1 Introduction

Can luck affect investor's trading behavior? A favorable trading outcome that results from *pure* luck should not provide any information about the investor's trading ability. As such, it should not alter the investor's trading pattern. In this paper, we study the effect of a purely random shock that resulted in a "lucky experience" to a sample of retail investors. We compare the subsequent trading activity of these lucky investors with that of other similar investors and find that the lucky investors do change their trading behavior: they increase portfolio turnover and worsen their trading performance.

Our results are consistent with behavioral reasons driving trading decisions. The behavioral literature has long discussed the effects of overconfidence on excessive trading and negative performance. Barber and Odean (2001) and Barber et al. (2020) find that investors that are naturally more overconfident tend to trade more and perform worse. Gervais and Odean (2001) attribute overconfidence to the process of incorrect learning, where investors facing a favorable sequence of trades mistakenly assume those as signals of their superior ability or information. To the extent that investors do not assume the lucky outcome as signal of ability and that the lucky treatment affects all investors equally, our results suggest a different interpretation. Our results are consistent with sensation seeking explanations proposed by Grinblatt and Keloharju (2009) instead.

We explore the unexpected price change produced by a salient environmental disaster to determine the lucky event. In November 5, 2015, a dam with iron tailings of the company Vale do Rio Doce in the city of Mariana suffered a catastrophic failure, immediately killing 19 people and releasing pollutants along 668 kilometres of watercourses.[1] In the following day of what has become known as the "Mariana disaster" the Vale stock fell by almost 6%[2] and the market value fell 8% [3]; after three months of the disaster, the stock price fell by 30%. The disaster was intensively covered by the media for several weeks in Brazil, hitting the frontpage of most Brazilian newspapers.[4] We use this unfortunate, and unanticipated, large negative shock to the stock price of Vale do Rio Doce to define our "lucky event". We say that investors who sold Vale stock one month prior to the event experienced pure luck.

In our case, the stock price change that we consider can be clearly attributed to an unanticipated event. As such, it should not entice any updates of the investor's own ability assessments. Any effect resulting from this price change would suggest that simple luck as opposed to any type of learning (biased or not) is behind the results. That is, the simple

[1] news 1, news 2

[2] news 3

[3] 14

[4] 13, news 4, news 5

feeling of success produced by the lucky event, regardless of how it was achieved, seem to induce the investor to trade more.

We explore a rich dataset with the complete stock trading activity by all Brazilian retail investors. Our dataset is at the investor-stock-day level and contains the number of shares the investor purchased or sold with the corresponding financial volume, from which we can infer average trading prices for both purchases and sales. We also have a number of other observables at the investor level such as the investor's gender, age, and experience, that we use to define matched groups of similar investors. In our benchmark definition, we define the lucky investors as those who's only trading decision on the month (21 trading days) prior to the disaster was to sell Vale's stock. Therefore, we exclude from this group individuals who sold and purchased other stocks to make the lucky event more salient to the investor. Importantly, we define the control group in a similar way. The control group in our benchmark exercise contains those investors who's only trading decision one month prior to the disaster was to sell a different stock than Vale's stock. As a robustness, we consider differing grouping of investors.

There are in total 501 retail investors in the treatment group and 5786 retail investors in the control group. Descriptive statistics reveal that investors in these two groups are similar to each other, except perhaps for the preference of the treatment group for the Vale stock. In both groups, the average investor's age is around 47, 85% of the investors are males, and the number of stocks they trade is 9. Treated investors have 5.8 years of experience and an average portfolio of 296 thousand reais. Investors in the control group have 5 years of experience and an average portfolio value of 277 thousand reais.

Once the control and treated groups are specified, we employ a difference-in-difference strategy. We compare the trading activity of control and treated investors before and after the Mariana disaster. In our benchmark analysis, we consider a window of three months before and three months after the disaster (excluding the month preceding the disaster that we use to classify investors into the groups). In our main regression, we compare the two groups using a number of control variables such as past trading activity, individual characteristics, total numbers of trades, the volume sold, and the proportion of the portfolio sold in the month preceding the disaster. Importantly, by including the volume and proportion of the portfolio sold as controls, we address the concern of a potential mechanical effect coming from Vale's holdings generating relatively larger sales proceed and subsequent higher turnover. As an additional step to ensure that we are comparing similar investors, we also run regressions where we use a propensity score to match investors with similar characteristics.

Our results indicate that lucky investors increase their portfolio turnover by 11% after the event when compared with the control group. Similarly, we also document an increase in 46% in the number of trades and in 35% the number of days with trade. Apart from an increase in the trading activity, we also document a worsening in investors performance. The portfolio return after the event is 6% lower for the luck investors when compared with the control group; 4% lower when returns are risk-adjusted.

We perform several robustness exercises. First, we rule out the concern of a spurious relation coming from a favorable draw of treated investors by randomly selecting investors from the control group and assigning them as if they were treated investors. We do this 10,000 times and compare the estimated coefficients with the obtained initially. We find that our estimate is in the extreme tail of this generated distribution of coefficients. Second, we look at differences in trading behavior between treated and controls around the same period in previous years. As expected, we find no statistically significant change in the observed trading activity and performance using past outcomes. Third, we look at what happens when we replace the Vale stock with the Petrobras stock to define the treated and non-treated investors using the original event date. Since Petrobras did not experience any relevant shock during this time, we should not find any result. Indeed, this is what we find.

The present work aims to contribute to two strands of the literature: one that studies the effects of behavioral biases in financial markets and the other that tries to explain the excessive trading puzzle by retail investors. We do so by exploring the impact of a large unexpected price shock that arguably led some investors to experience luck in the stock market. Our identification strategy is distinct from the closest articles that also study luck. For example, Anagol et al. (2021) explores how winning an IPO allotment lottery in India impacts the trading behavior of retail investors. Another paper that explores IPO allotment lotteries is Gao et al. (2021) that uses data from China. Instead, we employ a random shock to stock prices to infer a lucky experience. Although the purity of the luck event may vary (including how salient it is and how large it is), because the number of lucky experiences produced by random stock price changes is potentially much larger, we believe that understanding its effect on investor behavior is rather important.

The rest of this article is organized as follows: Section two discuss the related literature. Section three presents the data description and event description. In section four, have in detail the empirical strategy used. Section five reports empirical results and robustness checks. Section seven concludes.

1.1 Related Literature

A well-known stylized fact of the financial market is the high trading volume. As recognized by Odean (1999), rational explanations such as rebalancing portfolio, taxes, and liquidity demand seem insufficient to explain this pattern. Odean (1998) arguments that excessive trading is a consequence of investors that think they are better than they really are. This overconfidence bias makes investors engage in trades not driven by ability or information, lowering their performance. In contrast, rational investors would trade more carefully since they would engage in transitions only if the expected return is higher than trading costs Grossman and Stiglitz (1980). This duality of higher trade volume and underperformance is a well-documented fact in the literature that explores retail investors' performance, Barber and Odean (2013). Concerning this, Odean (1999) early noticed that retail investors make the wrong choice concerning portfolio selection since stocks they sold outperform those bought, including or not the trade cost and for many intervals after the trade.

Gervais and Odean (2001) theoretically investigates the source of overconfidence. In their model, investors process past trade information to infer their ability. Overconfidence is considered a biased learning process that outweighs knowledge's attribution as a source of successful trades. More importantly, the model generates two implications about overconfidence. First, positive past return increases this bias. That is, in bull markets, everyone is a genius. Second, overconfidence is higher at the beginning of the trader's career because investors do not have enough information about past success to update their beliefs correctly. Statman et al. (2006) gives empirical validity to these arguments by showing that market turnover volume responds positively to the lag of market returns.

Trading levels are likely to be distinct based on the investor characteristics. Barber and Odean (2001) documents that men have 45% higher portfolio turnover and lower returns than women. They argue that overconfidence can explain this pattern, since psychological studies showed that men in financial markets act more confidently than women.[5] However, gender is not a good instrument for overconfidence since it is subject to omitted variable bias. Because many differences between men and women are potentially correlated with excessive trading and overconfidence. For example, men could be more likely to engage in trades resembling gambling activities or think of trades more as excitement activity, thus chasing sensations.

Grinblatt and Keloharju (2009) explores overconfidence and sensation-seeking biases as explanations for excessive trading. They construct two precise measures for these biases using rich administrative data from Finland. The first is a self-confidence measure that identifies overconfident investors using questions about the individual's perceived ability relative to others contained in the mandatory admission process of Finland's army force. The second

[5] Indeed, the psychological evidence is that overconfidence is different when activity is perceived as male domain see Lundeberg et al. (1994), Deaux and Emswiller (1974), Lenney (1977), Beyer and Bowden (1997), Prince (1993)

measure uses excessive speed ticker convictions to classify sensation-seek investors. Their results indicate one additional ticker conviction is related to 7% more trades, and one point more in self-confidence measure increase 3% trade. Both effects are robust to controlling for investor characteristics such as wealth, age, and gender. But, sensation seeking seems to be better in explaining how much investors trade, not the decision to trade.

Liu et al. (2021) also proposes to test more than one behavioral bias capable of explaining excessive trading simultaneously, such as overconfidence, sensation seeking, gambling preferences[6], extrapolation[7], investor financial literacy[8], and realization utility[9]. They associate data from retail transactions and survey-based responses in China to propose a horserace between these biases. The results show that gambling trading and a perceived information advantage measures are the main reasons for excessive trading. Moreover, the explanatory power of sensation seeking is overturned when controlling for others' biases.

Another article that explores surveys to investigate behavior bias and trading is Glaser and Weber (2007), which investigates overconfidence bias and finds a higher portfolio turnover among this group relative to others. Moreover, Dorn and Sengmueller (2009) examines entertainment motives for trading using a survey constructed to elicit the enjoyment of trading by applying statements that identify compulsive gamblers. In line with the similarities of trading and gambling, their results indicate that investors who think of trades as entertainment show roughly twice portfolio turnover in relation to their peers, even after controlling for income, gender, age, and education. In the same direction, Gao and Lin (2015) reports a substitution effect of trading and gambling activities in Taiwan. On days with a hefty prize of lotteries, the trading volume in the exchange decreases between 5.2% and 9.1% among stocks that individuals respond to a significant fraction of trades. And the effect is higher for those stocks that are characterized as lottery-like stocks.

A further point is the impact of excessive trade on the return performance of retail investments. Barber and Odean (2000) noted that investors who trade too much have worse results than they don't trade, 11.4% against 18.5% annually. And remarkably, the performance is even worse than the market benchmark. On average, an investor turns 75% of the portfolio annually, which in the presence of higher transaction costs decreases their returns. Using a data set from Taiwan that contains trades from a diverse class of investors Barber et al. (2009) quantify the losses of excessive trading. They show an asymmetric pattern in the

[6]Preference for trading stocks that resembles gambling.Kumar (2009) noted that retail investors prefer stocks with low prices, high idiosyncratic volatility, and high specific skewness, which he defined as a lottery-like stock. Also related see Barber and Odean (2000), Shefrin and Statman (2000), Barberis and Huang (2008) Bordalo et al. (2012)

[7]Investors that extrapolates the trend. See Da et al. (2021), Barberis et al. (2018), Jin and Sui (2022)

[8]See Grinblatt et al. (2011), Lusardi and Mitchell (2014), Lusardi and Mitchell (2008)

[9]See Ingersoll and Jin (2013), Barberis and Xiong (2012)

returns of trades, where institutions gain and retail losses. Commissions and transactions taxes represents 66% the main charge, followed by trading losses (27%). Indeed, these results for retail are significant, with a penalty of 3.8 percentage points on the annual return.

Given the empirical evidence of the negative correlation between excessive trading and return performance, an important question is why investors keep trading to the detriment of their benefit. Barber et al. (2020) test two explanations. The first is learning to trade. Investor incurs a loss in the present to become an expert and profit in the future, leading to a positive final profit. The second is biased learning, which would generate losses in all periods because investors do not correctly learn how to trade. Their empirical result is consistent with the second explanation since a higher proportion of day-trading volume comes from day-traders that lose money. In the same line, Chague et al. (2020) studies day trade in the Brazilian context and finds that after 300 days, 97% of investors lose money, and only 1.1% earns more than the minimum wage.

Luck is often associated with positive outcomes in gambling activities, but many consider it a constant component of their lives, Schuster et al. (1989), Wiseman and Watt (2004). And when individuals believe in luck, they place more bets after a lucky shock, Darke and Freedman (1997). Also, Langer and Roth (1975) reports that in activities associated with luck, early success induces individuals to repeat this activity and increase the rate that they self-consider better in predicting the outcomes compared with individuals who did not experience early success. The similarity of trading activity with gambling suggests that investors who experienced a luck shock increase their trading activity, the central hypothesis in this study

Similar to our work that also examines the effect of luck on trading patterns among retail investors is Gao et al. (2021) which explores investors that won IPO allotment lotteries in China. The fact that IPO was underpriced in their samples implies that those investors got lucky. After overcoming this income shock, the results indicate that winning IPO induces investors to increase 28% the portfolio turnover. Consistent with the overconfidence literature, Gervais and Odean (2001), inexperienced investors have a higher impact. Moreover, lucky investors increase their portfolio volatilities and trade in lottery-like stocks consistent with sensation-seeking and gambling biases. The second paper is Anagol et al. (2021) which also explores IPO allotment lottery using data from India. Their sample contains IPO stocks that increase and decrease in value after the initial offer. Thus, using this feature, they show that winning a lottery generates a 7.2% higher trading volume when the IPO's stock increase in value. In the opposite case, the effect is negative. Furthermore, they argue that models in which investors misinterpret this noise signal (luck) as information about their ability better explain the results found compared with attention allocation and reinforcement learning models.

2 Event and Data description

This section is divided in two parts. First, we discuss why we choose the Mariana disaster as our random event. Second, we present the data that we have at our disposal.

