3-2021

Investment decisions and trading behavior of institutional and retail investors

Antonia KIRILOVA
Singapore Management University

INVESTMENT DECISIONS AND TRADING BEHAVIOR

OF INSTITUTIONAL AND RETAIL INVESTORS

ANTONIA KIRILOVA

SINGAPORE MANAGEMENT UNIVERSITY

2021

Investment Decisions and Trading Behavior of Institutional and Retail Investors

Antonia Kirilova

Submitted to Lee Kong Chian School of Business
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Business (Finance)

Dissertation Committee:

Ekkehart Boehmer (Chair)
Professor of Finance
Singapore Management University

Jianfeng Hu
Assistant Professor of Finance
Singapore Management University

Joe Zhe Zhang
Associate Professor of Finance
Singapore Management University

Seongkyu Gilbert Park
Assistant Professor
Hong Kong Polytechnic University

Singapore Management University
2021

I hereby declare that this PhD dissertation is my original work
and it has been written by me in its entirety.
I have duly acknowledged all the sources of information
which have been used in this PhD dissertation.

This PhD dissertation has also not been submitted for any degree
in any university previously.

Antonia Kirilova

15 March 2021

Investment Decisions and Trading Behavior of Institutional and Retail Investors

Antonia Kirilova

Abstract

This dissertation consists of three studies in the areas of empirical asset pricing, market microstructure, and behavioral finance. I study the trading behavior and portfolio choices of institutions and retail investors in the equity and derivatives markets. Examining the ways in which different market participants make investment decisions allows us to understand their role in shaping financial market dynamics. This is important in order to know how to structure markets for enhanced market efficiency, and to protect less sophisticated investors through better policies and regulations. Although there is a considerable amount of literature disputing the ability of retail investors and different types of institutions to make informed investment decisions, their trading patterns and the various effects they have on the markets are not fully understood. My dissertation aims to explore this broad issue from several different angles.

Chapter I examines retail investor activity in the extreme portfolios of well-known cross-sectional anomalies. This study is co-authored with Prof. Ekkehart Boehmer. We show that retail investors tend to trade in the opposite direction of anomalies (buying stocks in the short portfolios and selling stocks in the long portfolios), both before and after the anomaly variables become public information. However, we do not find evidence that retail trading is the cause of mispricing and subsequent return predictability. Stocks with high retail participation do not appear to be more mispriced after controlling for confounding factors. Instead of pushing prices away from fundamentals, contrarian retail trades are likely to provide liquidity to arbitrageurs after firm announcements. In addition, we show that retail short sellers exploit anomaly information and help to correct mispricing of overvalued stocks in the short portfolios of value-versus-growth anomalies. Overall, the goal of this study is to show that retail participation in equity markets is not detrimental to market efficiency and in certain settings can even be helpful in correcting anomalies.

Chapter II investigates trading styles and profitability of institutional and retail investors in a leading derivatives market. This study is co-authored with Prof. Jianfeng Hu, Prof. Seongkyu Gilbert Park, and Prof. Doojin Ryu. Using comprehensive account-level transaction data, we provide a detailed description of the options market and the different types of investors. We find that retail investors tend to stick to one trading style. About 70% of retail investors predominantly hold simple positions such as long calls or long puts. Institutional investors are more likely to use multiple strategies with various levels of complexity. We use trading style complexity as an ex-ante measure of trading skills and show that it significantly affects investment performance. Specifically, retail investors using simple strategies lose to the rest of the market. For both retail and institutional investors, volatility trading is the most profitable strategy, although subject to large downside risk. After adjusting for risk, Greek neutral strategies outperform. These style effects are persistent and cannot be explained by systematic risk exposure or known behavioral biases.

Chapter III is about rational regulation and irrational investors. It is co-authored with Prof. Jianfeng Hu. We show that irrational response to regulatory reforms aimed at investor protection can lead to these reforms having the opposite effect and hurting investors. After the August 2011 crisis in the Korean equity market, regulators increase the contract size of equity index options fivefold, hoping to limit retail participation and excessive speculation in the market. Contradicting the purpose of the reform, we find that investors' propensity to exit the market decreases after the reform. The dollar risk exposure of remaining investors significantly increases after the reform, consistent with investor inattention to the reform. Our estimation shows that it takes six months for risk taking activity to return to the pre-reform level but there is no significant decrease afterward. Heightened risk taking also leads to worse performance in the post-reform period. Although these effects are always stronger on retail investors, institutional investors are not spared either. In addition, we find that investors who are adversely affected by the reform exhibit self-attribution bias which causes them to extrapolate their performance into the future. They tend to outperform their peers before the crisis and their trading activity becomes more responsive to past performance after the crisis. However, limited attention to the market reform exacerbates their losses when their performance reverts to the mean. These results highlight the importance of considering behavioral biases in policy research and setting to avoid unintended consequences.

