A9 – Annual returns and proportions of investors with negative monetary result by the intersection of wealth groups, terciles of monthly turnover (%), and contrarian investment style - using time-series classification with past 12-month returns.

Panel A - Group returns (Year %)

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity
Contrarian	-1.0	-3.3	-16.4	0.0	-5.9	-17.5	0.1	-4.6	-10.9
Neutral	0.8	-1.8	-6.8	1.4	-0.9	-5.2	2.8	0.7	-3.8
Momentum	-1.9	-0.1	-4.5	0.8	0.9	0.4	1.8	1.4	0.1

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	0.8	-4.1	-17.7	0.5	-1.4	-13.6	-0.7	-4.0	-16.9
Neutral	3.0	0.7	-3.9	3.7	0.1	-4.8	4.6	1.3	-6.2
Momentum	2.1	1.9	-1.6	1.4	1.6	-0.1	2.7	1.1	-1.2

Panel B - Percentage of investors with negative return

	***Top* 1%**			***Top* Quintile**			**Quintile 4**		
	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity	Low Activty	Medium Activity	High Activity
Contrarian	50%	44%	66%	50%	51%	68%	51%	52%	64%
Neutral	38%	29%	48%	33%	31%	49%	29%	31%	47%
Momentum	39%	12%	26%	25%	13%	25%	21%	17%	30%

	Quintile 3			**Quintile 2**			***Bottom* Quintile**		
Contrarian	50%	53%	62%	53%	58%	62%	60%	63%	61%
Neutral	30%	35%	49%	33%	40%	52%	39%	44%	53%
Momentum	22%	22%	35%	25%	27%	40%	29%	32%	44%

Panel C - Observations

***Top* 1%**			***Top* Quintile**			**Quintile 4**		
1040	572	527	22200	12144	13087	22511	16249	16411
984	758	665	18322	15901	14673	15578	16976	16352
2449	4210	3805	49926	77322	74990	49817	74436	76219

Quintile 3			**Quintile 2**			***Bottom* Quintile**		
21510	19519	19315	20500	23255	24903	18178	30969	34424
12421	15920	16985	9355	13942	15780	4822	10276	12336
44450	68497	72016	38816	62275	66406	26150	52470	57948

Note: Investors wealth is defined by his wealth in direct equity holdings. Investors are ranked i) by their total wealth in December 2011 for the existing accounts and ii) by their maximum wealth within three years for new accounts (opened between 2012-2019). Investors are divided by their average monthly turnover (%) into three groups (terciles) of activity. Turnover is measured by the monetary volume of buy and sell orders divided by two times the average monthly holding. Investors' contrarian investment style is characterized by their purchases. Contrarians have 60% of their trades buying a loser stock, while momentum have less than 40% of their trades buying a loser stock. Loser stock are stocks with past 12-month return below zero. Only investors who buy at least one stock in the stock market can be classified. Investors who participate only in the derivatives market, or who only sold stocks for the entire sample period are left without a classification. The total number of observations within each group is reported on Panel C. Annual (%) is calculated accruing the daily returns obtained from the marginal daily gain and loss. Daily returns are calculated by the difference between the cumulative trading gain of day (t) minus (t-1) divided by the holdings value of day (t-1). Investors with negative results are investors with the cumulative trading measure below zero at the end of 2019.