2.1 Mariana disaster

To claim that a favorable trade was the result of pure luck we need to have an unpredictable shock to stock prices. In particular, we have to rule out the possibility of some investors having private information about the yet undisclosed piece of news.

To do so, we explore the disaster of Mariana dam on November 5, 2015. The dam was operated by Vale's subsidiary, Samarco. At the time, It was the most significant environmental disaster involving a mining company in the world. The dam break released 62 million cubic meters in rejection, which caused 19 deaths and affected 500 individuals that lived near the dam. Additionally, toxic residuals reached an essential river of the region (Rio Doce), impacting 230 cities downstream. The environment and economic damages were considerable.

We use this unfortunate disaster to determine a lucky event for those investors who sold Vale's stock before the disaster. We believe the disaster is suitable for our purposes for a couple of reasons. First, it was unanticipated by market participants. The collapse of the Dam was sudden and unexpected even for those with knowledge and working at the company. Second, the magnitude of the disaster, and its environmental consequences, implied that the event stayed on the media for several weeks. Figure 13 shows front page of the most important newspapers in Brazil. This enhances the salience of the Vale stock on the media, increasing the chances of investors noticing the event and realizing they were lucky for having sold the stock before. Finally, the size of the price fall also helps make the lucky experience more meaningful. The market value of Vale dropped 8% following the days after the disaster (about 6 billion of R$ in market capitalization, see Figure 14). The stock price continue to fall in the following days, reaching a relative devaluation of about 30% in the three months after the disaster.

The fact that we are using an unexpected price fall, as opposed to an unexpected price increase, is also important. A lucky event that increase the value of an investor's holdings could mechanically increase the investor's turnover. This is not the case for our lucky event, as the sales proceeds from selling the stock is not affected by the subsequent price changes that followed the disaster (we are assuming retail investors are not selling short the stock).

2.2 Retail investors trading activity data

Our dataset contains information about the trading activity of all retail investors on the Brazilian Stock Exchange. The source is CVM, the Brazilian Stock market regulator, equivalent to the SEC. It is at the level of investor-stock-day and contains the volume and quantity, sold or bought, for each investor-stock daily covering the period from January 2012 to December 2017. We also have information on about the age, gender, profession, state, city, and year of registration of the retail investors in our dataset. We combine this dataset with series of adjusted prices and split/inplit information from Economatica.

Following the literature on overconfidence, our main variable of interest is turnover. It is the amount traded by the investors during a period of time divided by the value of the investor's portfolio. A value larger than one indicates that the investors trades (purchases plus sales) more than the value of his portfolio in that period of time.

Since our dataset contains only trading flows, and not the investor's portfolios, we have to estimate their holdings. To do so, we proceed as follows. First, we compute the cumulative sum of the net daily flows for each investor-stock pair, accounting for inplits and splits, over the entire sample period 2012-2017. Then, we look at instances where the cumulative sum is negative and add the lowest value to the entire series to ensure the position is always possible. We implicitly assume that retail investors do not short-sell and that they are always selling stocks that they previously owned.

We also employ a liquidity filter to avoid having our results polluted by small and very illiquid stocks. Namely, we keep only stocks that had negotiations on more than 90% of the workdays each year from 2014 to 2016. From an initial set of 456 tickers, our final data set remains 186 tickers.

We follow Liu et al. (2021) and compute turnover in the following way. First, we sum all volumes bought or sold monthly and divide by the maximum value of the portfolio in that month. We also compute the portfolio return using Liu et al. (2021) as a reference, that is, the portfolio on the final of the month plus the volume sold and dividends minus the volume purchased and the value of the portfolio in the initial of the month, divided by the maximum value of the portfolio in that month. The fee is calculated by multiplying each transaction by a fixed cost of 3%, which is the custodian fee and a variable cost that depends on negotiated volume.

2.3 Treated and control groups

Our empirical strategy involves comparing the trading activity of retail investors who got luck (the treated group) with the other similar retail investors (the control group). Ideally, we want both group of investors to be identical except for the fact that the treated group experimented a lucky shock.

We define the treated group as those investors who only sold Vale's stocks in the window of one month (i.e., the 21 trading days before the event, excluding the disaster day, November 5, 2015), and did not trade (buy or sell) any other security. For the control group, we consider those investors who only sold stocks other than Vale at the same window as the treated group. Also, to avoid the control group being affected by the devaluation of Vale's stock, we considered only those who did not hold the stock in the portfolio at the event time.

We include in the control group only investors who were also selling to ensure that we are comparing between investors who were in "selling mode" before the disaster. Selling and buying decisions are likely to have different motivations, and comparing investors who are selling Vale with investors who are buying other stocks will likely suffer from the omission of important variables.

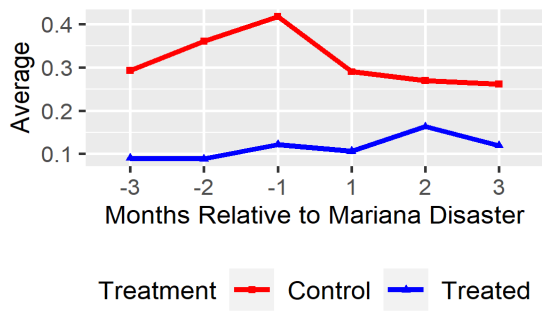
Moreover, we refrain from examining investors who purchased Vale stock before the disaster, i.e. the unlucky investors, because their holdings would be mechanically affected by the subsequent price fall in Vale stock. Additionally, the well-known disposition effect implies that these unlucky investors would be prone to trade less and to avoid realizing the loss, polluting the impact of luck on investor's turnover. Furthermore, we also compute the trading activity excluding Vale's stock in the computation to avoid this mechanical effect.[10]

Figures 1a, 1b, 2a and 2b reports the average of the outcomes series for treated and controls group. Visualizing a large decrease in trading activity for the control group after the event is possible. This pattern is consistent with market behavior during the period. For example, Figure 3a shows the number of trades and Figure 3b shows the volume of all investors. In contrast, the treated investors remained at the same level of trading after the event, with a slight increase in the second month after. This indicates that lucky investors do not follow the overall market by reducing their trading activities. Figure 2b is consistent with the luck impacting the investor performance since before the event, lucky investors have a higher level of portfolio return than control investors, and after the event, the relative level change between treated and control.

[10]Results are robust to the inclusion of Vale in the trading activity measures.

Figure 1: Turnover and Number of trades

(a) Turnover Dynamics



(b) Number of Trades Dynamics

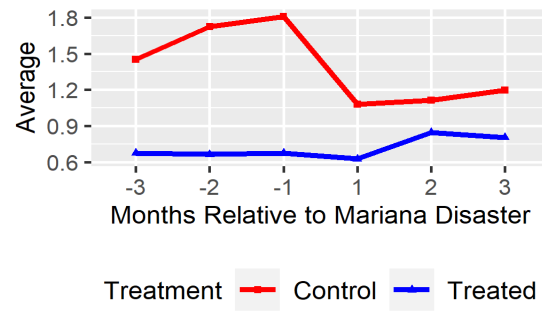
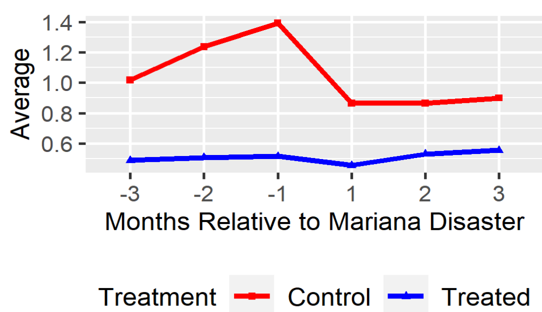


Figure 2: Number of Days with Trade

(a) Number of Days with Trade Dynamics



(b) Portfolio Return Dynamics

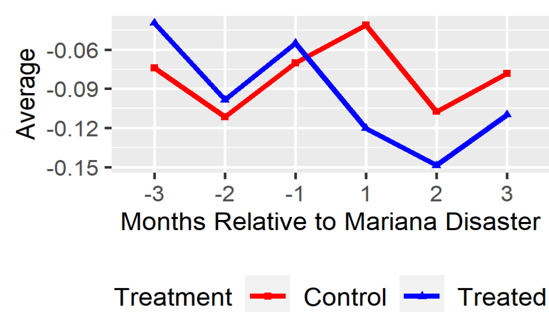
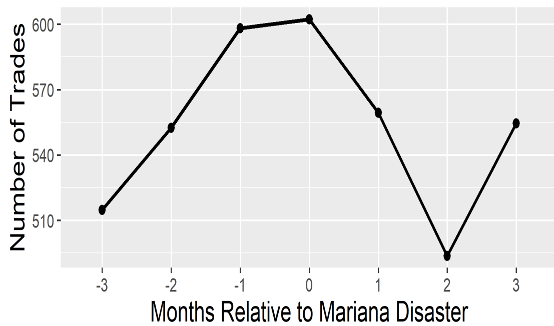


Figure 3: Number of Trades and Volume in overall market

(a) Number of Trades Dynamics



(b) Volume Dynamics

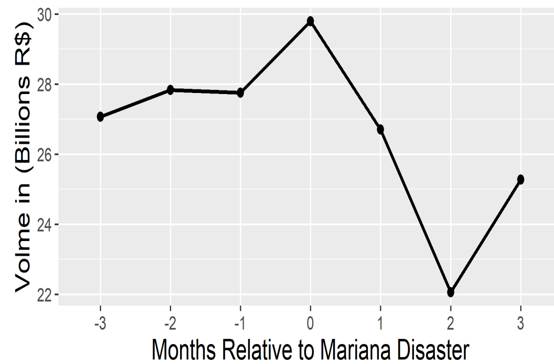


Table 1 reports the summary statistics from our sample, divided into treated and controls groups. Particularly, Panel A contains the trading behavior of treated and control in the pre-treatment window (from January of 2012 to one month before the luck event). The median-treated investor traded 29 days in 10 months, held a portfolio of R$ 31000, and made 35 trades in 7 different stocks. By comparison, the median-control investor traded 24 days in 12 months, held a portfolio of R$ 19.000, and made 28 trades in 6 different stocks.

A remarkable difference between the groups is the preference for Vale's Stocks. The sum of all volume traded by the median treated investor in Vale is R$ 91.000. By comparison, control investors are close to zero. However, focusing on the average investor, the preference is less remarkable, an indication of investors in control groups with a preference for Vale stocks. The last two rows of Panel A show the distribution for two covariates that measures investor ability, first is the Alpha of the investor's portfolio[11]. The second is the Return Risk-Adjusted 30 days ahead, which measures each investor's stock-picking ability. In concern to portfolio alphas, both groups seems to be similar. However, the stock-picking in a short time is relatively lower for the control group. Panel B reports trading activity in one month (21 trading days before the event). The treated average investor has lower activity in comparison with the control group. Moreover, the proportion of the portfolio sold during the interval is higher for the control group.

Panel C reports personal characteristics such as age, gender, and experience. About those covariates, both groups seem to be very similar. The average investor has 47 years, with 85% of them being male and having five years of experience. Panel D reports that descriptive summary of the outcomes. It can be noted that a higher average monthly turnover is 20% for

[11]That is, the constant coefficient of a Fama-French regression in each investor portfolio using 36 months of data previous to the unanticipated event.

treated and 33% for the control group when included Vale in the portfolio. For comparison, our measure of portfolio turnover is higher than Odean (1999), which documents a monthly portfolio turnover of 6.5%. But is lower than Liu et al. (2021), which documents an average of 84%. Performance measures also align with the stylized evidence that retail investors perform poorly. During the period, treated investors had an average return of -8%, almost the same for the control group -7%. The performance measure does not include Vale stocks because this stock suffered a devaluation and could lead to an automatic lower performance of the treated group. The same pattern occurs with risk-adjusted returns.

3 Empirical Strategy

The empirical strategy consists of the estimation of a difference-in-difference model. Based on Gao et al. (2021)[12] we select a small interval of time around the event, three months before and after (excluding the month preceding the event that was used to identify treated and control investors).

In the first specification, we estimate the average effect by collapsing the period before in a dummy variable that equals $Post_t = 0$ and $Post_t = 1$ after the event. The coefficient of interest is given by β and the treatment is $Treated_i * Post_{it}$, see equation (1). In the second specification, equation (2), we include a linear trend and its interaction with $Treated_i$ as a simple test of the parallel trends hypothesis.

$$Y_{i,t} = \beta Treated_i Post_t + \gamma Treated_i + \lambda Post_t + \alpha + \varepsilon_{i,t} \tag{1}$$

$$Y_{i,t} = \beta Treated_i Post_t + \gamma Treated_i + \lambda Post_t + \alpha + Trend_t + Trend_t Treated_i + \varepsilon_{i,t} \tag{2}$$

In a third specification, we explore the monthly frequency of our data by estimating a dynamic difference-in-difference, see equation (3):

$$Y_{i,t} = \sum_{K=-3}^{K=-2} \beta^{lag} D_{i,t}^{K} + \sum_{L=1}^{L=3} \beta^{leads} D_{i,t}^{L} + \gamma_i + \lambda_t + \epsilon_{i,t} \tag{3}$$

where $D_{i,t} = Treated_i \times 1\{t = k\}$ captures the treatment effect in each moment of time, in comparison with the first month before the event. In this case, the parallels trends hypothesis is infered from the significance of the β^{lag} coefficient and treatment effect by the significance of β^{lag} coefficients. Time fixed effect that absorbs the shocks that are common

[12] In their paper, the main effect of luck happens in the three months after the lucky event

to all investors are given by λ_t, and γ_i is a fixed effect that captures the time-invariant characteristics of the investors.

We would like to compare identical investors except that one group of investors receive the treatment and the other not. In the diff-and-diff specifications of (1) and (2), all the investor level characteristics that are common at the group level (and fixed over time) are taken into account when we take the differences over time. In equation (3), we directly include investor fixed effects to control for fixed characteristics at the investor level. The specifications above are able to capture differences in the group of investors that are fixed over time. They do not capture, however, investor-level characteristics that are varying over time and that could pollute our findings. To address this concern, we define a tighter control group that only includes investors that are very similar to those in the treatment group using a propensity score model.