Table of Contents

Acknowledgements

Over the course of my PhD, many people have offered their kind support, and I would like to express my deepest gratitude to everyone who helped me along the way.

First, I would like to thank my supervisor, Prof. Ekkehart Boehmer, for his patient guidance and expert advice. His feedback and expertise helped me to expand my research abilities and successfully complete my dissertation.

I owe a big thank you to Prof. Jianfeng Hu for his mentorship and continuous support throughout my PhD. I learned a lot from him, and his encouragement motivated me to constantly improve as an academic.

In addition, I would like to acknowledge my committee members Prof. Joe Zhang and Prof. Gilbert Park for their assistance and insightful comments on my research. I am also thankful to Prof. Liang Hao for helping me and offering valuable advice on multiple occasions.

Moreover, I am very grateful for all the new friends I met in Singapore, as well as the friends and relatives back in Europe, who always root for me.

A special thank you goes to my family who are always by my side, even when I am so far from home. Their encouragement helped me to overcome many challenges, and the countless hours we spent talking and laughing on Skype made me feel closer to home. I am also forever grateful to my grandparents, who unfortunately are not with us anymore, but I remember their loving support and will always keep them in my heart.

Finally, I would especially like to thank my partner Ange for being an important part of my journey. He is a great support and puts a smile on my face every single day. Whether we are on an island adventure or on lockdown at home, I am always happy to share these moments with him.

Chapter I: Retail Trading on Anomalies

1. Introduction

Do retail investors trade on cross-sectional return anomalies? Or is their activity the cause of stock mispricing and the subsequent return predictability associated with anomalies? Although there is a large strand of literature contending the ability of retail investors to acquire and trade on private information, the question of whether and how retail investors trade on public accounting information used to construct anomaly portfolios is rather new and unexplored. Our study investigates retail activity in the extreme portfolios of well-known accounting anomalies and presents important evidence of the role that retail trading plays in market efficiency.

Numerous studies have shown that cross-sectional sorts based on observable variables can predict future stock returns.[1] Different views have been expressed regarding the source of such anomalies, but previous literature provides evidence that anomaly returns are at least partly caused by mispricing due to biased expectations which is corrected after firms release their financial information to the public. For example, Engelberg, McLean, and Pontiff (2018) show that anomaly returns are six times higher on earnings announcement days, and Bowles et al. (2020) show that in recent years anomaly returns are concentrated in the first 5-30 days after announcements. Within this framework which relates anomalies to stock mispricing, there are three different roles that an investor or a group of investors may play in the market. First, they may be the biased investors who push prices away from fundamentals and create stock mispricing before anomaly variables are public information. Second, they may help to correct

[1] See e.g. Green, Hand, Zhang (2013) and Hou, Xue, Zhang (2015) for an extensive list of return predictive signals.

mispricing directly by trading on the anomalies after announcements of accounting information. And third, they may help to correct mispricing indirectly by providing liquidity to arbitrageurs. Several recent papers explore institutions' and other market participants' role when it comes to trading on anomalies.[2] The role of retail investors in this context has not been studied extensively and is not well understood. For this reason, we investigate whether and how retail investors trade on mispriced stocks and the effect they have on market efficiency.

The most closely related study to ours is a working paper by McLean, Pontiff, and Reilly (2019). They construct an index of 131 stock return anomalies and show that retail traders tend to buy stocks in anomaly short portfolios and sell stocks in anomaly long portfolios. While our analysis confirms their results in general, we aim to go one step further and examine retail trading on anomalies in greater detail, with the aim of uncovering retail investors' role in mispricing and market efficiency. Ultimately, our goal is to provide evidence on the different roles that retail investors play in the market when it comes to trading on anomalies: whether they cause the anomaly mispricing before the release of accounting information, and whether they help to correct the mispricing after firm announcements, either directly by trading on anomaly information, or indirectly by providing liquidity to arbitrageurs. To do this, our study differs from McLean, Pontiff, and Reilly (2019) in several important aspects. First, we investigate retail trading using the traditional June-rebalancing of anomaly portfolios (following Fama and French, 1993) as well as a new announcement-rebalancing strategy (following Bowles et al., 2020). This allows us to observe any changes in retail behavior in windows before and after firms announce their accounting information. Second, we examine the relation between retail activity and stock mispricing by comparing the returns of stocks with high and low retail participation (measured as retail trading volume as a fraction