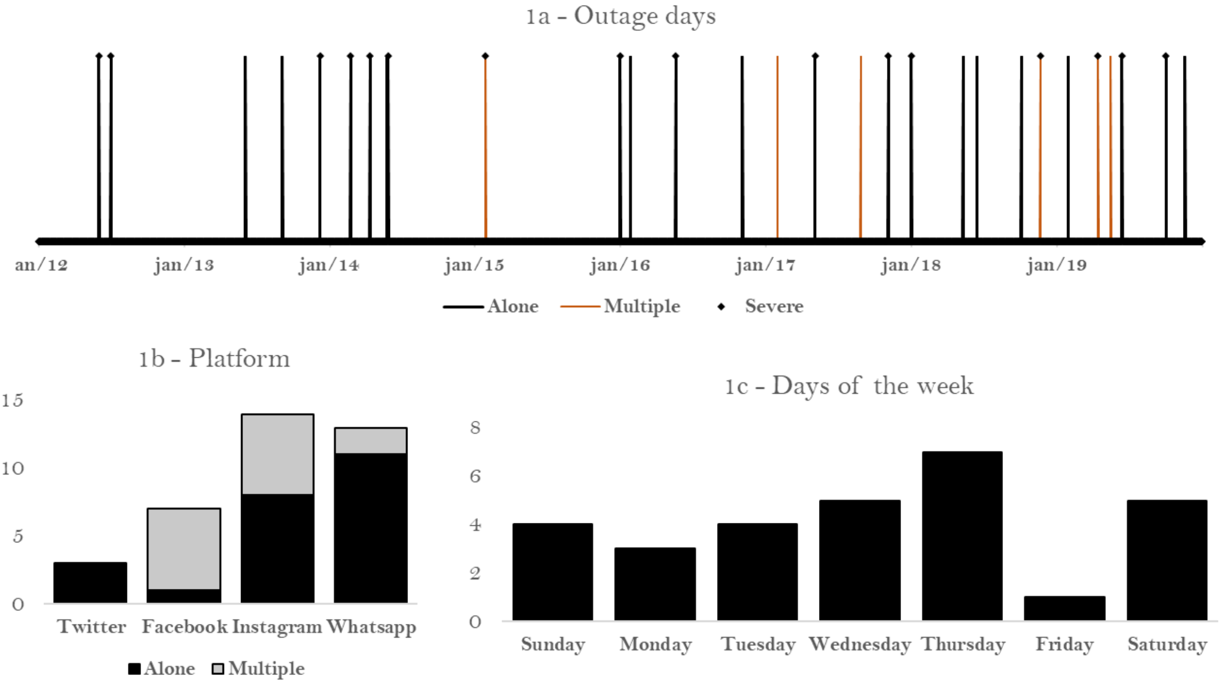
Table A10 – Definitions of the counterfactual exercise using net flows.

Counterfactual	**Counterfactual exercise**	**Which redistribution factors are shut down?**
Baseline (A)	Investors hold the **same exact portfolio** in the beginning and **do not** trade	All.
Initial holdings return heterogeneity (B)	Investors hold their own portfolio in the beginning and **do not** trade	New trading across individual stocks, derivatives contracts, and net flows.
Net Flows (C)	Investors hold their own portfolio in the beginning and **trade only in/out the equity market**, with negative or positive net flows being invested in a **zero-return** asset.	New trading across individual stocks and derivatives contracts.
Benchmark (D)	Investors hold their own portfolio in the beginning and **trade freely** across different individual stocks (at their average intraday execution price), derivatives contracts and in/out stock market.	None.

Note: Table A10 defines the counterfactual exercises for the decomposition of changes in the inequality measure assuming that positive net flows are created instantly, and negative net flows are consumed right away. Under this hypothesis the decomposition will have a new counterfactual, (C) “Net Flow”, which is the counterfactual Gini Index when investors hold their own portfolio and trade in/out of equity market but investing in a zero-return asset. The assumption of investment in a zero-return asset is made to shun away in this counterfactual from the effects of market timing (differences in return between risky and risk-free asset), which is then captured in (D), within the new trading contribution. The difference between the Gini Index calculated from (B) and (A) gives the contribution to the total variation from initial holdings return heterogeneity. Differences between the Gini Index calculated from (C) and (B) give the contribution of net flows in/out of equity market. Differences between the Gini Index calculated from (D) and (C), give the contribution from new trading in the stock and derivatives market.