In this model, we include a number of investor-level characteristics such as age, gender, and experience and, importantly, a proxy that capture the investors ex-ante preference for Vale stock. After estimating the propensity score, treated and control investors are matched using the nearest neighbor method. To improve matching quality, we impose a maximum of 5% in the distance between the propensity score. Moreover, the match is done 1-to-1 and without replacement.[13] Because of the tight imposition match, the sample size decreased from 501 treated investors to 418 and control from 7896 to 418.

The matching procedure allows the construction of a much more similar control group. The only difference is that treated sold Vale's stocks, and control sold other stock. Therefore, the identification strategy relies on matching[14] to balance covariates and a difference-in-difference to estimate the effect of luck on trading behavior.

Table 2 reports the quality of matching procedure. Before the match, treated and control investors are statistically different in many covariates, especially in the trades and volumes during the identification period. And the similarity of both groups remains in the age, the proportion of males, and the number of trades made in the previous 12 months, indicating that both groups were active in the market. Figure 4 shows the matching quality graphically. The plot reports the standardized difference between groups. Before the match, many covariates have a higher than 10% difference, as indicating by the points outside the black

[13]Additionally, we considered matching with replacement. In this case, we use the match fixed effect to capture the variation within the match. The results are reported in the appendix.

[14]Also, as robustness to the selection of sample by matching we estimate the specifications 1 and 2 using the inverse of propensity score as weights. However, this procedure generates extreme weights due to some units with lower propensity score estimation. To avoid this, we trimmed the sample using the procedure suggested by Imbens and Rubin (2015). After the trimming, we verified the covariance balance between the treated and control groups and compared the results with the complete and trimmed samples. The results are robust to this exercise and shown in the appendix.

dashed lines. After the match, none of the covariates are statistically significant, and all are in the 10% range of standard difference. Also, Table 2 shows the p-values for the difference between groups, which are all above 10%.

Thus, any time-varying shock correlated with the treatment because the difference in those covariates after the match have similar effects in treated and control. This similarity gives strength to the identification strategy. Since, conditional on the matched sample, treatment works as well as a randomly assigned. Therefore, our identification hypothesis is the absence of time-varying shock correlated with treatment after the matching procedure.

The inference needs to take into account the sample selection due to match. This previous procedure creates a dependence between the matched treated and control investor. Abadie and Spiess (2021) showed that if the model is corrected specified, this dependence is entirely captured in the model, such as the residuals are correct to inference. However, misspecification of the model might lead to severe biases inference. Hence, we follow Abadie and Spiess (2021) and correct this concern by using the cluster at the match level, which allows for the correlation between treated and control investors due to match. Since it is a 1-to-1 match, the cluster is at the pair's level. Moreover, we also use clusters on the investor level to correct the auto-correlation.

Finally, we employ three robustness exercises to verify the plausibility from our estimations. First, we identify as "treated" investors that sold Petrobras stock before the disaster. Second, we address the possibility that our results are due spurious relation by randomly selecting 501 investors of the control group and assigning them as treated investors. Third, we estimate the effect using data from the previous years.

4 Results

This section is divided into three parts.[15] The first part shows the results from a static estimation, equations 1 and 2. The second part reports the results from the dynamic diff-in-diff estimation in an event study plot. Third, represent the robustness tests.

4.1 Static Results

The central hypothesis is that luck triggers a sensation in the investors, leading them to increase their trading activity. Table 3 shows the results of the static estimation for three outcomes portfolio turnover, numbers of trades, and days with at least one trade. The

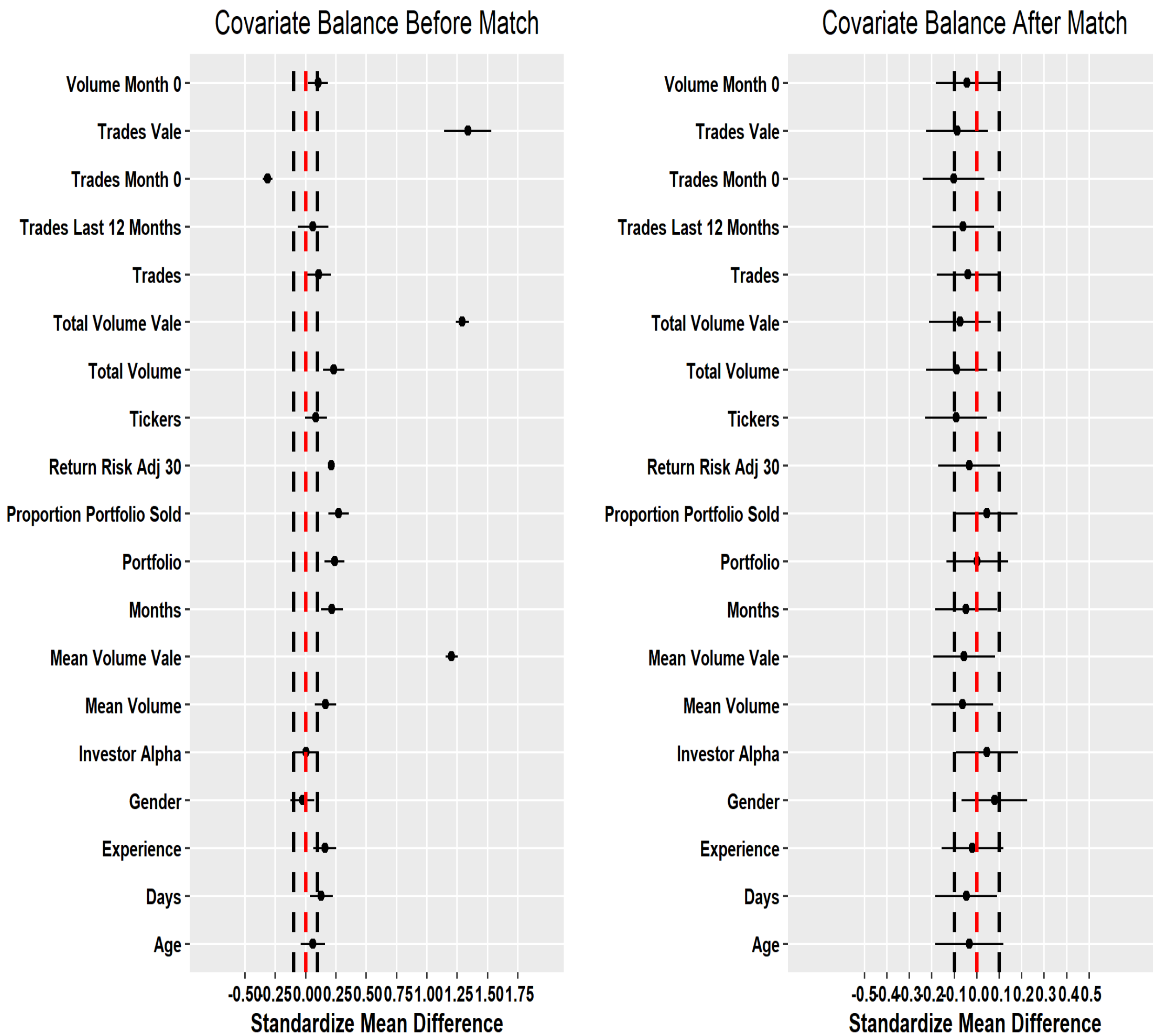
[15] The appendix contains the results of matching with replacement and weighting by the inverse of the propensity score instead of sample selection procedure.

Table 1: Descriptive summary

Variable	Treated ($N_t = 501$) Mean	SD	Median	Pct25	Pct75	Control ($N_c = 7568$) Mean	SD	Median	Pct25	Pct75
Panel A: Trading Activity										
# Days	46.918	56.493	29.000	15.000	55.000	39.858	53.144	24.000	12.000	46.000
# Months	17.655	10.638	16.000	8.000	25.000	15.304	10.678	12.000	7.000	22.000
Portfolio (1000$RS)	296.772	1118.172	31.852	8.402	135.749	277.872	2770.472	19.592	3.679	88.341
# Tickers	9.970	10.282	7.000	3.000	12.000	9.087	10.220	6.000	3.000	12.000
# Trades	66.613	116.660	35.000	17.000	70.000	54.956	103.916	28.000	14.000	57.000
# Trades Last 12 Months	20.723	45.505	11.000	5.000	21.000	18.682	30.570	11.000	5.000	21.000
# Trades Vale	14.259	19.198	7.000	3.000	18.000	2.543	6.982	0.000	0.000	2.000
Mean Volume (1000$RS)	36.749	81.499	11.946	4.851	33.280	40.324	260.379	10.361	4.155	23.614
Mean Volume Vale (1000$RS)	43.248	104.504	12.477	4.571	30.970	18.320	115.956	0.001	0.001	6.326
Total Volume (1000$RS)	4229.619	16023.831	425.307	129.554	1431.303	2534.138	15196.506	299.948	81.967	1060.792
Total Volume Vale (1000$RS)	1117.079	4557.229	91.616	19.042	399.243	190.499	2375.219	0.001	0.001	19.561
Alpha %	-0.013	0.029	-0.013	-0.027	0.003	-0.013	0.024	-0.013	-0.021	-0.006
Return Risk Adj 30 days ahead %	-0.237	1.307	0.000	-0.385	0.000	-0.894	4.375	-0.279	-1.415	0.143
Fee	276.388	2020.902	0.000	0.000	108.975	409.910	7080.870	28.998	0.000	130.656
Panel B: Trading Activity in Identification Interval										
Volume Month (1000$RS)	59.185	571.591	9.025	3.239	19.525	70.420	1035.341	8.801	3.176	19.735
Trades Month	1.114	0.411	1.000	1.000	1.000	1.441	1.078	1.000	1.000	2.000
Proportion Portfolio Sold %	-0.363	0.339	-0.229	-0.549	-0.094	-0.466	0.375	-0.343	-0.899	-0.122
Panel C: Personal Characteristics										
Age	47.951	15.055	45.000	36.000	58.000	47.013	15.178	45.000	35.000	57.000
Gender	0.854	0.353	1.000	1.000	1.000	0.862	0.345	1.000	1.000	1.000
Experience	5.824	4.617	6.000	2.000	8.000	5.098	4.506	4.000	1.000	8.000
Panel D: Outcomes										
Turnover %	0.208	0.398	0.095	0.022	0.228	0.334	0.504	0.167	0.066	0.405
Turnover Without Vale %	0.115	0.294	0.022	0.000	0.112	0.315	0.469	0.166	0.064	0.384
# Trades	1.193	2.143	0.667	0.167	1.333	1.463	2.293	0.833	0.333	1.667
# Trades Without Vale	0.716	1.650	0.167	0.000	0.833	1.396	2.170	0.833	0.333	1.667
Portfolio Return %	-0.082	0.091	-0.088	-0.132	-0.038	-0.069	0.066	-0.062	-0.109	-0.033
Portfolio Return After Fee %	-0.087	0.091	-0.093	-0.137	-0.040	-0.073	0.063	-0.063	-0.111	-0.034
Risk Adjusted Return %	-0.053	0.071	-0.055	-0.091	-0.022	-0.040	0.051	-0.035	-0.067	-0.012
Risk Adjusted Return After Fee %	-0.053	0.071	-0.055	-0.092	-0.020	-0.041	0.051	-0.036	-0.068	-0.013
# Days With Trade	0.876	1.108	0.500	0.167	1.167	1.079	1.172	0.667	0.333	1.333
# Days With Trade Without Vale	0.509	0.911	0.167	0.000	0.667	1.047	1.137	0.667	0.333	1.333

Notes: This table reports the descriptive summary of the cross-section distribution. In panel A: Days is the number of days, months the number of months with at least one trade, portfolio is the value of portfolio in thousand of R$. Ticker is the number of different stock trades. Investor's alpha is average of the constant coefficient of a Fama-French four factor on portfolio return for each investor using 36 months prior the event. Risk adjusted 30 days ahead is the return of the stock on investor portfolio in the period t, evaluated at the next month, t+1. Panel B captures trading activity in the 21 days before the event. Proportion of portfolio sold is the volume sold divided by the initial portfolio value in the interval. Panel C contains the personal characteristics. Panel D reports the outcomes during the period used in the estimations, three months before and three months after the event.