[2] See Akbas et al. (2015), Edelen, Ince, Kadlec (2016), Gao, Wang (2019), Cui et al. (2019), Calluzzo, Moneta, Topaloglu (2019), Chen, Da, Huang (2019), and others.

of total market volume). Third, we decompose retail trading into buy volume, sell volume, and short sell volume, to show the different patterns of retail trading in long and short anomaly portfolios. In addition, focusing on retail short sellers separately allows us to uncover a subset of more sophisticated retail traders. Finally, we also compare different categories of anomalies and discover that retail behavior is not uniform across all of them.

We follow the method of Boehmer et al. (2019) to identify marketable retail trades in the transaction-level data provided by TAQ and FINRA. Retail trades in TAQ are defined as off-exchange transactions with subpenny price improvements relative to the National Best Bid or Offer (NBBO). We separate these into retail buy trades (transactions at prices just below a round penny) and retail sell trades (transactions at prices just above a round penny). Using the same methodology, we identify retail short sale transactions in the FINRA data. We aggregate retail transactions into daily retail trading data from January 2010 to December 2018, and merge it with stock market data from CRSP and accounting data from Compustat. Using the Compustat data, we construct portfolios based on 25 anomaly variables which fall into four broad categories: Value-versus-growth anomalies, Investment anomalies, Profitability anomalies, and Intangibles anomalies. These include some of the most well-known accounting anomalies and constructing them is fairly straightforward. We only use accounting anomalies in our analysis because we want to examine how retail investors trade on public accounting information contained in a company's annual report. First, we sort stocks into decile portfolios based on each individual anomaly variable. Then, we construct two composite indexes that combine the information from all 25 anomalies. The first is an average portfolio index. For each stock, we take the equal-weighted average of its portfolio rankings based on the 25 individual anomalies. Stambaugh and Yuan (2017) suggest that averaging helps to remove the noise in identifying stock mispricing. The second is an extreme portfolios index. We follow McLean, Pontiff, and Reilly (2019) and compute the number of long portfolios minus the number of short portfolios that a stock belongs to. This approach focuses on the extreme portfolio deciles and disregards the rest.

The traditional approach for constructing anomaly portfolios (following Fama and French, 1993) consists of sorting portfolios at the end of June each year, since most firms in the broad cross section have released their annual accounting statements by that time. Hence, at the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on each anomaly variable measured at the fiscal year ending in calendar year $t-1$. However, using this traditional June-rebalancing strategy, which rebalances portfolios only once a year, tends to yield insignificant returns in recent years for many of the anomalies. Since our retail data only covers the period after 2010, we are constrained in our ability to analyze the full sample period available in Compustat. Regardless, why should we wait to rebalance portfolios at the end of June if many firms release their accounting information several months before that? It is reasonable to expect that investors would trade around the time when firms file their annual reports. Instead of waiting for all firms in the broad cross-section to release their financial statements and then forming anomaly portfolios, investors may be trading immediately after each firm announcement. Especially in recent years, when trading strategies that exploit anomaly information have risen in popularity, arbitrageurs may be trading to eliminate mispricing earlier than the traditional portfolio formation at the end of June. Hence, we construct announcement-rebalanced portfolios following the method of Bowles et al. (2020). The announcement-rebalancing strategy consists of rebalancing the anomaly portfolios every time one of the firms in the cross-section files its annual 10-K report. Compared to the June-rebalancing strategy, the announcement-rebalancing strategy requires higher turnover and incurs higher transaction costs, but delivers significantly higher positive abnormal returns. An event study reveals that anomaly returns are highest and most significant in the first month after announcement, although many anomalies continue to generate positive returns afterwards as well. Moreover, we find that anomaly returns tend to be negative in the year before announcement. In reality, trading during this period does not constitute a viable strategy because the accounting information is not public yet. However, if anomalies arise due to mispricing, then we would expect this to be the period when the mispricing is created. The negative anomaly results in that period confirm this assumption. That is, the stocks in the short (long) portfolio have their prices pushed up (down) and become overvalued (undervalued). Since retail investors are often accused of pushing prices away from fundamentals, we are interested in examining this period to check whether they are indeed the cause of mispricing and subsequent return predictability.