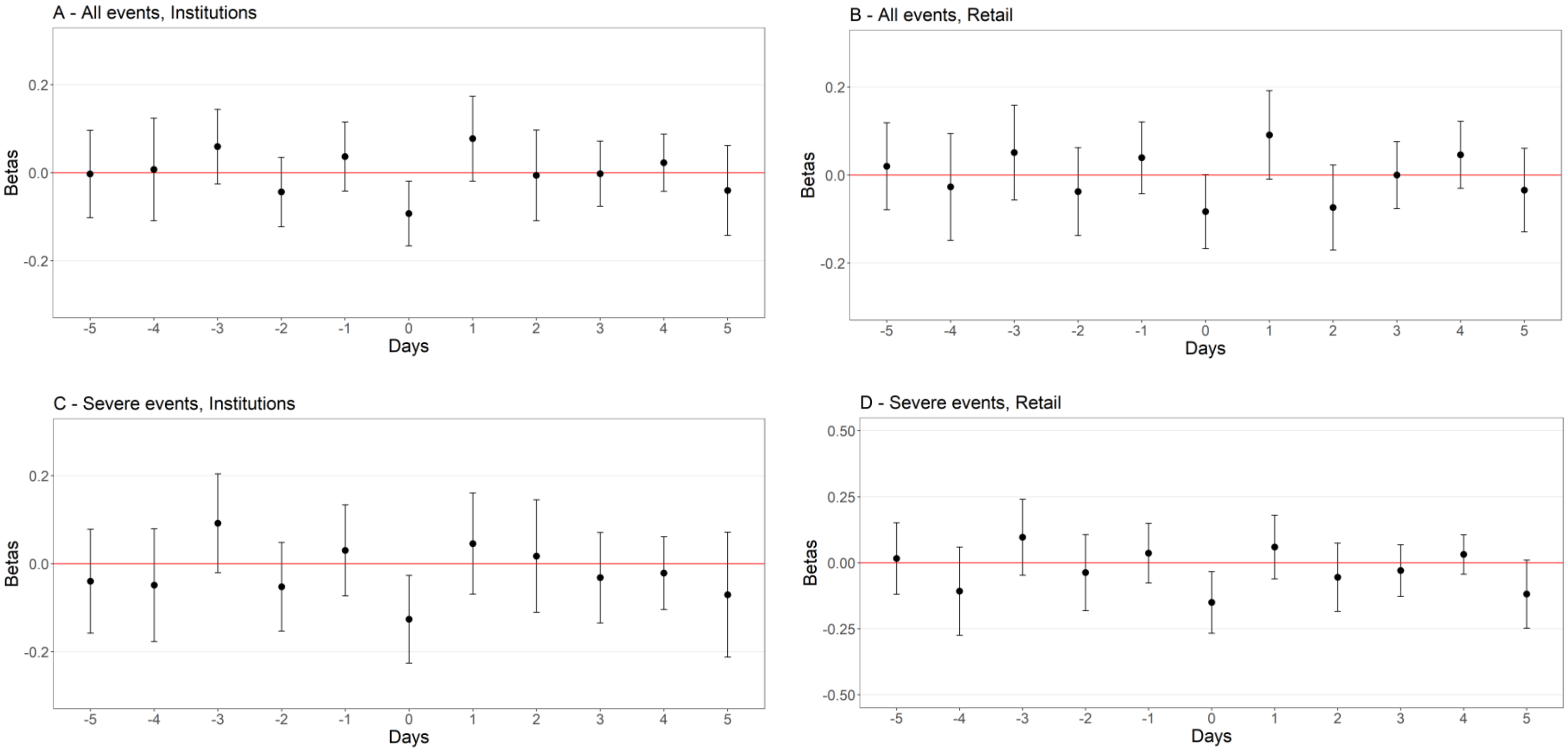
APPENDIX B – Chapter 4

Figure A1 – Timeline, social media platforms and days of the week of interruptions occurring outside market hours



Note: Information on the interruptions is collected from news articles. Alone outages are when there is only one platform that had interruptions problems on that day. Multiple outages are when at least two platforms had interruptions on that day. Severe interruptions last at least 1 hour, are reported in the news as a generalized outage (not only for some users) and are officially acknowledged by the company representatives.

Figure A2 - Social media outages affecting trading monetary volume.



Note: The figure reports point estimates and the respectively 95% level of confidence intervals for the outage dummies of the time-series model of equation (1) and for ten new dummies that are equal when it is one to five days before the outages, and one to five days after the outages, and zero otherwise. All models equally weight the stocks for each day of the stock-day panel and are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method. Models include 20 day of the month dummies, 5 day of the week dummies, and 96 monthly dummies. “All events” consider as outages all 33 interruptions described on table 1, while “severe events” consider the 18 severe outages described in table 1.

Table A1 – Interruptions of social media platforms outside market hours.