Figure 4: Covariate Balance Before and after Matching



Note: This figure shows the covariates balance. Before the match, there were 501 treated investors and 7568 controls. After matching, there are 418 in both groups. All variables are normalized. Thus the difference is measured in terms of standard deviations. The red dashed line is 0 and the black dashed lines are 0.1 and -0.1

Table 2: Covariate Balance

	Before Match							After Match						
	Treated		Control		Difference			Treated		Control		Difference		
Variables	Mean	Std.Err	Mean	Std.Err	Mean	Std Error	P-value	Mean	Std.Err	Mean	Std.Err	Mean	Std Error	P-value
# Days	46.92	56.54	39.86	53.15	7.06	2.60	0.007***	44.88	56.47	47.36	53.07	-2.48	3.84	0.517
# Months	17.66	10.65	15.30	10.68	2.35	0.49	0***	16.55	10.32	17.03	10.79	-0.48	0.74	0.516
Log Portfolio	10.26	2.45	9.60	2.74	0.66	0.11	0***	10.03	2.49	10.02	2.60	0.01	0.18	0.965
Different Tickers	9.97	10.29	9.09	10.22	0.88	0.47	0.062*	10.75	10.82	11.79	12.22	-1.04	0.81	0.198
# Trades	66.61	116.76	54.96	103.92	11.66	5.35	0.029**	65.01	120.55	69.22	96.53	-4.21	7.65	0.582
# Trades Last 12 Months	20.72	45.54	18.68	30.57	2.04	2.06	0.323	20.54	46.86	22.99	36.19	-2.45	2.93	0.403
# Trades Vale	14.26	19.21	2.54	6.98	11.72	0.86	0***	9.95	12.59	11.20	15.81	-1.25	1.00	0.214
Log Mean Volume	9.51	1.35	9.28	1.40	0.23	0.06	0***	9.43	1.35	9.52	1.46	-0.09	0.10	0.372
Log Mean Volume Vale	9.20	2.36	3.45	4.67	5.75	0.12	0***	9.03	2.49	9.17	2.72	-0.14	0.18	0.43
Log Total Volume	13.08	1.98	12.60	2.04	0.48	0.09	0***	12.95	1.97	13.13	2.13	-0.18	0.14	0.209
Total Volume Vale	11.16	3.12	3.96	5.40	7.20	0.15	0***	10.71	3.17	10.95	3.42	-0.24	0.23	0.295
Investor's Alpha %	-0.01	0.03	-0.01	0.02	0.00	0.00	0.932	-0.01	0.03	-0.01	0.03	0.00	0.00	0.5
Risk Adjusted 30 Days Ahead %	-0.02	0.13	-0.36	1.59	0.33	0.02	0***	-0.03	0.13	-0.02	0.58	-0.01	0.03	0.644
Log Volume Month 0	9.24	1.41	9.07	1.60	0.17	0.06	0.011**	9.20	1.41	9.27	1.63	-0.07	0.11	0.539
# Trades Month 0	1.11	0.41	1.44	1.08	-0.33	0.02	0***	1.12	0.44	1.16	0.43	-0.04	0.03	0.147
Proportion Portfolio Sold %	-0.36	0.34	-0.47	0.38	0.10	0.02	0***	-0.40	0.35	-0.41	0.36	0.02	0.03	0.515
Age	47.95	15.07	47.01	15.18	0.94	0.77	0.224	47.50	14.91	47.97	14.35	-0.47	1.13	0.68
Gender	0.85	0.35	0.86	0.34	-0.01	0.02	0.647	0.88	0.32	0.86	0.35	0.03	0.03	0.282
Experience	5.82	4.62	5.10	4.51	0.73	0.21	0.001***	5.59	4.50	5.67	4.72	-0.09	0.32	0.791

Notes: This table reports the covariate balance between treated and control before and after matching. There are 501 treated investors and 7568 control investors before the match and 418 treated and controls after matching. Days are the number of days, months the number of months with trades from the beginning of the data set to one month before the event. Log of Portfolio is the average of the log of the portfolio during the same periods. Ticker is the average number of different stock trades. Volume variables are averages in using the same period, the same occurs for trades. Investor's alpha is the average of the constant coefficient of a Fama-French four-factor on portfolio return for each investor used 36 months prior to the event. Risk-adjusted 30 days ahead is the return of the stock on the investor portfolio in the period t, evaluated at the next month, t+1. Volume and Trades in month 0 capture the trading activity in the 21 days before the event, and the proportion of portfolio sold is the volume sold divided by the initial portfolio value during the same period. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

coefficient of $Trated_i \times Post_t$ captures the luck effect on trading activity. Columns 1 to 6 show the estimations result before controlling for covariates (unmatched sample). Columns 7 to 12 contain the results after controlling for covariates (matched sample). Consistent with the hypothesis, all outcomes show a positive coefficient. It is worth noting that treated investors have a lower trading level given by the coefficient of $Treated_i$. Moreover, after the event, the trading activity is a lower coefficient $Post_t$.

Remarkably, lucky investors, on average, increase their portfolio turnover by 11% in the three months after the event. For comparison, Gao et al. (2021) shows that investors who won an IPO allotment lottery in China, which is a stroke of good luck because the lottery is under-priced, have an increase in portfolio turnover by 28% in the three months after the luck event. However, the average portfolio turnover was approximately 365%, which gives a relative increase of 7%. Our results are economically more significant since the average portfolio of the control group before the event for the control group was 35%, which means a approximately increase of 31%

Furthermore, Anagol et al. (2021) also explores winning IPO allotment lotteries in India. They show that when the stock increase in value after the IPO, lucky investors increase their trading volume by 4,3%, the number of trades by 1,5%, and the likelihood of participation in subsequent IPOs by 0,2% three months after the event.

More than the impact on portfolio turnover, there is a significant increase in the number of trades. Table 3 column 9 indicates an increase by 0.75. The effect is economically meaningful compared with the intercept coefficient of 1.64, which gives the average mean of control investors before the event. Thus, the number of trades shows a relative increase around of 46%.

Moreover, the increase in trades is associated with investors spending more days making trades. The number of days with trades increase by 0.42, in percentage, around of 35%. Sensation-seeking or gambling activity could be associated with this behavior. Gao et al. (2021) also found that lucky shock is associated with investor increasing their portfolio volatility, both idiosyncratic and systematic, and an expansion in the number of lottery stocks in their portfolio.

Table 3: Trade Activity without Vale

	Unmatched Sample						Matched Sample					
Dependent Variables:	Turnover		Trades		Days With Trades		Turnover		Trades		Days With Trades	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables												
treated × post	0.1128***	0.1508***	0.6180***	0.8382***	0.3499***	0.5625***	0.1101***	0.1821***	0.7353***	0.9559***	0.4257***	0.6021***
	(0.0252)	(0.0369)	(0.1243)	(0.1675)	(0.0617)	(0.0954)	(0.0384)	(0.0662)	(0.1935)	(0.2776)	(0.1006)	(0.1639)
treated	-0.2566***	-0.2313***	-0.9890***	-0.8421***	-0.7129***	-0.5711***	-0.2109***	-0.1629***	-0.9632***	-0.8162***	-0.7100***	-0.5923***
	(0.0150)	(0.0205)	(0.0795)	(0.1264)	(0.0491)	(0.0691)	(0.0259)	(0.0419)	(0.1431)	(0.2507)	(0.0855)	(0.1301)
post	-0.0836***	-0.1550***	-0.5302***	-0.8866***	-0.3399***	-0.6468***	-0.0657**	-0.1758***	-0.5580***	-0.9091***	-0.3725***	-0.6575***
	(0.0082)	(0.0134)	(0.0403)	(0.0545)	(0.0186)	(0.0303)	(0.0266)	(0.0516)	(0.1360)	(0.2047)	(0.0771)	(0.1257)
(Intercept)	0.3568***	0.3092***	1.661***	1.423***	1.217***	1.012***	0.3179***	0.2445***	1.641***	1.406***	1.222***	1.032***
	(0.0074)	(0.0108)	(0.0350)	(0.0494)	(0.0175)	(0.0245)	(0.0225)	(0.0363)	(0.1211)	(0.2146)	(0.0702)	(0.1105)
trend		0.0238***		0.1188***		0.1023***		0.0367**		0.1170		0.0950**
		(0.0037)		(0.0158)		(0.0086)		(0.0155)		(0.0750)		(0.0394)
trend × treated		-0.0126		-0.0734		-0.0709***		-0.0240		-0.0735		-0.0588
		(0.0090)		(0.0451)		(0.0257)		(0.0187)		(0.0894)		(0.0479)
Fit statistics												
Adjusted R^2	0.00697	0.00756	0.00793	0.00865	0.01407	0.01614	0.01861	0.01962	0.01630	0.01656	0.03125	0.03224
R^2	0.00703	0.00766	0.00799	0.00875	0.01413	0.01625	0.01921	0.02062	0.01690	0.01757	0.03185	0.03323
Observations	48,414	48,414	48,414	48,414	48,414	48,414	4,896	4,896	4,896	4,896	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2). All columns with unmatched sample are clustered at the individual level. All columns with matched sample uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

There is extensive literature relating excessive trading with lower returns. For example, Barber and Odean (2000) founds that investors who trade more have an annual performance of 11.4 percent as opposed to 17.9 of market returns. In a more recent study, Barber et al. (2009) estimates a penalty for excessive trading is 3.8 percentage points in annual performance. Accordingly, our second hypothesis is that investors impacted by luck would experience lower performance because the increase in trading activity is not supported by superior ability, and the negative performance compounds to produce even lower performance. Table 4 reports the static effect of luck in portfolio performance before transaction costs are considered. Consistent with the hypothesis, lucky investors have a 6% lower portfolio returns than control investors after the event. Moreover, columns 5 to 8 show that the finding is robust to using portfolio risk-adjusted [16]; they have 4% lower risk-adjusted returns.

Table 4: Return Variables without Vale

Dependent Variables:	Portfolio Return				Portfolio Risk Adjusted Return			
	Unmatched Sample		Matched Sample		Unmatched Sample		Matched Sample	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0656***	-0.0814***	-0.0601***	-0.0923***	-0.0449***	-0.0752***	-0.0383***	-0.0758***
	(0.0066)	(0.0167)	(0.0097)	(0.0253)	(0.0062)	(0.0159)	(0.0093)	(0.0238)
treated	0.0188***	0.0083	0.0232***	0.0017	0.0094**	-0.0108	0.0158**	-0.0092
	(0.0050)	(0.0101)	(0.0076)	(0.0157)	(0.0040)	(0.0088)	(0.0062)	(0.0135)
post	0.0109***	0.0345***	0.0119*	0.0635***	0.0432***	0.0200***	0.0447***	0.0393**
	(0.0013)	(0.0030)	(0.0064)	(0.0166)	(0.0013)	(0.0029)	(0.0061)	(0.0157)
(Intercept)	-0.0711***	-0.0554***	-0.0712***	-0.0368***	-0.0614***	-0.0769***	-0.0639***	-0.0675***
	(0.0010)	(0.0020)	(0.0052)	(0.0113)	(0.0008)	(0.0017)	(0.0043)	(0.0097)
trend		-0.0079***		-0.0172***		0.0077***		0.0018
		(0.0009)		(0.0049)		(0.0008)		(0.0045)
trend × treated		0.0053		0.0108		0.0101**		0.0125*
		(0.0046)		(0.0072)		(0.0044)		(0.0067)
Fit statistics								
Adjusted R^2	0.00361	0.00523	0.00837	0.01113	0.02216	0.02463	0.00815	0.01010
R^2	0.00367	0.00533	0.00898	0.01214	0.02222	0.02473	0.00876	0.01111
Observations	48,414	48,414	4,896	4,896	48,414	48,414	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2). All columns with unmatched sample are clustered at the individual level. All columns with matched sample uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

[16]Risk adjusted returns are calculated as residuals of the 4-factor Fama-French estimation in the investors' portfolio return

Table 5 reports the effect of luck on portfolio performance after accounting for transactions cost. Lucky investors present a 6.14% lower returns not risk-adjusted and 3.94% for risk-adjusted. Thus, comparing the table 4 and 5, transactions cost implies in a reduction of 13 basis points and 9 basis points on returns, not adjusted and adjusted, respectively.[17]

Importantly, luck potentially has two mechanisms implying a lower return. First reduces the investor's ability. Second, it induces more trades that do not cover the fees. In the literature, both effects seem to be present, Gao et al. (2021) reports that their lucky investors have 1.3% lower returns, and the security selection ability decreases by 0.2%. Our results are consistent with both mechanisms. Since differences in the investor's ability (portfolio's alpha and stock-picking ability) are controlled in the matched sample. Thus, the lucky effect decreasing performance before the transaction costs indicates that investors decrease their ability. And the lower performance after transaction costs means that more fees are paid to the detriment of returns.

[17]This slight difference can be due to underestimating transaction costs. Since it does not include taxes paid by investors. Furthermore, we do not have deal-level data. We only observe the aggregate transaction of each stock by an investor in a day. Thus, if there is more than one transaction in the same stock the fixed cost is underestimated.

Table 5: Return Variables without Vale After Fee

Dependent Variables:	Portfolio Return Aft Fee				Portfolio Risk Adjusted Return Aft Fee			
	Unmatched Sample		Matched Sample		Unmatched Sample		Matched Sample	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0668***	-0.0831***	-0.0614***	-0.0942***	-0.0458***	-0.0769***	-0.0394***	-0.0772***
	(0.0065)	(0.0167)	(0.0097)	(0.0253)	(0.0062)	(0.0158)	(0.0093)	(0.0237)
treated	0.0212***	0.0103	0.0251***	0.0033	0.0107***	-0.0101	0.0168***	-0.0085
	(0.0049)	(0.0101)	(0.0075)	(0.0156)	(0.0040)	(0.0088)	(0.0061)	(0.0134)
post	0.0121***	0.0367***	0.0131**	0.0656***	0.0440***	0.0220***	0.0455***	0.0409***
	(0.0013)	(0.0029)	(0.0064)	(0.0166)	(0.0013)	(0.0029)	(0.0061)	(0.0157)
(Intercept)	-0.0756***	-0.0592***	-0.0752***	-0.0402***	-0.0628***	-0.0775***	-0.0649***	-0.0681***
	(0.0010)	(0.0019)	(0.0052)	(0.0112)	(0.0008)	(0.0017)	(0.0043)	(0.0097)
trend		-0.0082***		-0.0175***		0.0073***		0.0016
		(0.0009)		(0.0049)		(0.0008)		(0.0045)
trend × treated		0.0054		0.0109		0.0104**		0.0126*
		(0.0046)		(0.0072)		(0.0044)		(0.0067)
Fit statistics								
Adjusted R^2	0.00379	0.00556	0.00850	0.01141	0.02303	0.02532	0.00847	0.01038
R^2	0.00385	0.00567	0.00911	0.01242	0.02310	0.02542	0.00908	0.01140
Observations	48,414	48,414	4,896	4,896	48,414	48,414	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2). All columns with unmatched sample are clustered at the individual level. All columns with matched sample uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

4.2 Dynamic Results

This section reports the results from dynamic diff-in-diff estimation equation 3. This specification explores the effect of luck over time while using fixed effects of investor and time. To identify treated and controlled investors, they must have traded in the month (21 trading days) before the disaster. This month has a higher trading activity for both groups compared to others. Thus is not a clear benchmark of the trading activity in regular months. Consequently, using this month as a reference in the dynamic diff-in-diff strategy would result in an polluted estimation of the effects.[18] Then we use as a reference in the event study plots, Figure 5 to Figure 9, the month before the disaster.