Our analysis uncovers several interesting patterns in retail trading on anomalies. We start by examining the relation between retail trading activity and our composite anomaly indexes, to see if retail investors appear to be trading on anomalies using either the June-rebalancing strategy or the announcement-rebalancing strategy. For both strategies, we find consistent evidence that the composite anomaly indexes are significantly negatively associated with retail buy volumes and significantly positively associated with retail sell volumes that close existing positions. This points to a tendency of retail investors to trade in the opposite direction of anomalies, buying stocks in the short portfolios and selling stocks in the long portfolios. We find that retail investors buy the overvalued stocks in the short anomaly portfolios both before firm announcements and during the one month after announcement. They correct their trading afterwards and stop buying these stocks, but there is no evidence that they start buying stocks in the long portfolios. Instead, they sell the undervalued stocks in the long anomaly portfolios during all time windows, both before and after firm announcements.

After observing these trading patterns, we investigate the question of whether retail investors are responsible for creating the mispricing associated with anomalies. Retail buy and sell volumes tend to be very close in quantity, meaning that any retail order imbalances are generally a very small percentage of overall market activity. This makes it unlikely that retail investors could have a large impact on market prices and push prices away from fundamentals in a significant way. However, we observe that retail volume as a percentage of total market volume is highest in the subsample of small-capitalization stocks. This means that retail investors may have an impact on the prices of small stocks. Combined with the fact that small stocks tend to be the most mispriced and generally drive anomaly returns, it is easy to see why retail traders are often blamed for creating mispricing. However, direct causal evidence of this has not been shown in the literature. This leads us to examine retail trading before firm announcements, which is the period when mispricing is created. Stocks that later end up in the long (short) anomaly portfolios have their prices pushed down (up) and become undervalued (overvalued). We split stocks within the extreme anomaly portfolios into stocks with high retail participation and stocks with low retail participation during the period before announcements (retail participation is measured as retail trading volume as a percentage of total market volume). We do this separately for each market capitalization quintile in order to control for the effect of firm size on retail participation. Then, we compare the returns of stocks with high and low retail participation, both before and after announcements. We find that, after controlling for firm size, there is no association between retail activity and stock mispricing. Although retail investors are often accused of pushing prices away from fundamentals, the evidence does not point to retail trading as the cause of mispricing and the subsequent return predictability associated with anomaly variables.

When it comes to the period after firm announcements, the fact that retail investors continue to trade in the opposite direction of the anomalies while anomaly portfolios generate positive abnormal returns, suggests that retail trading is contrarian in nature. This is consistent with a liquidity provision role. It is likely that retail investors provide liquidity to the arbitrageurs who are trading on anomaly information. In this way, retail investors may help to improve market efficiency by indirectly alleviating the mispricing associated with anomalies.

A subset of retail investors who appear to be more sophisticated are retail short sellers. This result is consistent with the literature which finds that short sellers in general act as arbitrageurs and correct the mispricing in short anomaly portfolios (see e.g. Chen, Da, and Huang, 2019). Retail short sale volumes are significantly negatively associated with the composite anomaly indexes, indicating that retail short sellers exploit anomaly information correctly. They consistently avoid shorting stocks in the long portfolios of anomalies. In addition, after the announcement of accounting variables to the public, they start to trade stocks in the short portfolios of value-versus-growth anomalies. Retail shorting of these overvalued stocks helps to correct mispricing and improve market efficiency.

Our study contributes to the literature in two main ways. First, we add to the literature on the characteristics and trading patterns of retail investors by showing how they trade on publicly available information about accounting anomaly variables. Second, we add to the literature on stock mispricing and the sources of return predictability. Retail investors are often assumed to be unsophisticated, uninformed, and biased traders who commit systematic mistakes and push prices away from fundamentals. However, there is debate in the literature, as outcomes are sensitive to the adopted methodology. For example, Barber and Odean (2000) argue that excessive trading, resulting from overconfidence, leads to the underperformance of retail investors. While this may be true if we examine returns net of transaction costs, it is important to emphasize that the poor returns are not necessarily the result of poor portfolio choices. In fact, households in their sample tend to hold high-beta, small-cap, and value stocks. In other words, they trade in the correct direction of the beta, size, and book-to-market anomalies. When it comes to tests of return predictability, Barber, Odean, and Zhu (2009) demonstrate that outcomes can differ depending on the chosen horizon. They show that, over weekly horizons, retail order imbalances positively predict future returns, while over annual horizons, retail trading is negatively related to future returns. The authors interpret this as evidence of retail traders moving markets and causing prices to deviate from underlying fundamentals. Though this is a possibility, there is no direct causal evidence in the literature to support the claim that retail investors are responsible for creating stock mispricing. Multiple other studies show that aggregate retail trading and retail order imbalances positively predict future stock returns (see Kaniel, Saar, and Titman, 2008; Kaniel et al., 2012; Kelley and Tetlock, 2013, 2017; Boehmer et al., 2019; Song, 2019). They interpret this as partly informed retail trading, and partly liquidity provision to other informed market participants, both of which enhance market efficiency. The focus of the above-mentioned studies is extracting any private information that may be contained in retail order imbalances and showing its predictability for future stock returns. On the other hand, our study focuses on the ability of retail investors to trade on publicly available information, in particular accounting anomaly variables contained in firms' annual reports.