Date	Platform	Start	End	Time Span (Min)	Market Hour Time Span (Min)	Weekday	Type of problem	Multiple Platform	Severe
31/05/2012	facebook	18:45	20:30	105	0	Thursday	outage	No	1
30/06/2012	instagram	1:00	14:00	780	0	Saturday	outage	No	1
03/06/2013	twitter	17:30	18:15	45	0	Monday	outage	No	0
04/09/2013	twitter	17:40	18:20	40	0	Wednesday	outage	No	0
07/12/2013	whatsapp	19:00	21:30	150	0	Saturday	outage	No	1
22/02/2014	whatsapp	15:30	19:30	240	0	Saturday	outage	No	1
12/04/2014	instagram	13:00	15:00	120	0	Saturday	outage	No	1
25/05/2014	whatsapp	15:00	16:00	60	0	Sunday	feature instability	No	0
27/05/2014	instagram	17:45	19:30	105	0	Tuesday	outage	No	1
27/01/2015	facebok; instagram	4:15	5:10	60	0	Tuesday	outage	Yes	1
31/12/2015	whatsapp	13:30	15:00	90	0	Thursday	feature instability	No	1
26/01/2016	whatsapp	0:30	2:00	90	0	Tuesday	feature instability	No	0
19/05/2016	whatsapp	17:00	18:30	90	0	Thursday	outage	No	1
02/11/2016	whatsapp	20:00	21:00	60	0	Wednesday	feature instability	No	0
29/01/2017	facebok; instagram	11:00	11:45	45	0	Sunday	outage	Yes	0
03/05/2017	whatsapp	17:10	19:30	140	0	Wednesday	outage	No	1
26/08/2017	facebok; instagram	10:15	11:15	60	0	Saturday	access instability	Yes	0
03/11/2017	whatsapp	5:45	7:15	90	0	Friday	outage	No	1
31/12/2017	whatsapp	16:00	17:40	100	0	Sunday	outage	No	1
10/05/2018	instagram	9:30	10:00	30	0	Thursday	feature instability	No	0
14/06/2018	whatsapp	2:00	2:30	30	0	Thursday	access instability	No	0
03/10/2018	instagram	2:15	3:00	45	0	Wednesday	outage	No	0
20/11/2018	facebok; instagram	10:30	11:30	60	0	Tuesday	feature instability	Yes	1
29/01/2019	instagram	22:30	23:30	60	0	Monday	access instability	No	0
14/04/2019	facebok; instagram; whatsapp	8:00	10:00	120	0	Sunday	access instability	Yes	1
16/05/2019	facebok; instagram; whatsapp	18:20	19:00	40	0	Thursday	access instability	Yes	0
13/06/2019	instagram	18:30	20:00	90	0	Thursday	feature instability	No	1
02/10/2019	twitter	22:30	23:59	120	0	Wednesday	feature instability	No	1
18/11/2019	instagram	19:00	20:00	60	0	Monday	access instability	No	0

Note: Information on the interruptions is collected from news articles. The common factors of different news articles are used to define the time the outage began, the duration, the typical problem users are experiencing, and if this is a generalized interruption or if it only affects part of users. ‘Feature instability’ is when users have difficulties posting or reading on social media, ‘access instability’ is when users have intermittent difficulties logging on, while ‘outage’ is when the app and website are completely down. Severe interruptions last at least 1 hour, are reported in the news as a generalized outage (not only for some users) and are officially acknowledged by the company representatives.

Table A2 – Social media outages affecting trading monetary volume with different time dummies.

Dependent Variable: ΔLog(Monetary Volume + 1)

Investor Category	**Institutions**		**Retail**	
Model	EW	EW	EW	EW
Panel A - All events				
Outage$_t$	-0.096*** (-2.75)	-0.075* (-1.89)	-0.113** (-2.39)	-0.092* (-1.71)
Log(Market Capitalization)$_{t-1}$	-0.015 (-0.521)	-0.028 (-0.929)	-1.16 (-1.26)	-1.84* (-1.69)
Return$_{t-1}$	-0.007 (-0.692)	0.010 (1.13)	-0.010 (-0.705)	0.011 (0.767)
Trend$_t$	-5.72e-5 (-0.020)	-0.003 (-0.985)		
Day of the month dummies	Yes	Yes	Yes	Yes
Day of the week dummies	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No
Month FE	Yes	Yes	No	No
Weekly FE	No	No	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974
R2	0.22	0.08	0.27	0.15
Panel B - Severe events				
Outage$_t$	-0.119*** (-2.65)	-0.139** (-2.51)	-0.153*** (-2.61)	-0.167** (-2.14)
Log(Market Capitalization)$_{t-1}$	-0.015 (-0.500)	-0.028 (-0.910)	-1.19 (-1.29)	-1.85* (-1.70)
Return$_{t-1}$	-0.007 (-0.689)	0.011 (1.15)	-0.010 (-0.722)	0.011 (0.767)
Trend$_t$	9.52e-5 (0.033)	-0.003 (-0.900)		
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974
R2	0.22	0.08	0.27	0.15
Panel C - Weak events				
Outage$_t$	-0.067 (-1.23)	0.0010 (0.017)	-0.068 (-0.940)	-0.008 (-0.098)
Log(Market Capitalization)$_{t-1}$	-0.016 (-0.542)	-0.028 (-0.945)	-1.20 (-1.29)	-1.88* (-1.72)
Return$_{t-1}$	-0.007 (-0.714)	0.010 (1.12)	-0.010 (-0.722)	0.010 (0.757)
Trend$_t$	-0.0003 (-0.096)	-0.003 (-1.02)		
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974
R2	0.21	0.07	0.27	0.15
Panel D - Out of market hours events				
Outage$_t$	0.007 (0.209)	-0.017 (-0.461)	0.030 (0.637)	-0.019 (-0.396)
Log(Market Capitalization)$_{t-1}$	-0.015 (-0.518)	-0.029 (-1.02)	-1.22 (-1.31)	-1.88*(-1.71)
Return$_{t-1}$	-0.007 (-0.713)	0.010 (1.11)	-0.010 (-0.730)	0.010 (0.754)
Trend$_t$	-0.0002 (-0.076)	-0.003 (-1.04)		
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974
R2	0.21	0.07	0.27	0.15