[18] In Gao et al. (2021) some investors can potentially have a higher trading volume in the previous month before to IPO allotment lotteries. Thus to avoid this problem, they defined the pre-treatment period as months -2 to -4 before the lottery. We also exclude the month before the event but control for trading activity in this month in matched sample.

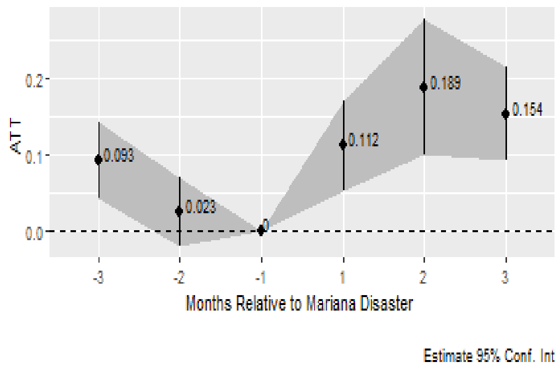
The left side of these plots reports the estimation using the unmatched sample and the right side the matched sample. Trading activities, portfolio turnover Figure 5b, number of trades Figure 6b and number of days with trade Figure 7b in the unmatched sample presents an violation in the parallel hypothesis. However, the parallel trends are satisfied using the matched sample, where the distribution of covariates is balanced among treated and control investors, see Figure 4. Also, worth noting that this hypothesis is satisfied for return variables Figure 8 and Figure 9 using both samples.

Figure 5b shows that one month after the luck shock, turnover is 13% higher for lucky investors compared to unlucky investors one month before the shock, increases to almost 20% in the second month, and returns to 12% in the third month. The effect is statistically significant even after three months. Number of trades and days with trades reports almost the same pattern on the Figure 6b and Figure 7b, respectively. In the literature, Gao et al. (2021) reports that luck has a medium-term impact that decreases over time. Luckier investors increase their portfolio turnover by 28%, 9%, and 3% after three, six, and nine months subsequent the event. Although all impact is statistically positive, the last two are not economically large once the median portfolio turnover in their sample is 180%. A similar pattern is reported by Anagol et al. (2021). Winning an IPO lottery in India generates an increase in transaction volume relative to the portfolio value of 3% and 2% in the three and six months subsequent to the event. At the same interval the number of trades increases by 1% and 7%.

Figure 8b reports the effect of luck on portfolio return after fees. The dynamic shows a large negative effect in the first month after the lucky shock. Moreover, this effect monotonically decreases to zero in the third month. Figure 9b shows the dynamic effect for risk-adjusted returns. Using this measure, the lucky investors have a negative shock one month after the event but a lower effect in the second month that remains almost the same in the third month. To summarize, the results indicate that luck positively impacts trading activity while is also associated with worsening returns, and the effect remains statistically significant in the three months following the luck shock

Figure 5: Dynamic Effect on Turnover

(a) Before Match



(b) After Match

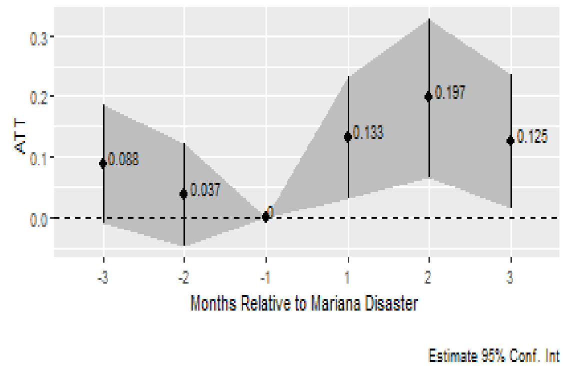
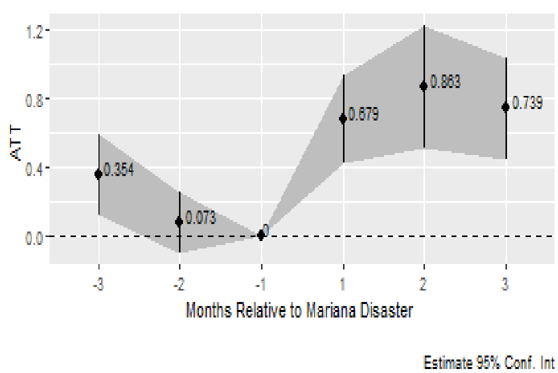


Figure 6: Dynamic Effect on Number of Trades

(a) After Match



(b) Before Match

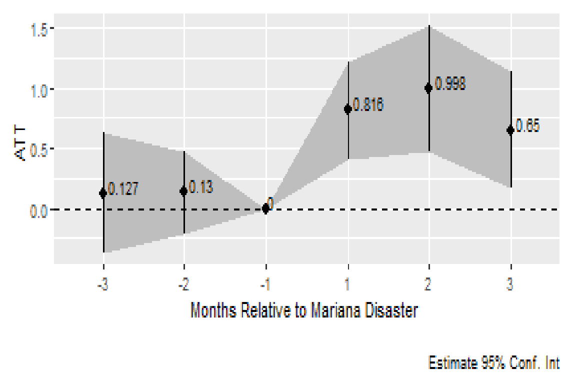
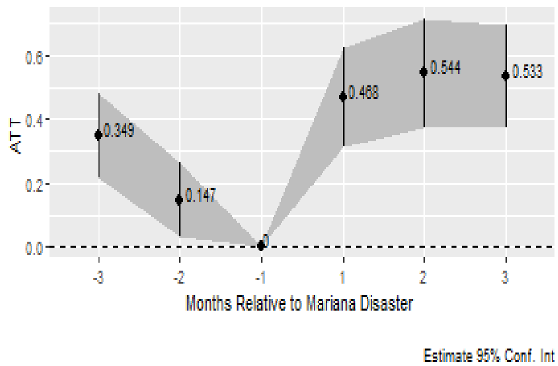


Figure 7: Dynamic Effect on Number of Days with Trades

(a) Before Match



(b) After Match

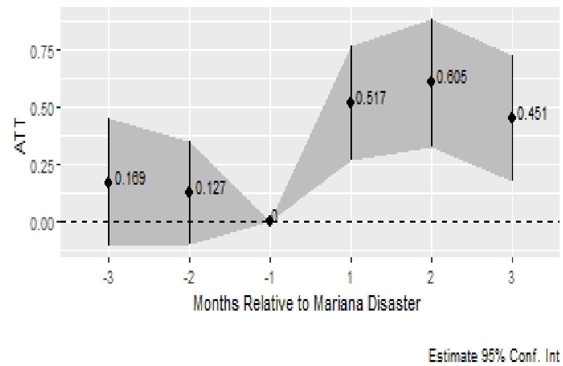
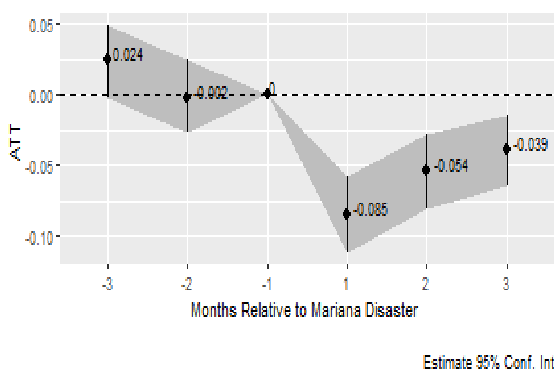


Figure 8: Dynamic Effect on Number of Portfolio Return

(a) Before Match



(b) After Match

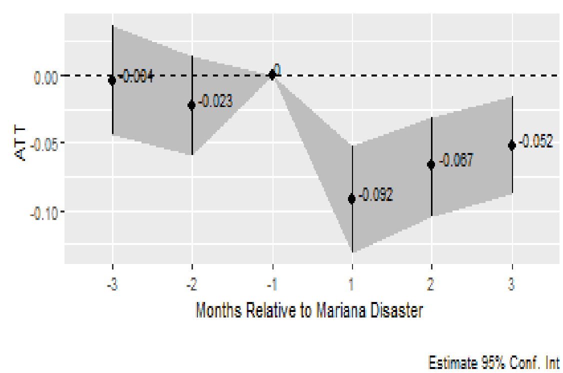
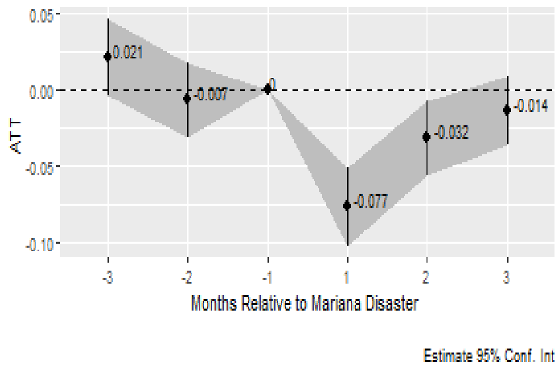
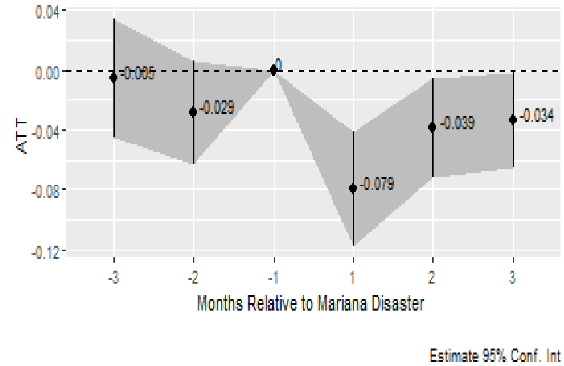