The rest of the paper proceeds as follows. Section 2 develops alternative hypotheses. Section 3 describes the data and presents summary statistics. Section 4 includes an analysis of the main empirical results. Section 5 concludes.

2. Hypotheses

One source of stock return predictability has been shown to be mispricing due to biased expectations which is corrected after firms announce their accounting information (see e.g. Engelberg, McLean, and Pontiff, 2018; Bowles et al., 2020). This means that, in the period before firm announcements, when accounting variables are not public information yet, biased investors push prices away from fundamentals and create mispricing. Stocks that later end up in the long (short) anomaly portfolios have their prices pushed down (up) and become undervalued (overvalued). When firms release their accounting information to the public, arbitrageurs start trading to correct the mispricing, and this results in positive anomaly returns. Within this framework, there are three different roles that retail investors may take in the market when it comes to trading on anomalies. These three possible roles, which are not mutually exclusive, correspond to our three main hypotheses.

Hypothesis 1: First, retail investors may be the biased investors who push prices away from fundamentals and create stock mispricing. If this is the case, then before announcements of accounting information, we should observe retail buying (retail selling) of stocks that will later end up being in the short (long) anomaly portfolios. In addition, stocks with high retail participation must be more mispriced.

Hypothesis 2: Second, retail investors may help to correct mispricing directly by acting as arbitrageurs and trading to exploit anomaly information after the announcement of accounting variables to the public. In this case, we should observe retail buying of stocks in long anomaly portfolios and/or retail selling of stocks in short anomaly portfolios after firm announcements.

Hypothesis 3: Third, retail investors may help to correct mispricing indirectly by providing liquidity to arbitrageurs. If this is the case, then after firm announcements, we should observe contrarian retail trading, consisting of buying stocks in short anomaly portfolios and/or selling stocks in long anomaly portfolios.

3. Data and summary statistics

We combine daily stock data from CRSP with data about the anomaly variables from Compustat, and data on retail activity from TAQ and FINRA. First, we download CRSP data for all common stocks (with share code 10 or 11) listed on the NYSE, NYSE MKT (formerly Amex), and NASDAQ. To exclude penny stocks, we require stocks to have a price of at least $1 at the previous month-end. We also exclude financial firms and firms with negative book equity. Second, we construct anomaly portfolios using annual accounting data from Compustat. We examine the following 25 anomaly variables split into four broad categories: 1) Value-versus-growth anomalies, which include Book-to-Market (BM), Assets to Market Equity (AM), Earnings to Price (EP), Cash Flow to Market Equity (CFP), Payouts to Market Equity (POM), and Net Payouts to Market Equity (NPOM); 2) Investment anomalies, which include Asset Growth (AG), Investment to Assets (IA), Investment Growth (IG), Net Operating Assets (NOA), Net Stock Issues (NSI), Net External Financing (NXF), Inventory Changes (INV), Operating Accruals (OA), and Percent Operating Accruals (POA); 3) Profitability anomalies, which include Gross Profitability (GP), Gross Margins (GM), Asset Turnover (ATRN), Capital Turnover (CTRN), and Taxable Income to Book Income (TIBI); and 4) Intangibles anomalies, which include Advertising Expenses to Market Equity (ADM), R&D to Sales (RDS), R&D to Market Equity (RDM), Operating Leverage (OL), and Hiring Rate (HR). Details about the anomalies and their calculation are provided in the Appendix. We choose these anomalies because they include some of the most well-known accounting anomalies, and they fall into different categories to ensure that our analysis is not dominated by one type of anomaly. In addition, constructing these anomalies is fairly straightforward, and if our goal is to investigate whether retail investors trade on any of the anomalies, it is reasonable to begin the analysis with those that are widely known and relatively easy to exploit. Many of the anomalies that we include in our analysis are analyzed by Hou, Xue and Zhang (2015), and several of them are from the list of 11 anomalies that Stambaugh and Yuan (2017) use to construct their mispricing factors. We only use accounting anomalies in our analysis because we want to examine how retail investors trade on public accounting information contained in a company's annual report.