Note: All models are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation. Day of the month dummies are 20 dummies for different trading days of the month, from -10 to +10. Day of the week dummies are five weekday dummies. Month FE are twelve-month dummies and Year FE are eight-year dummies. Trend is a daily linear trend. Weekly FE are dummies for every week of the sample (417 dummies). FE$^{\#}$ in panels B to D indicates the use of the same dummies as in panel A. EW model equally weights the stocks for each day of the stock-day panel. Panel A consider as outages all 33 interruptions described on table 1. Panel B considers the 18 severe outages described in table 1, while Panel C considers the 15 non-severe outages described in table 1. Panel D considers the 29 outages described in table A1. Outages outside market hours are shifted to the next available trading day.

Table A3 – Social media outages affecting trading monetary volume - different dependent variables.

Dependent Variable	**Log(Monetary Volume + 1)**		**Monetary Vol. / Market Cap.**	
Investor Category	**Institutions**	**Retail**	**Institutions**	**Retail**
Panel A - All events				
Outage$_t$	-0.051* (-1.70)	-0.032 (-1.11)	-0.004 (-0.661)	0.006 (0.732)
Log(Market Capitalization)$_{t-1}$	1.43*** (4.54)	2.45*** (5.93)	-0.014 (-0.382)	0.151** (2.46)
Return$_{t-1}$	0.014* (1.82)	0.033*** (3.43)	-0.001 (-0.795)	0.008*** (5.47)
Ave. Log(Mon. Vol. + 1)$_{t-1:t-22}$	-0.045 (-0.446)	0.022 (0.215)		
Ave. (Mon. Vol. / Mkt Cap.)$_{t-1:t-22}$			-0.086 (-0.721)	0.382** (2.08)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,953	1,953	1,952	1,952
R2	0.91	0.95	0.51	0.63
Panel B - Severe events				
Outage$_t$	-0.058 (-1.34)	-0.057* (-1.79)	-0.012** (-2.57)	-0.006 (-0.765)
Log(Market Capitalization)$_{t-1}$	1.42*** (5.09)	2.45*** (6.04)	-0.014 (-0.387)	0.151** (2.46)
Return$_{t-1}$	0.014* (1.88)	0.033*** (3.49)	-0.0010 (-0.777)	0.008*** (5.44)
Ave. Log(Mon. Vol. + 1)$_{t-1:t-22}$	-0.043 (-0.479)	0.024 (0.239)		
Ave. (Mon. Vol. / Mkt Cap.)$_{t-1:t-22}$			-0.085 (-0.714)	0.385** (2.09)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,953	1,953	1,952	1,952
R2	0.91	0.95	0.51	0.63
Panel C - Weak events				
Outage$_t$	-0.041 (-0.909)	-0.004 (-0.073)	0.006 (0.567)	0.020 (1.12)
Log(Market Capitalization)$_{t-1}$	1.43*** (4.55)	2.45*** (5.84)	-0.015 (-0.403)	0.150** (2.43)
Return$_{t-1}$	0.014* (1.81)	0.033*** (3.36)	-0.001 (-0.802)	0.008*** (5.33)
Ave. Log(Mon. Vol. + 1)$_{t-1:t-22}$	-0.046 (-0.464)	0.022 (0.213)		
Ave. (Mon. Vol. / Mkt Cap.)$_{t-1:t-22}$			-0.088 (-0.742)	0.383** (2.05)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,953	1,953	1,952	1,952
R2	0.91	0.95	0.51	0.63
Panel D - Out of market hours events events				
Outage$_t$	0.013 (0.373)	-0.010 (-0.293)	-0.002 (-0.481)	0.001 (0.192)
Log(Market Capitalization)$_{t-1}$	1.43*** (3.96)	2.45*** (5.94)	-0.014 (-0.393)	0.151** (2.42)
Return$_{t-1}$	0.014* (1.81)	0.033*** (3.42)	-0.001 (-0.805)	0.008*** (5.26)
Ave. Log(Mon. Vol. + 1)$_{t-1:t-22}$	-0.045 (-0.401)	0.021 (0.213)		
Ave. (Mon. Vol. / Mkt Cap.)$_{t-1:t-22}$			-0.088 (-0.739)	0.384** (2.03)
Fixed-Effects$^{\#}$	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW
Observations	1,953	1,953	1,952	1,952
R2	0.91	0.95	0.51	0.63