Figure 9: Dynamic Effect on Abnormal Return

(a) Before Match



(b) After Match

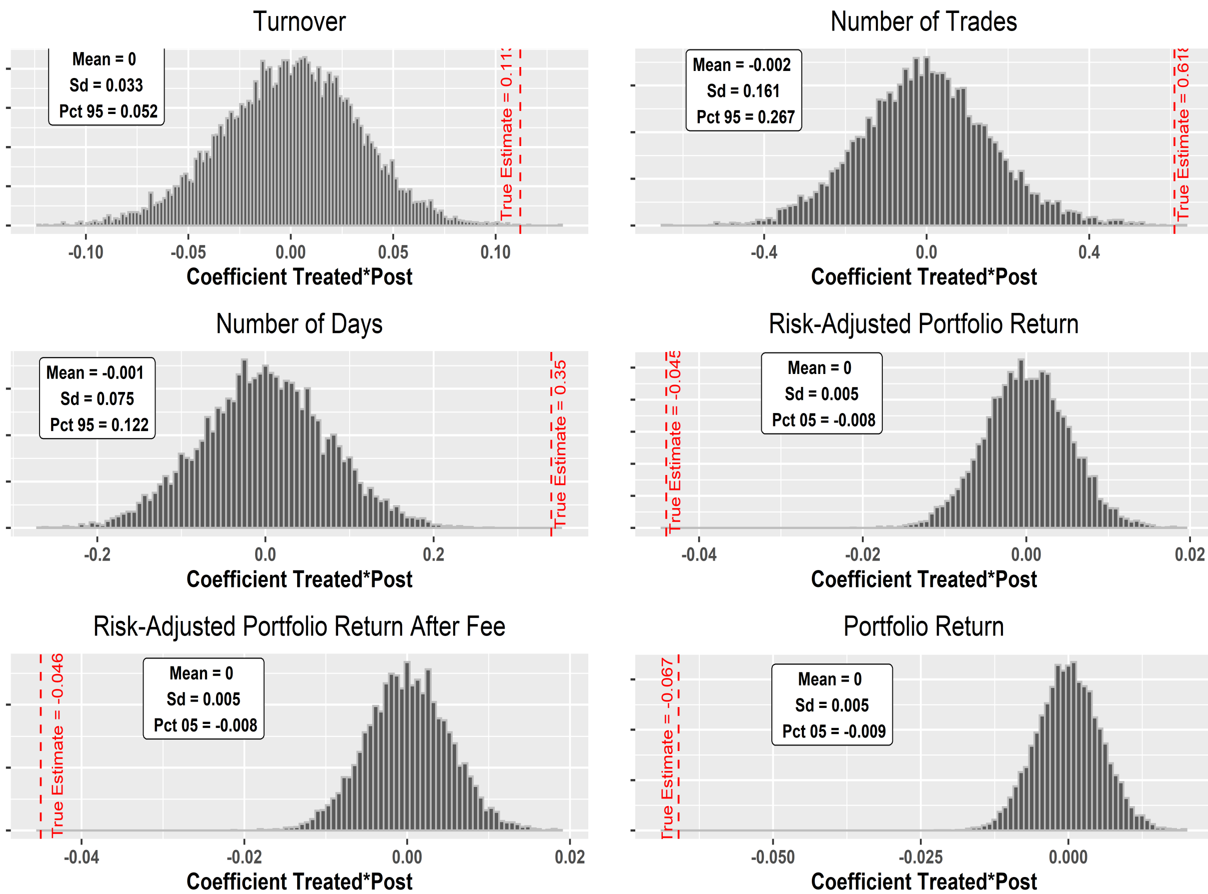


5 Robustness

5.1 Pseudo-treatment

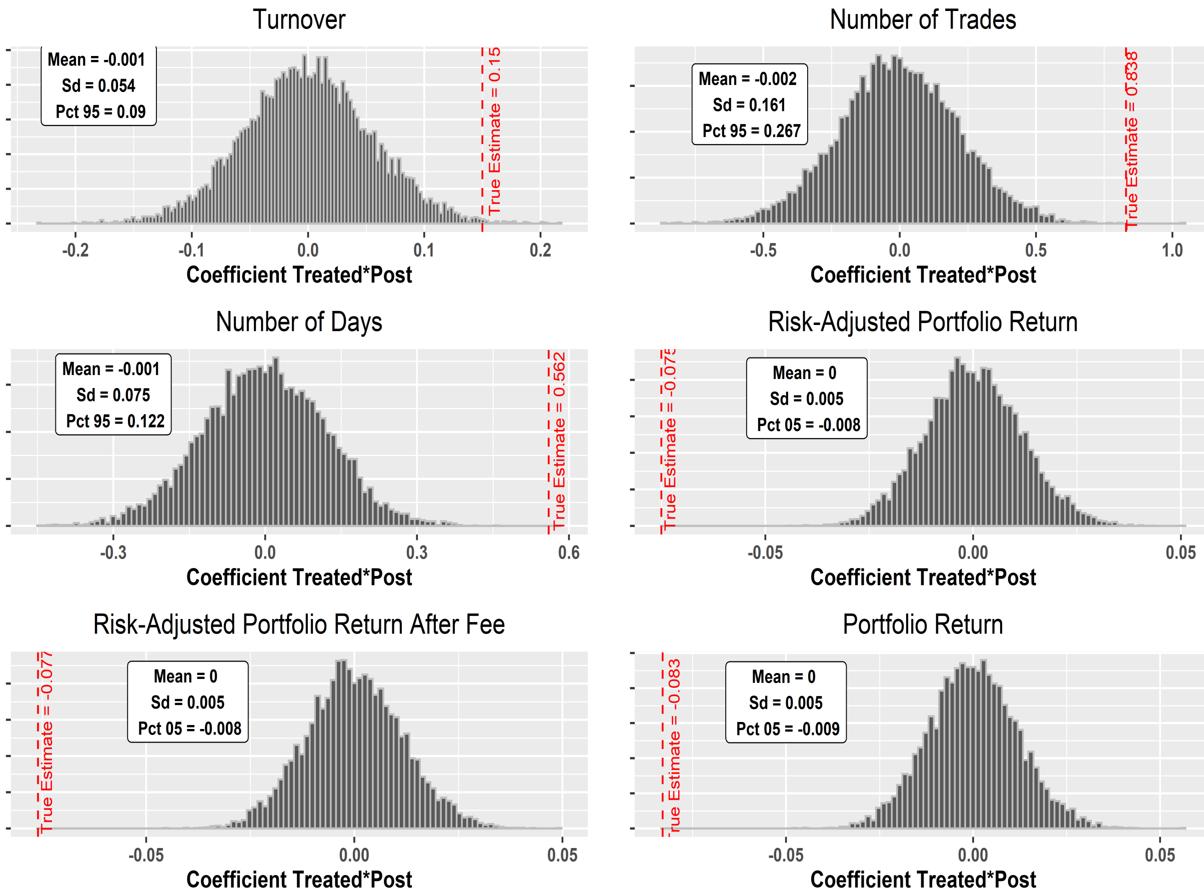
The Figure 1a and Figure 1b show that individuals in the control group decrease their trading activity after the event, consistent with the overall market Figures 3a and 3b, while individuals in the treated group maintain their trading level. Therefore, a possible concern is that our treatment effect comes to a favorable draw of investors that maintains the same activity trading whether the market is decreasing. To address this concerns, we randomly selected from the control group 501 individuals and assigned them a "pseudo-treated" group and kept the remaining 7067 individuals in the control group. We then estimate equation 1 and equation2, repeat the same process 10000 times, and construct the distribution of the parameter of interest. Figure 10 and Figure 10 plots the histogram of this distribution and compares them to the original coefficients given by the horizontal red line. For turnover, there is less than a 5% probability of the pseudo-treatment being large than the original estimate. For the number of trades and days with trade, the histograms indicate a probability lower than 1% of the times. Finally, the histogram of variables returns shows that original coefficients are much less likely to occur. This placebo indicates that our results are due to the investors who get luck by selling Vale's stock before the large disaster that devaluated the stock's value.

Figure 10: Pseudo-Treatment



Note: This figure reports the distribution of the coefficients estimated in the equation 1 using control investor as a treated investors. The line in red contains the value of the original effect.

Figure 11: Pseudo-Treatment



Note: This figure reports the distribution of the coefficients estimated in the equation 2 using control investor as a treated investors. The line in red contains the value of the original effect.

5.2 Past-Outcomes

Regress the treatment in past outcomes is often recommended as a placebo test Imbens and Rubin (2015). Because past outcomes are fixed when treatment is assigned, we should not find any effect. Since our main strategy is a difference-in-difference, equation 1 and equation 2, we estimate the difference in trading behavior using data from previous years before the event, maintaining the same treated and control groups fixed. Moreover, considering our identification strategy relies on matching, this placebo focuses on the matched sample. Table 6 reports the results of those equations for all outcomes, trading, and return variables. For brevity, only the coefficient of interest is reported. The table is divided into three panels, using data from one to three years before the disaster. As expected, there was no treatment effect one year before, Panel A. No trading activity outcomes are significant in Panel B, but performance returns are positive at 5%, the opposite signal of what was expected. However, this significance is not robust using data from three years before, Panel C. To conclude, this placebo test rules out explanations that our results are driven by any previous difference in trading patterns between treated and control groups. Which strengthens the argument that luck is the treatment affecting the trading and performance outcomes only after the event.

Table 6: Placebo in Time

Dependent Variables:	Turnover		Trades		Days with Trade		Portfolio Return		Return Portfolio Aft Fee	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Panel A: One Year Before										
Variables										
treated × post	-0.0494	0.0414	0.3489	0.8672	-0.0874	0.1773	-0.0083	0.0120	-0.0077	0.0115
	(0.0399)	(0.0614)	(0.4334)	(0.6770)	(0.1237)	(0.1828)	(0.0080)	(0.0187)	(0.0079)	(0.0186)
Fit statistics										
Adjusted R^2	0.00248	0.00250	0.00292	0.00272	0.00261	0.00257	0.03730	0.06865	0.03844	0.07008
Panel B: Two Years Before										
treated × post	0.0095	0.0672	0.1087	0.0259	0.0940	0.0921	-0.0064	0.0390**	-0.0067	0.0387**
	(0.0263)	(0.0544)	(0.2209)	(0.3619)	(0.0955)	(0.1596)	(0.0081)	(0.0194)	(0.0081)	(0.0193)
Fit statistics										
Adjusted R^2	0.00031	0.00020	-0.00010	-0.00045	1.96×10^{-5}	-0.00035	0.05109	0.05273	0.05113	0.05278
Panel C: Three Years Before										
treated × post	-7.57×10^{-5}	-0.0072	0.0964	-0.5359	0.1062	-0.1806	-0.0010	0.0275	-0.0014	0.0276
	(0.0338)	(0.0476)	(0.3605)	(0.4366)	(0.1025)	(0.1318)	(0.0075)	(0.0190)	(0.0075)	(0.0190)
Fit statistics										
Adjusted R^2	-0.00020	-0.00010	4.61×10^{-5}	1.78×10^{-5}	-0.00044	1.39×10^{-6}	0.00040	0.02082	0.00033	0.02026
Observations	4,896	4,896	4,896	4,896	4,896	4,896	4,896	4,896	4,896	4,896

Note: This table reports the placebo estimation from the models 1 columns (1) and 2 columns (2). All columns uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Panel A: reports the estimation using the period t-12. Panel B: reports the estimation using the period t-24. And panel C:reports the estimation using the periods t-36. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

5.3 Placebo Firm

In this section, we considered treated investors as those who sold Petrobras stocks instead of Vale stocks one month (21 trading days) before the Mariana disaster. Since Petrobras stocks do not devalued after the event, these newly treated investors do not get luck by selling the stock before the event. Therefore it is expected not to find any difference in trading activity between those investors and the control groups.

Petrobras and Vale have similarities that allow a better comparison. Both are salient companies among Brazilians, especially among retail investors. It is important to control this feature because of the extensive evidence that salience is essential in deciding which stocks to trade.[19] Second, both firms are commodities producers and are large companies relative to the Brazilian market. Therefore, the stocks are closer concerning their value and size risks. Hence, investors of those companies are more likely to be similar in their financial investments.

Table 7 reports the effect defining treated investor using Petrobras. As expected, none of the trading activity variables have a statistically significant coefficient. Except for days with trades, which are negative, a signal opposite of that was expected. Table 8 presents the results for returns variables. Column (2) shows a positive effect, as opposed to the expected. Therefore, using investors that sold stocks of a company that is salient to retail investors, is large, and with almost the same sector does not explain the results. Our results come from investors who experience luck, which we define as selling a stock before a large salient event devaluation of the company's stocks.

[19] Barber and Odean (2008) confirms that retail investor are net buyer of firms on the news. Barber et al. (2021) notice Hobbinhood user are more likely to engage in attention-bias trades due to the app's unique features. Hartzmark (2015) shows that stocks in the tails of portfolio returns are more likely to be negotiated. Reports that information in the news with the same relevance is incorporated in asset pricing with different speeds depending on their position on newsletter Fedyk (2018)

Table 7: Placebo Petrobras

Dependent Variables:	Turnover		Trades		Days with Trade	
Model:	(1)	(2)	(1)	(2)	(1)	(2)
Variables						
treated × post	0.0149	-0.0828	0.2371	-0.2908	0.0740	-0.4532**
	(0.0560)	(0.0952)	(0.1687)	(0.1865)	(0.1555)	(0.1358)
post	-0.0858***	-0.1246**	-0.5721***	-0.7359***	-0.3429***	-0.4640***
	(0.0189)	(0.0408)	(0.0594)	(0.0503)	(0.0521)	(0.1083)
(Intercept)	0.3157***	0.2898***	1.666***	1.557***	1.171***	1.090***
	(0.0186)	(0.0210)	(0.0540)	(0.0546)	(0.0513)	(0.0629)
treated	0.1235*	0.0583	-0.1834	-0.5353**	0.0526	-0.2988
	(0.0534)	(0.0914)	(0.1662)	(0.1804)	(0.1484)	(0.1546)
trend		0.0129		0.0546**		0.0404
		(0.0103)		(0.0150)		(0.0288)
trend × treated		0.0326		0.1760**		0.1757**
		(0.0350)		(0.0680)		(0.0561)
Fit statistics						
Adjusted R^2	0.00887	0.00966	0.00568	0.00668	0.00894	0.01220
Observations	43,176	43,176	43,176	43,176	43,176	43,176

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2). Using treated as those investor whose sold Petrobras stock, instead of Vale stocks before the Mariana Disaster. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 8: Placebo Petrobras Returns

Dependent Variables:	Portfolio Return		Portfolio Return Aft Fee	
Model:	(1)	(2)	(1)	(2)
Variables				
treated × post	f1 -0.0305	0.1224**	-0.0305	0.1238**
	(0.0463)	(0.0370)	(0.0464)	(0.0362) f1
post	0.0213	-0.0012	0.0225	0.0005
	(0.0261)	(0.0434)	(0.0260)	(0.0431)
(Intercept)	-0.0646**	-0.0796**	-0.0686**	-0.0833**
	(0.0188)	(0.0280)	(0.0187)	(0.0280)
treated	-0.0193	0.0827	-0.0206	0.0823
	(0.0207)	(0.0486)	(0.0209)	(0.0478)
trend		0.0075		0.0073
		(0.0111)		(0.0111)
trend × treated		-0.0510**		-0.0514**
		(0.0168)		(0.0165)
Fit statistics				
Adjusted R^2	0.01210	0.02675	0.01318	0.02831
Observations	43,176	43,176	43,176	43,176

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2). Using treated as those investor whose sold Petrobras stock, instead of Vale stocks before the Mariana Disaster. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

6 Conclusions

This article studies the effect of luck on the financial market, evaluating the impact on retail investors' trading activity. We define luck as selling a stock that suffers a significant devaluation due to an environmental disaster. This event had the necessary conditions to generate luck. First was a salient event, making it likely that investors be aware of it. Second, it greatly impacts the stock, generating the sensation of luck. The third was unpredictable, implying that an investor's skill did not generate this trade.

Empirically, compared similar investor, we document that lucky investors increase their trading activity, portfolio turnover 11%, number of trades 46% and days with trades 35% after the luck event. The positive effect shows some persistence, remaining statistically significant after three months, and is economically meaningful. Moreover, consistent with the literature, this excessive trading behavior is correlated with negative returns. Lucky investor performs 6% lower and 4% after adjusting for risk factors.

Furthermore, we report a series of robustness. First, we showed that our main estimates are in the extreme tails of a pseudo-treatment distribution. That is constructed by assigning control investors as treated, which rules out any spurious relation. Second, we do not find any treatment effect using past outcomes. Thus the difference in trading arises only after the disaster. Third, we use the same identification strategy (investors that sold a stock 21 days before the disaster) but changed Vale's stock for Petrobras'stock. Since this company did not suffer a significant devaluation because of the disaster, there was no luck in this case. No statistically significant effect is found in this case.

We contribute to the literature that tries to explain the excessive trading puzzle by using a different identification strategy than previous studies that investigate luck's impact on trading behavior. For comparison, Gao et al. (2021), and Anagol et al. (2021) uses IPO's lotteries. Instead, we used an unpredictable shock that caused a large devaluation in stock. Furthermore, we add to the literature the impact of behavioral biases in the trading activity of retail investors in the Brazilian context.