We first sort stocks into deciles based on each accounting variable. After we assign each stock to its corresponding decile based on each individual anomaly, we also construct two composite indexes that combine the information from all 25 anomalies. The first is an average portfolio index. We follow the method of Stambaugh and Yuan (2017) who construct their factors by averaging rankings across multiple anomalies. They suggest that averaging helps to remove the noise in identifying stock mispricing. Hence, for each stock, we take the average of its portfolio assignments based on the 25 individual anomalies. We assign an equal weight to each anomaly in the index. The second is an extreme portfolios index. We follow McLean, Pontiff, and Reilly (2019) and for each stock-month, we calculate the number of long portfolios minus the number of short portfolios that a stock belongs to in a given month. For example, if on a given month stock XYZ is assigned to the top Book-to-Market decile, the top Gross Profitability decile, and the bottom Operating Leverage decile, then its score for the number of long portfolios is equal to 2, its score for the number of short portfolios is equal to 1, and its overall anomaly score is 2 minus 1, equal to 1. Since for the rest of the anomalies the stock falls within deciles 2-9, rather than the extreme deciles 1 or 10, we disregard this information in the calculation of the anomaly score. Hence, the score tells us the number of extreme anomaly portfolios that each stock belongs to. The correlation between the average portfolio index and the extreme portfolios index is about 86% in the pooled sample. The high correlation indicates that they capture very similar information, but the latter puts more emphasis on the extreme deciles and disregards the rest.

The traditional approach for constructing anomaly portfolios (following Fama and French, 1993) consists of using accounting data from companies' financial statements submitted after each fiscal year end, and sorting portfolios at the end of June each year, since most firms have released their annual reports at that time, meaning that all the accounting variables for the broad cross section are public information. Hence, at the end of June of each year t, we use NYSE breakpoints to sort the stocks in our sample into deciles based on each anomaly variable, which is measured at the fiscal year ending in calendar year $t-1$. However, using this traditional June-rebalancing strategy which rebalances portfolios only once a year tends to generate insignificant returns in recent years for many of the anomalies. Since our retail data only covers the period after 2010, we are constrained in our ability to analyze the full sample period available in Compustat. If accounting information is released several months earlier, why should we wait to rebalance portfolios at the end of June? We can expect that many investors would trade around the time when firms file their annual reports instead. Especially in recent years when trading strategies exploiting market anomalies have risen in popularity, arbitrageurs may be trading to eliminate mispricing earlier than the traditional portfolio formation at the end of June. Instead of waiting for all firms to release their annual financial statements and then forming anomaly portfolios to exploit mispricing in the broad cross-section, investors could be trading immediately following firm news. Engelberg, McLean, and Pontiff (2018) show that anomaly returns are six times higher on earnings announcement days. They argue that the anomaly returns are driven by mispricing due to biased expectations, which is corrected at least partially when firms release news. Bowles et al. (2020) show that in recent years, anomaly returns are concentrated in the first 5-30 trading days after announcements of accounting information. Following their method, we also construct announcement-rebalanced portfolios. The announcement-rebalancing strategy consists of rebalancing the anomaly portfolios every time a firm files its annual 10-K report. Usually firms file their reports between February and April, and since most announcements are concentrated in those months, the anomaly portfolios may potentially be rebalanced daily. Since some news may be announced after the closing of the market, we rebalance the portfolios on the day after the announcement date. Compared to the June-rebalancing strategy, the announcement-rebalancing strategy requires higher turnover and incurs higher transaction costs, but delivers significant positive returns, especially in the first month after announcement. Table 1, Panel A shows mean cumulative abnormal returns (CARs) of anomaly long-minus-short portfolios in windows around 10-K announcements. The anomaly portfolios are sorted in calendar time and rebalanced on the day after each announcement. We align all firms' 10-K announcement dates at t=0 and compute equal-weighted average CARs for each anomaly portfolio in each monthly window, starting from 10 months before the announcement and ending 10 months after. We can see that for most of the anomalies, the announcement-rebalancing strategy provides significant abnormal returns after the firm announcement.[3] The CARs are largest and most significant in the (0, +20) window, which is the first month after announcement. The returns of the composite anomaly indexes are strongly significant and larger than the individual anomalies. This confirms that a combination of several anomalies into one removes noise and captures anomaly information in a better way. Therefore, we believe that the composite anomaly indexes are useful in our analyses as a signal that is strongly related to returns. Although we do not expect that each individual investor would be able to trade on all 25 anomalies, we can expect that their aggregate trading may be related to the composite anomaly signal. Table 1 also shows that most anomalies continue to generate positive abnormal returns after the first month post-announcement. This means that hard-to-arbitrage mispricing persists even after the accounting variables become public information, and arbitrageurs continue to