Note: All models are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation. FE$^{\#}$ includes 20 days of the month dummies, 5 day of the week dummies, and 96 monthly dummies. EW model equally weights the stocks for each day of the stock-day panel. Panel A considers as outages all 33 interruptions described on table 1. Panel B considers the 18 severe outages described in table 1, while Panel C considers the 15 non-severe outages described in table 1. Panel D considers the 29 outages described in table A1. Outages outside market hours are shifted to the next available trading day. For dependent variables which are not in their first difference we include the term $\rho * \overline{Y_{t-1:t-22}}$ in equation (1) to account for the persistency of the dependent variable.

Table A4 – Different social media outages affecting number of active investors and monetary volume for different investors categories.

Panel A

Dependent			**ΔLog(Monetary Volume + 1)**					
Investor Category		**Institutions**				**Retail**		
Social Media	Facebook	Instagram	Twitter	Whatsapp	Facebook	Instagram	Twitter	Whatsapp
Outage_t	-0.245*** (-7.29)	-0.070** (-2.73)	-0.207* (-1.69)	-0.111 (-1.28)	-0.194*** (-5.11)	0.044 (1.18)	-0.311* (-1.72)	-0.128 (-1.63)
$\text{Log(Market Cap)}_{t-1}$	-0.578*** (-3.06)	-0.574*** (-3.04)	-0.576*** (-3.05)	-0.571*** (-3.02)	-0.705*** (-3.39)	-0.700*** (-3.36)	-0.704*** (-3.37)	-0.698*** (-3.36)
Return_{t-1}	-0.008 (-0.715)	-0.007 (-0.702)	-0.007 (-0.695)	-0.007 (-0.684)	0.011 (1.04)	0.011 (1.05)	0.011 (1.06)	0.011 (1.07)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974	1,974	1,974	1,974	1,974
R2	0.22	0.22	0.22	0.22	0.08	0.08	0.08	0.08

Panel B

Dependent			**ΔLog(#Investors + 1)**					
Investor Category		**Institutions**				**Retail**		
Social Media	Facebook	Instagram	Twitter	Whatsapp	Facebook	Instagram	Twitter	Whatsapp
Outage_t	-0.071*** (-7.58)	-0.009 (-1.21)	-0.054* (-1.92)	-0.038** (-2.32)	-0.120*** (-7.24)	-0.013 (-0.716)	-0.122 (-1.33)	-0.090** (-2.03)
$\text{Log(Market Cap)}_{t-1}$	-0.138*** (-2.74)	-0.137*** (-2.71)	-0.137*** (-2.71)	-0.136*** (-2.69)	-0.296*** (-2.94)	-0.294*** (-2.93)	-0.295*** (-2.93)	-0.292*** (-2.91)
Return_{t-1}	-0.002 (-0.707)	-0.001 (-0.686)	-0.001 (-0.677)	-0.001 (-0.658)	0.016** (3.18)	0.016** (3.20)	0.016*** (3.20)	0.016*** (3.27)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov type	NW	NW	NW	NW	NW	NW	NW	NW
Observations	1,974	1,974	1,974	1,974	1,974	1,974	1,974	1,974
R2	0.30	0.30	0.30	0.30	0.11	0.11	0.11	0.11

Note: All models are estimated by OLS with standard errors calculated using Newey and West variance-covariance matrix estimation, with the lag being automatically selected by the Newey and West (1994) method. Fixed-Effects include 20 day of the month dummies, 5 day of the week dummies, and 96 monthly dummies. All models equally weight the stocks for each day of the stock-day panel. All models consider only the severe events which happened only for that particular social media (“alone” outage on that day) as described in table 1.

Table A5 - Firm level heterogeneity of social media outages affecting number of investors.