7 Appendix

7.1 Estimation using fixed effect of match

Table 9: Trading Activity Without Vale with Replacement

Dependent Variables:	Turnover		Trades		Days With Trades	
Model:	(1)	(2)	(1)	(2)	(1)	(2)
Variables						
treated × post	0.1053**	0.1053**	0.6060***	0.6060***	0.3289***	0.3289***
	(0.0409)	(0.0409)	(0.1993)	(0.1994)	(0.1167)	(0.1167)
treated	-0.2013***	-0.2013***	-0.9013***	-0.9013***	-0.6645***	-0.6645***
	(0.0270)	(0.0270)	(0.1373)	(0.1373)	(0.0865)	(0.0865)
post	-0.0694**	-0.0020	-0.4759***	-0.9320***	-0.2990***	-0.4108**
	(0.0304)	(0.0677)	(0.1513)	(0.2964)	(0.0971)	(0.1924)
trend		0.0409***		0.0186		0.0483
		(0.0139)		(0.0657)		(0.0389)
trend × post		-0.0380**		0.0800		-0.0066
		(0.0172)		(0.0749)		(0.0458)
Fixed-effects						
Match	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Adjusted R^2	0.11034	0.11180	0.12506	0.12523	0.14076	0.14105
R^2	0.18482	0.18648	0.19831	0.19878	0.21269	0.21327
Observations	5,472	5,472	5,472	5,472	5,472	5,472

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) after matching with replacement. To capture only the variation within match we add fixed effect match, this inclusion absorb the intercept coefficient in the estimations. All columns uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 10: Return Variables with Replacement

Dependent Variables:	Portfolio Return		Portfolio Return Aft Fee		Portfolio Risk Adjusted Return		Portfolio Risk Adjusted Return After Fee	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0659***	-0.0659***	-0.0670***	-0.0670***	-0.0437***	-0.0437***	-0.0446***	-0.0446***
	(0.0107)	(0.0107)	(0.0108)	(0.0108)	(0.0100)	(0.0100)	(0.0101)	(0.0101)
treated	0.0274***	0.0274***	0.0292***	0.0292***	0.0163**	0.0163**	0.0171***	0.0171***
	(0.0079)	(0.0079)	(0.0078)	(0.0078)	(0.0066)	(0.0066)	(0.0066)	(0.0066)
post	0.0176**	0.0241	0.0186**	0.0250	0.0491***	-0.1463***	0.0499***	-0.1464***
	(0.0082)	(0.0260)	(0.0083)	(0.0259)	(0.0076)	(0.0229)	(0.0077)	(0.0229)
trend		-0.0145***		-0.0149***		-0.0152***		-0.0156***
		(0.0053)		(0.0053)		(0.0052)		(0.0052)
trend × post		0.0074		0.0077		0.0482***		0.0486***
		(0.0072)		(0.0072)		(0.0064)		(0.0064)
Fixed-effects								
Match	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Adjusted R^2	0.03004	0.03249	0.02861	0.03122	-0.00119	0.01555	-0.00180	0.01520
R^2	0.11124	0.11384	0.10993	0.11267	0.08263	0.09832	0.08206	0.09800
Observations	5,472	5,472	5,472	5,472	5,472	5,472	5,472	5,472

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) after matching with replacement. To capture only the variation within match we add fixed effect match, this inclusion absorb the intercept coefficient in the estimations. All columns uses cluster at individual and pair match level, as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

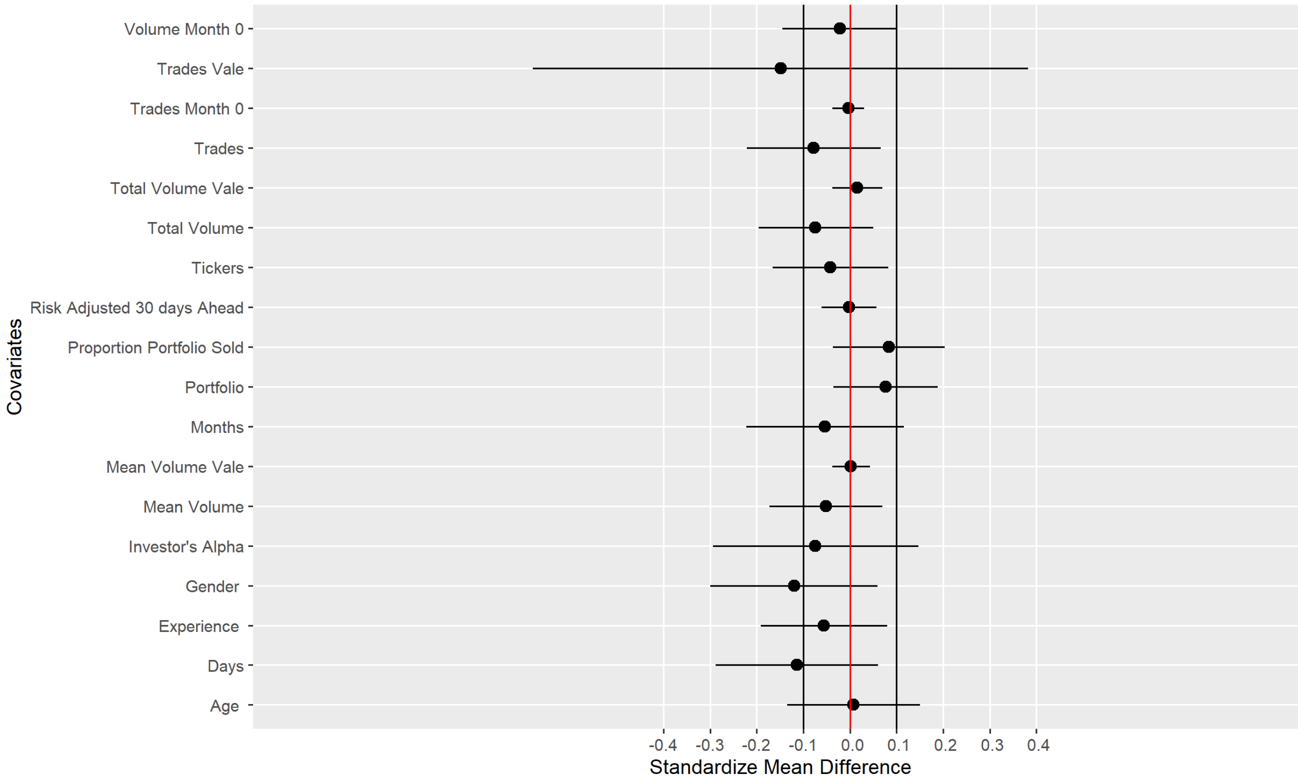
Table 11: Main estimation using Fixed effect

	Unmatched Sample						Matched Sample					
Dependent Variables:	Turnover		Trades		Days With Trades		Turnover		Trades		Days With Trades	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables												
treated × post	0.1128***	0.1508***	0.6180***	0.8382***	0.3499***	0.5625***	0.1101***	0.1821***	0.7353***	0.9559***	0.4257***	0.6021***
	(0.0252)	(0.0369)	(0.1243)	(0.1675)	(0.0617)	(0.0954)	(0.0384)	(0.0662)	(0.1935)	(0.2776)	(0.1006)	(0.1639)
treated	-0.2566***	-0.2313***	-0.9890***	-0.8421***	-0.7129***	-0.5711***	-0.2109***	-0.1629***	-0.9632***	-0.8162***	-0.7100***	-0.5923***
	(0.0150)	(0.0205)	(0.0795)	(0.1264)	(0.0491)	(0.0691)	(0.0259)	(0.0419)	(0.1431)	(0.2507)	(0.0855)	(0.1301)
post	-0.0836***	-0.1550***	-0.5302***	-0.8866***	-0.3399***	-0.6468***	-0.0657**	-0.1758***	-0.5580***	-0.9091***	-0.3725***	-0.6575***
	(0.0082)	(0.0134)	(0.0403)	(0.0545)	(0.0186)	(0.0303)	(0.0266)	(0.0516)	(0.1360)	(0.2047)	(0.0771)	(0.1257)
(Intercept)	0.3568***	0.3092***	1.661***	1.423***	1.217***	1.012***						
	(0.0074)	(0.0108)	(0.0350)	(0.0494)	(0.0175)	(0.0245)						
trend		0.0238***		0.1188***		0.1023***		0.0367**		0.1170		0.0950**
		(0.0037)		(0.0158)		(0.0086)		(0.0155)		(0.0750)		(0.0394)
trend × treated		-0.0126		-0.0734		-0.0709***		-0.0240		-0.0735		-0.0588
		(0.0090)		(0.0451)		(0.0257)		(0.0187)		(0.0894)		(0.0479)
Fixed-effects												
Match	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics												
Adjusted R^2	0.00697	0.00756	0.00793	0.00865	0.01407	0.01614	0.09282	0.09395	0.12151	0.12184	0.14212	0.14325
R^2	0.00703	0.00766	0.00799	0.00875	0.01413	0.01625	0.16881	0.17021	0.19509	0.19576	0.21398	0.21536
Observations	48,414	48,414	48,414	48,414	48,414	48,414	4,896	4,896	4,896	4,896	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) for unmatched sample. In the matched sample, we include a fixed effect of the match to capture the variation within the match. Since the match is one-to-one, including the fixed effect absorbs the intercept coefficient. In unmatched sample, the cluster is at the individual level. In matched sample is clustered in individual and pair matches as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

7.2 Inverse of Propensity Score Weighting Estimations

Figure 12: Covariance Balance after trimming



Note: This figure show the covariance balance between the two groups, treated and control after trimming the sample using the procedure suggested by Imbens and Rubin (2015). All variables are normalized. Thus the difference is measured in terms of standard deviations. The black vertical lines are -10% and 10% standardize mean difference.

Table 12: Return Estimation using Fixed effect

	Unmatched Sample				Matched Sample			
Dependent Variables:	Portfolio Return		Portfolio Risk Adjusted Return		Portfolio Return		Portfolio Risk Adjusted Return	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0656***	-0.0814***	-0.0449***	-0.0752***	-0.0601***	-0.0923***	-0.0383***	-0.0758***
	(0.0066)	(0.0167)	(0.0062)	(0.0159)	(0.0097)	(0.0253)	(0.0093)	(0.0238)
treated	0.0188***	0.0083	0.0094**	-0.0108	0.0232***	0.0017	0.0158**	-0.0092
	(0.0050)	(0.0101)	(0.0040)	(0.0088)	(0.0076)	(0.0157)	(0.0062)	(0.0135)
post	0.0109***	0.0345***	0.0432***	0.0200***	0.0119*	0.0635***	0.0447***	0.0393**
	(0.0013)	(0.0030)	(0.0013)	(0.0029)	(0.0064)	(0.0166)	(0.0061)	(0.0157)
(Intercept)	-0.0711***	-0.0554***	-0.0614***	-0.0769***				
	(0.0010)	(0.0020)	(0.0008)	(0.0017)				
trend		-0.0079***		0.0077***		-0.0172***		0.0018
		(0.0009)		(0.0008)		(0.0049)		(0.0045)
trend × treated		0.0053		0.0101**		0.0108		0.0125*
		(0.0046)		(0.0044)		(0.0072)		(0.0067)
Fixed-effects								
Match	No	No	No	No	Yes	Yes	Yes	Yes
Fit statistics								
Adjusted R^2	0.00361	0.00523	0.02216	0.02463	0.02145	0.02446	-0.00540	-0.00328
R^2	0.00367	0.00533	0.02222	0.02473	0.10341	0.10657	0.07881	0.08116
Observations	48,414	48,414	48,414	48,414	4,896	4,896	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) for unmatched sample. In the matched sample, we include a fixed effect of the match to capture the variation within the match. Since the match is one-to-one, including the fixed effect absorbs the intercept coefficient. In unmatched sample, the cluster is at the individual level. In matched sample is clustered in individual and pair matches as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 13: Return After Fee Estimations using Fixed Effect

	Unmatched Sample				Matched Sample			
Dependent Variables:	Portfolio Return Aft Fee Wo Vale		Portfolio Return Risk Adj After Fee		Portfolio Return Aft Fee Wo Vale		Portfolio Return Risk Adj After Fee	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0668***	-0.0831***	-0.0458***	-0.0769***	-0.0614***	-0.0942***	-0.0394***	-0.0772***
	(0.0065)	(0.0167)	(0.0062)	(0.0158)	(0.0097)	(0.0253)	(0.0093)	(0.0237)
treated	0.0212***	0.0103	0.0107***	-0.0101	0.0251***	0.0033	0.0168***	-0.0085
	(0.0049)	(0.0101)	(0.0040)	(0.0088)	(0.0075)	(0.0156)	(0.0061)	(0.0134)
post	0.0121***	0.0367***	0.0440***	0.0220***	0.0131**	0.0656***	0.0455***	0.0409***
	(0.0013)	(0.0029)	(0.0013)	(0.0029)	(0.0064)	(0.0166)	(0.0061)	(0.0157)
(Intercept)	-0.0756***	-0.0592***	-0.0628***	-0.0775***				
	(0.0010)	(0.0019)	(0.0008)	(0.0017)				
trend		-0.0082***		0.0073***		-0.0175***		0.0016
		(0.0009)		(0.0008)		(0.0049)		(0.0045)
trend × treated		0.0054		0.0104**		0.0109		0.0126*
		(0.0046)		(0.0044)		(0.0072)		(0.0067)
Fixed-effects								
Match	No	No	No	No	Yes	Yes	Yes	Yes
Fit statistics								
Adjusted R^2	0.00379	0.00556	0.02303	0.02532	0.01985	0.02303	-0.00588	-0.00380
R^2	0.00385	0.00567	0.02310	0.02542	0.10195	0.10526	0.07838	0.08069
Observations	48,414	48,414	48,414	48,414	4,896	4,896	4,896	4,896

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) for unmatched sample. In the matched sample, we include a fixed effect of the match to capture the variation within the match. Since the match is one-to-one, including the fixed effect absorbs the intercept coefficient. In unmatched sample, the cluster is at the individual level. In matched sample is clustered in individual and pair matches as suggested by Abadie and Spiess (2021). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 14: Trading Activity using Inverse of Propensity Score