[3] It also appears that anomalies start generating positive abnormal returns in the months before announcement. Part of the reason is the sorting of firms into deciles that would tend to be traded as a portfolio. A firm may be assigned to a certain portfolio after the announcement of another firm and before its own announcement, and hence would begin to be traded by arbitrageurs before t=0. Another reason is that some accounting variables may be public before t=0, since some firms release their earnings announcement up to 3 months before their 10-K announcement. Unfortunately, we do not have access to data about which accounting variables are contained in each firm's earnings announcement. However, we conduct a robustness check using only the firms that release their earnings announcement and 10-K announcement on the same date, and our main results remain similar. In addition, we control for post-earnings-announcement drift (PEAD) after each quarterly earnings announcement.

trade on anomalies for months after firm announcements. This is consistent with limits to arbitrage, especially due to short selling constraints, and previous literature shows that mispricing of overvalued stocks remains unarbitraged for a longer time compared to undervalued stocks (see e.g. Miller, 1977; Shleifer and Vishny, 1997; Stambaugh, Yu, and Yuan, 2012). We also see that anomaly portfolio CARs tend to be negative in the windows before announcement. In reality, trading in these windows does not constitute a viable strategy because the accounting information is not public at this time, so investors are not able to trade on the anomalies yet. However, if anomalies arise due to mispricing, then we would expect this to be the period when the mispricing is created. That is, the stocks in the short (long) portfolio would have their prices pushed up (down) and become overvalued (undervalued). Since retail investors are often accused of pushing prices away from fundamentals, we are interested in examining this period to check whether they are indeed the cause of mispricing and subsequent return predictability.

Panel B of Table 1 shows mean CARs of the two composite anomalies in subsample of stocks sorted by firm size. The results clearly show that anomaly returns tend to be driven by small-capitalization stocks. Small-cap stocks have significantly higher CARs, especially in the firm month after announcement, and the returns appear to be more persistent compared to those of large-cap stocks. In Panel C, we split stocks into subsamples based on their post-earnings-announcement drift (PEAD) portfolio to control for the known phenomenon of stock CARs' tendency to drift in the direction of an earnings surprise for several weeks or even months after the earnings announcement. Quarterly standardized unexpected earnings (SUEs) are calculated using Compustat seasonal changes in earnings per share (EPS), following Livnat and Mendenhall (2006) and Hirshleifer et al. (2008). Then, daily SUE-based portfolios are formed and held from the announcement day until the next announcement. We mark negative-SUE stocks as those in decile 1 or 2, zero-SUE stocks as those in decile 5 or 6, and positive-SUE stocks as those in decile 9 or 10. We can see that the CARs of the two composite anomalies are significant and positive in the month after 10-K announcement, regardless of their earnings surprise. However, it seems that positive-SUE stocks are the ones that start generating positive and significant returns in the months before 10-K announcement. This is to be expected, since firms usually release their earnings announcement before the 10-K announcement. Hence, PEAD appears to drive at least part of the returns of the accounting anomalies, and may also be related to retail trading. Therefore, we need to control for PEAD in our subsequent regression analyses.

[Table 1 about here]

In Table 2, we use regression analysis to test the return patterns observed around 10-K announcements. We calculate mean daily return of each stock in three-month windows before its annual 10-K announcements, in the one-month window after the announcement, and in three-month windows afterwards. Then, for each of the time windows, we regress the mean daily return on two independent variables that measure the number of long anomaly portfolios and the number of short anomaly portfolios that the stock enters at the time of its 10-K announcement (t=0). We also include several control variables. Although our study focuses on retail investor activity in the extreme portfolios of accounting anomalies, we need to control for the effect of the momentum anomaly which may also be related to retail trading. Momentum decile portfolios are created daily based on compounded returns over the past six months, skipping one month between the formation period and the holding period, following Jegadeesh and Titman (1993). We also control for PEAD associated with quarterly earnings announcements. Hence, we include a variable for the SUE portfolio that the stock falls into, to control for any effect of earnings announcements on retail trading. Some retail investors may base their investment decisions on analyst recommendations, so we control for the consensus recommendation from IBES. We transform the scale as follows: 2=Strong Buy, 1=Buy, 0=Hold, -1=Underperform, -2=Sell. If a stock has no analyst coverage at a given time, we set the consensus recommendation to zero. We also control for institutional order imbalance, which is computed using all TAQ transactions excluding retail trades, signed with the Lee and Ready (1991) algorithm. The rest of the control variables include firm size, past-month return, past-month return volatility, and share turnover. The regression results in Table 2 show a similar pattern to the one in Table 1. Stocks that enter more long (short) anomaly portfolios at announcement have significant positive (negative) returns in the month after announcement, as well as afterwards in the longer run. In addition, the fact that stocks in long (short) anomaly portfolios have significant negative (positive) returns before announcements is consistent with the idea that these stocks are mispriced and the mispricing is corrected upon release of accounting information to the public.