ΔLog(#Investors + 1)

Panel A - Retail	**All Events**						**Severe Events**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outage$_t$	-0.030 (-1.6)	-0.030 (-1.6)	-0.030 (-1.6)	-0.030 (-1.6)	-0.030* (-1.7)	-0.030* (-1.7)	-0.073*** (-2.7)	-0.073*** (-2.7)	-0.072*** (-2.7)	-0.073*** (-2.7)	-0.071*** (-2.7)	-0.070*** (-2.7)
Return$_{f,t-1}$	0.003*** (7.5)	0.003*** (7.5)	0.003*** (7.6)	0.003*** (7.5)	0.003*** (7.5)	0.003*** (7.6)	0.003*** (7.5)	0.003*** (7.5)	0.003*** (7.6)	0.003*** (7.5)	0.003*** (7.5)	0.003*** (7.6)
Log(Market Capitalization)$_{f,t-1}$	-0.127*** (-7.8)	-0.127*** (-7.8)	-0.127*** (-7.8)	-0.126*** (-7.8)	-0.126*** (-7.8)	-0.125*** (-7.8)	-0.127*** (-7.8)	-0.127*** (-7.8)	-0.127*** (-7.8)	-0.126*** (-7.8)	-0.126*** (-7.8)	-0.125*** (-7.8)
Outage$_t$ * Log(Market Capitalization)$_{f,t-1}$		0.009 (1.0)				0.008 (0.99)		0.002 (0.18)				0.002 (0.15)
Outage$_t$ * Return$_{f,t-1}$			-0.003 (-1.2)			-0.003 (-1.2)			-0.006* (-1.8)			-0.005* (-1.7)
Return$_{f,t-2:t-5}$				0.001 (1.5)		0.0008 (1.3)				0.001 (1.6)		0.0009 (1.4)
Outage$_t$ * Return$_{f,t-2:t-5}$				-0.0005 (-0.15)		-0.0003 (-0.10)				-0.004 (-0.88)		-0.003 (-0.71)
Return$_{f,t-6:t-22}$					0.004* (1.9)	0.003* (1.8)					0.004* (1.9)	0.003* (1.9)
Outage$_t$ * Return$_{f,t-6:t-22}$					-0.001 (-0.15)	-0.0002 (-0.03)					-0.013 (-1.2)	-0.011 (-1.1)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	807,215	807,215	807,215	807,215	807,214	807,214	807,215	807,215	807,215	807,215	807,214	807,214
R2	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007

Panel D - Domestic Institutions	**All Events**						**Severe Events**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outage$_t$	-0.026* (-2.4)	-0.026* (-2.4)	-0.025** (-2.4)	-0.026** (-2.5)	-0.026** (-2.5)	-0.026** (-2.5)	-0.049*** (-3.1)	-0.049*** (-3.1)	-0.048** (-3.1)	-0.049*** (-3.1)	-0.049*** (-3.2)	-0.048*** (-3.2)
Return$_{f,t-1}$	-0.0009*** (-3.8)	-0.0009*** (-3.8)	-0.0008*** (-3.5)	-0.0009*** (-3.8)	-0.0009*** (-3.8)	-0.0008*** (-3.6)	-0.0009*** (-3.8)	-0.0009*** (-3.8)	-0.0008*** (-3.6)	-0.0009*** (-3.8)	-0.0009*** (-3.8)	-0.0009*** (-3.7)
Log(Market Capitalization)$_{f,t-1}$	-0.054*** (-6.6)	-0.054*** (-6.6)	-0.054*** (-6.6)	-0.053*** (-6.6)	-0.053*** (-6.6)	-0.053*** (-6.6)	-0.054*** (-6.7)	-0.054*** (-6.7)	-0.054*** (-6.7)	-0.053*** (-6.6)	-0.053*** (-6.6)	-0.053*** (-6.6)
Outage$_t$ * Log(Market Capitalization)$_{f,t-1}$		-0.003 (-0.47)				-0.003 (-0.51)		-0.009 (-1.4)				-0.010 (-1.4)
Outage$_t$ * Return$_{f,t-1}$			-0.003 (-1.5)			-0.003 (-1.6)			-0.003 (-1.5)			-0.004 (-1.6)
Return$_{f,t-2:t-5}$				0.0006* (1.8)		0.0005 (1.5)				0.0007* (1.8)		0.0005 (1.6)
Outage$_t$ * Return$_{f,t-2:t-5}$				0.001 (0.68)		0.001 (0.73)				-0.0003 (-0.15)		-0.0004 (-0.2)
Return$_{f,t-6:t-22}$					0.002** (2.0)	0.001* (1.9)					0.002** (2.0)	0.001* (1.94
Outage$_t$ * Return$_{f,t-6:t-22}$					0.0007 (0.17)	0.0009 (0.24)					0.0006 (0.15)	0.002 (0.43)
Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	807,215	807,215	807,215	807,215	807,214	807,214	807,215	807,215	807,215	807,215	807,214	807,214
R2	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
Within R2	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003