	Complete Sample						Trimmed Sample					
Dependent Variables:	Turnover		Trades		Days With Trades		Turnover		Trades		Days With Trades	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables												
treated × post	0.0964***	0.1511***	0.6278***	0.7878***	0.3714***	0.5962***	0.1033***	0.1523**	0.6067***	0.7383***	0.3679***	0.5975***
	(0.0284)	(0.0538)	(0.1465)	(0.2177)	(0.0789)	(0.1428)	(0.0320)	(0.0601)	(0.1635)	(0.2438)	(0.0869)	(0.1605)
treated	-0.2127***	-0.1763***	-0.9771***	-0.8704***	-0.7359***	-0.5861***	-0.2162***	-0.1836***	-0.9666***	-0.8788***	-0.7442***	-0.5912***
	(0.0214)	(0.0254)	(0.1061)	(0.1643)	(0.0648)	(0.0951)	(0.0230)	(0.0271)	(0.1197)	(0.1877)	(0.0720)	(0.1073)
post	-0.0671***	-0.1553***	-0.5400***	-0.8362***	-0.3615***	-0.6805***	-0.0627***	-0.1530***	-0.5333***	-0.8408***	-0.3607***	-0.6926***
	(0.0155)	(0.0414)	(0.0874)	(0.1494)	(0.0526)	(0.1104)	(0.0176)	(0.0475)	(0.0997)	(0.1712)	(0.0602)	(0.1267)
(Intercept)	0.3129***	0.2541***	1.649***	1.452***	1.240***	1.027***	0.3054***	0.2452***	1.648***	1.443***	1.251***	1.029***
	(0.0169)	(0.0186)	(0.0785)	(0.1159)	(0.0458)	(0.0697)	(0.0193)	(0.0209)	(0.0895)	(0.1318)	(0.0524)	(0.0798)
trend		0.0294**		0.0987**		0.1063***		0.0301**		0.1025**		0.1106***
		(0.0114)		(0.0425)		(0.0311)		(0.0131)		(0.0485)		(0.0356)
trend × treated		-0.0182		-0.0533		-0.0749*		-0.0163		-0.0439		-0.0765*
		(0.0141)		(0.0599)		(0.0394)		(0.0158)		(0.0689)		(0.0450)
Fit statistics												
Adjusted R^2	0.02222	0.02315	0.01906	0.01952	0.03693	0.03851	0.02308	0.02404	0.01971	0.02016	0.03828	0.03991
R^2	0.02228	0.02325	0.01912	0.01962	0.03699	0.03861	0.02335	0.02449	0.01998	0.02061	0.03854	0.04035
Observations	48,414	48,414	48,414	48,414	48,414	48,414	10,926	10,926	10,926	10,926	10,926	10,926

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) using the inverse of probability as weights. The complete sample includes all treated and control investors are the same as the unmatched sample. In the trimmed sample, we exclude treated and control investors with a lower probability of treatment since the weights generated by those investors may distort our estimation. Following the suggestion of Imbens and Rubin (2015) trimmed the sample using an optimal cut. In our case, we remove investors with a lower than 7% of probability of being treated. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

7.3 Figures

7

Table 15: Portfolio Return using Inverse of Propensity Score

Dependent Variables:	Portfolio Return				Portfolio Risk Adjusted Return			
	Complete Sample		Trimmed Sample		Complete Sample		Trimmed Sample	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0765***	-0.1371***	-0.0833***	-0.1469***	-0.0507***	-0.1197***	-0.0575***	-0.1295***
	(0.0079)	(0.0206)	(0.0089)	(0.0233)	(0.0087)	(0.0191)	(0.0098)	(0.0216)
treated	0.0322***	-0.0082	0.0351***	-0.0074	0.0177***	-0.0283**	0.0188***	-0.0292**
	(0.0061)	(0.0127)	(0.0068)	(0.0146)	(0.0057)	(0.0124)	(0.0065)	(0.0141)
post	0.0218***	0.0902***	0.0232***	0.0982***	0.0490***	0.0644***	0.0496***	0.0709***
	(0.0047)	(0.0124)	(0.0054)	(0.0141)	(0.0062)	(0.0110)	(0.0071)	(0.0126)
(Intercept)	-0.0844***	-0.0388***	-0.0870***	-0.0371***	-0.0697***	-0.0594***	-0.0712***	-0.0570***
	(0.0037)	(0.0079)	(0.0042)	(0.0091)	(0.0042)	(0.0088)	(0.0048)	(0.0101)
trend		-0.0228***		-0.0250***		-0.0051		-0.0071*
		(0.0037)		(0.0042)		(0.0034)		(0.0039)
trend × treated		0.0202***		0.0212***		0.0230***		0.0240***
		(0.0059)		(0.0067)		(0.0055)		(0.0062)
Fit statistics								
Adjusted R^2	0.01326	0.01829	0.01478	0.02060	0.01070	0.01477	0.01114	0.01474
R^2	0.01332	0.01839	0.01505	0.02105	0.01076	0.01488	0.01141	0.01519
Observations	48,414	48,414	10,926	10,926	48,414	48,414	10,926	10,926

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) using the inverse of probability as weights. The complete sample includes all treated and control investors are the same as the unmatched sample. In the trimmed sample, we exclude treated and control investors with a lower probability of treatment since the weights generated by those investors may distort our estimation. Following the suggestion of Imbens and Rubin (2015) trimmed the sample using an optimal cut. In our case, we remove investors with a lower than 7% of probability of being treated. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 16: Portfolio Returns after fee using Inverse Propensity Score

Dependent Variables:	Portfolio Return Aft Fee				Portfolio Risk Adjusted Return After Fee			
	Complete Sample		Trimmed Sample		Complete Sample		Trimmed Sample	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Variables								
treated × post	-0.0777***	-0.1383***	-0.0845***	-0.1479***	-0.0517***	-0.1208***	-0.0585***	-0.1305***
	(0.0079)	(0.0206)	(0.0088)	(0.0233)	(0.0087)	(0.0191)	(0.0097)	(0.0216)
treated	0.0342***	-0.0062	0.0370***	-0.0053	0.0187***	-0.0274**	0.0197***	-0.0283**
	(0.0060)	(0.0127)	(0.0067)	(0.0145)	(0.0057)	(0.0124)	(0.0065)	(0.0142)
post	0.0230***	0.0919***	0.0245***	0.0998***	0.0499***	0.0659***	0.0506***	0.0723***
	(0.0046)	(0.0124)	(0.0053)	(0.0142)	(0.0062)	(0.0111)	(0.0071)	(0.0127)
(Intercept)	-0.0886***	-0.0427***	-0.0911***	-0.0409***	-0.0708***	-0.0602***	-0.0723***	-0.0578***
	(0.0037)	(0.0079)	(0.0042)	(0.0091)	(0.0042)	(0.0090)	(0.0049)	(0.0102)
trend		-0.0230***		-0.0251***		-0.0053		-0.0073*
		(0.0037)		(0.0043)		(0.0035)		(0.0039)
trend × treated		0.0202***		0.0211***		0.0230***		0.0240***
		(0.0059)		(0.0067)		(0.0055)		(0.0062)
Fit statistics								
Adjusted R^2	0.01341	0.01853	0.01492	0.02084	0.01106	0.01512	0.01151	0.01511
R^2	0.01347	0.01863	0.01519	0.02129	0.01112	0.01522	0.01179	0.01556
Observations	48,414	48,414	10,926	10,926	48,414	48,414	10,926	10,926

Note: This table reports the estimation from the models 1 columns (1) and 2 columns (2) using the inverse of probability as weights. The complete sample includes all treated and control investors are the same as the unmatched sample. In the trimmed sample, we exclude treated and control investors with a lower probability of treatment since the weights generated by those investors may distort our estimation. Following the suggestion of Imbens and Rubin (2015) trimmed the sample using an optimal cut. In our case, we remove investors with a lower than 7% of probability of being treated. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Figure 13: Main Newspaper's Front Page Day after the disaster

References

Abadie, A. and J. Spiess (2021). Robust post-matching inference. *Journal of the American Statistical Association*, 1–13.

Anagol, S., V. Balasubramaniam, and T. Ramadorai (2021). Learning from noise: Evidence from india's ipo lotteries. *Journal of Financial Economics 140*(3), 965–986.

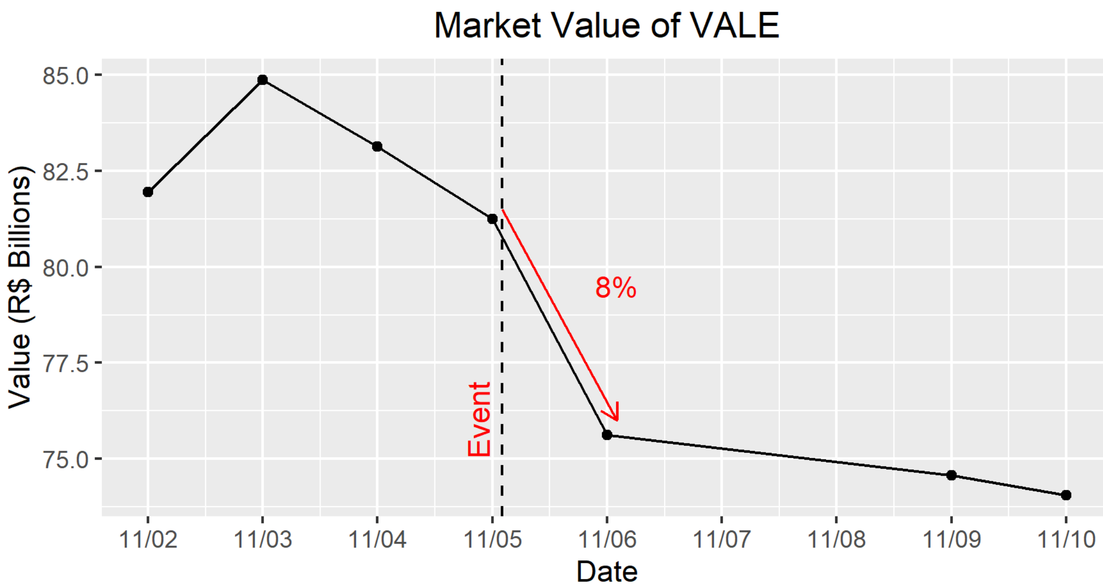
Barber, B. M., X. Huang, K. J. Ko, and T. Odean (2020). Leveraging overconfidence. *Available at SSRN 3445660*.

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

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

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

Figure 14: Vale's Value



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

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116*(1), 261–292.

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

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

Barberis, N. and M. Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review 98*(5), 2066–2100.

Barberis, N. and W. Xiong (2012). Realization utility. *Journal of Financial Economics 104*(2), 251–271.

Beyer, S. and E. M. Bowden (1997). Gender differences in seff-perceptions: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin 23*(2), 157–172.

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

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

Da, Z., X. Huang, and L. J. Jin (2021). Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics 140*(1), 175–196.

Darke, P. R. and J. L. Freedman (1997, apr). Lucky Events and Beliefs in Luck: Paradoxical Effects on Confidence and Risk-Taking. *Personality and Social Psychology Bulletin 23*(4), 378–388.

Deaux, K. and T. Emswiller (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of personality and social psychology 29*(1).

Dorn, D. and P. Sengmueller (2009). Trading as entertainment? *Management Science 55*(4), 591–603.

Fedyk, A. (2018). Front page news: The effect of news positioning on financial markets. Technical report, Working paper.

Gao, H., D. Shi, and B. Zhao (2021). Does good luck make people overconfident? Evidence from a natural experiment in the stock market. *Journal of Corporate Finance 68*(March), 101933.

Gao, X. and T. C. Lin (2015). Do Individual Investors Treat Trading as a Fun and Exciting Gambling Activity? Evidence from Repeated Natural Experiments. *Review of Financial Studies 28*(7), 2128–2166.

Gervais, S. and T. Odean (2001). Learning to be overconfident. *Review of Financial Studies 14*(1), 1–27.

Glaser, M. and M. Weber (2007). Overconfidence and trading volume. *GENEVA Risk and Insurance Review 32*(1), 1–36.

Grinblatt, M. and M. Keloharju (2009). Sensation seeking, overconfidence, and trading activity. *Journal of Finance 64*(2), 549–578.

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

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American economic review 70*(3), 393–408.

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

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Ingersoll, J. E. and L. J. Jin (2013). Realization utility with reference-dependent preferences. *The Review of Financial Studies 26*(3), 723–767.

Jin, L. J. and P. Sui (2022). Asset pricing with return extrapolation. *Journal of Financial Economics 145*(2), 273–295.

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

Langer, E. J. and J. Roth (1975). Heads i win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of personality and social psychology 32*(6), 951.

Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological bulletin 84*(1), 1.

Liu, H., C. Peng, W. A. Xiong, and W. Xiong (2021). Taming the bias zoo. *Journal of Financial Economics, forthcoming*.

Lundeberg, M. A., P. W. Fox, and J. Punćcohaŕ (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of educational psychology 86*(1), 114.

Lusardi, A. and O. S. Mitchell (2008). Planning and financial literacy: How do women fare? *American economic review 98*(2), 413–17.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *Journal of economic literature 52*(1), 5–44.

Odean, T. (1998, oct). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance 53*(5), 1775–1798.

Odean, T. (1999, dec). Do Investors Trade Too Much? *American Economic Review 89*(5), 1279–1298.

Prince, M. (1993). Women, men, and money styles. *Journal of economic Psychology*.

Schuster, B., F. Forsterlung, and B. Weiner (1989, jun). Perceiving the Causes of Success and Failure. *Journal of Cross-Cultural Psychology 20*(2), 191–213.

Shefrin, H. and M. Statman (2000). Behavioral portfolio theory. *Journal of financial and quantitative analysis 35*(2), 127–151.

Statman, M., S. Thorley, and K. Vorkink (2006). Investor overconfidence and trading volume. *Review of Financial Studies 19*(4), 1531–1565.

Wiseman, R. and C. Watt (2004). Measuring superstitious belief: Why lucky charms matter. *Personality and individual differences 37*(8), 1533–1541.