Note: All models are estimated with standard errors clustered by day and stock. Day of the month dummies are 20 dummies for different trading days of the month, from -10 to +10. Day of the week dummies are five weekday dummies. Firm-Monthly FE are a set of dummies for every month of the sample (96 dummies) for each firm (485), a total of 46,560 dummies. Fixed-Effects in panel B indicate the use of the same dummies as in panel A. “All events” consider as outages all 33 interruptions described on table 1, while “severe events” consider the 18 severe outages described in table 1.

Table A6 – Social media outages affecting trading activity for investors with different frequency of trading.

Panel A - Retail

Dependent	**ΔLog(Monetary Volume + 1)**		**ΔLog(# Investors + 1)**	
Events	All	Severe	All	Severe
	(1)	(2)	(3)	(4)
Outage$_t$ * Below Median Activity	-0.072 (-1.36)	-0.138* (-1.81)	-0.019 (-0.999)	-0.067** (-2.50)
Outage$_t$ * Above Median Activity	-0.086. (-1.86)	-0.124** (-2.02)	-0.030 (-1.60)	-0.062** (-2.32)
Outage$_t$ * Top Decil Activity	-0.077* (-1.69)	-0.137** (-2.42)	-0.030 (-1.64)	-0.062** (-2.41)
Return$_{f,t-1}$	0.005*** (4.62)	0.005*** (4.63)	0.003*** (7.84)	0.003*** (7.85)
Log(Market Capitalization)$_{f,t-1}$	-0.316*** (-8.40)	-0.316*** (-8.40)	-0.127*** (-8.69)	-0.127*** (-8.70)
Above Median Activity	0.0007 (0.168)	0.0003 (0.078)	0.0004 (0.271)	0.0002 (0.144)
Top Decil Activity	0.0005 (0.125)	0.0004 (0.105)	0.0004 (0.221)	0.0002 (0.096)
Fixed-Effects	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	2,421,645	2,421,645	2,421,645	2,421,645
R2	0.002	0.002	0.006	0.006

Panel B - Domestic Institutions

Dependent	**ΔLog(Monetary Volume + 1)**		**ΔLog(# Investors + 1)**	
Events	All	Severe	All	Severe
	(1)	(2)	(3)	(4)
Outage$_t$ * Below Median Activity	-0.109 (-1.08)	-0.351*** (-3.13)	-0.030 (-1.42)	-0.075*** (-2.85)
Outage$_t$ * Above Median Activity	-0.174*** (-2.91)	-0.219*** (-2.75)	-0.034** (-2.21)	-0.051*** (-3.07)
Outage$_t$ * Top Decil Activity	-0.055 (-1.50)	-0.115** (-2.07)	-0.012 (-1.49)	-0.027** (-2.00)
Return$_{f,t-1}$	-0.003*** (-2.85)	-0.003*** (-2.85)	-0.0003** (-2.28)	-0.0003** (-2.27)
Log(Market Capitalization)$_{f,t-1}$	-0.245*** (-6.97)	-0.246*** (-6.99)	-0.033*** (-4.89)	-0.033*** (-4.91)
Above Median Activity	0.0008 (0.076)	-0.001 (-0.130)	6.99e-5 (0.026)	-0.0002 (-0.078)
Top Decil Activity	-0.0007 (-0.066)	-0.002 (-0.178)	-0.0002 (-0.058)	-0.0003 (-0.114)
Fixed-Effects	Yes	Yes	Yes	Yes
VarCov Clustered	Firm & Day	Firm & Day	Firm & Day	Firm & Day
Observations	2,421,645	2,421,645	2,421,645	2,421,645
R2	0.002	0.002	0.005	0.005

Note: All models are estimated with standard errors clustered by day and stock. Fixed-Effects include twenty days of the month dummies, five days of the week dummies, and 46,560 firm-monthly dummies. “All events” consider as outages all 33 interruptions described on table 1, while “severe events” consider the 18 severe outages described in table 1. For each month and within each broad investor category, retail and domestic institutions, investors are split into three additional groups by their frequency of trading (days in a month). The first group have the investors that trade less, below the median for that month. The second group are investors above the median and below the top decile. While the last group are very active investors, above the top decile for that month. Domestic institutions are at the fund level, we cannot recover which funds have the same ‘parent’ institution.