[Table 2 about here]

To motivate our expectation that smart investors would trade on anomalies around firm announcements, we proceed to compare the returns of the announcement-rebalancing strategy with those of the June-rebalancing strategy. Figure 1 shows the difference in cumulative abnormal returns (CARs) between the two strategies when used to trade on the composite anomaly indexes. A stock's abnormal return is equal to its realized return minus expected return, where expected returns are calculated using the market model by regressing past one-year returns on market returns. The figures in Panel A plot the mean yearly CAR of the two composite anomaly indexes. Abnormal returns are cumulated starting from 1 July of year t until 30 June of year t+1. This is the traditional investment timeline when portfolios are rebalanced at the end of June and held over the next one year. We can see that the CARs of the two strategies closely follow each other from July up until January. However, starting from February when firms begin to release their 10-K reports, the announcement-rebalancing strategy starts to significantly outperform the June-rebalancing strategy. This shows that there are benefits to rebalancing portfolios as soon as accounting variables are released. The figures in Panel B plot the CAR earned with the two strategies over the whole sample period from 2010 to 2018. At the end of this period, the announcement-rebalancing strategy has cumulative performance twice as large as that of the June-rebalancing strategy.

[Figure 1 about here]

We follow the method of Boehmer et al. (2019) to identify marketable retail trades in the TAQ data. We start by separating off-exchange trades and extracting the trades with subpenny price improvements. Boehmer et al. (2019) explain that, in the U.S. equity markets, most of the orders initiated by retail investors do not take place on stock exchanges. Instead, they are either internalized by the retail brokers or sold to wholesalers. Such transactions that take place off-exchange are reported to a FINRA Trade Reporting Facility (TRF) and included in the TAQ data with exchange code "D". Typically, an order which is executed by a wholesaler or internalized receives a small price improvement relative to the National Best Bid or Offer (NBBO), usually a fraction of a cent. This allows us to separate retail order flow in the data. In particular, we identify retail buy trades as transactions at prices just below a round penny, and retail sell trades as transactions at prices just above a round penny. To avoid including institutional trades, such as those executed in dark pools, we exclude transactions reported at a round penny or near the midquote (at 0.4, 0.5, or 0.6 cent). Since Regulation National Market System (Reg NMS) requires all limit orders to be submitted at round pennies, this method is able to identify only the marketable retail orders. In other words, we study trades initiated by retail investors, rather than retail limit orders.

TAQ data on retail trades can be extracted starting from 2005 when Reg NMS was introduced, but the practice of providing subpenny price improvements became widespread only after 2009. For this reason, we study retail trading activity in the period from 2010 to 2018. The U.S. Securities and Exchange Commission (SEC) implemented a Tick Size Pilot

Program from October 2016 until September 2018. The program affected the tick size, and therefore the practice of providing subpenny price improvements, for approximately 1200 small-capitalization stocks. For this reason, we exclude from our sample the observation of the stocks that were in the pilot test groups during the period when their tick sizes were affected.

We also obtain retail short sales from the FINRA website[4]. Stock exchanges in the U.S. report short sale transactions executed off-exchange to FINRA Trade Reporting Facilities (TRFs) and FINRA makes that data available to the public. We download data from the Monthly Short Sale Transaction Files for the period 2010-2018. We use transactions reported by the NYX TRF or the NASDAQ TRF during normal trading hours, and only keep the non-exempt transactions. The FINRA short sales data is a subset of the TAQ off-exchange transactions data, so we use the same method to identify retail transactions (following Boehmer et al. (2019) and Song (2019)). That is, retail short sell trades are those transactions executed at prices just above a round penny (again, excluding transactions at a round penny or at 0.4 or 0.5 cent). Then, we aggregate retail transactions to calculate daily retail trading volume.

Finally, we merge the retail trading data from TAQ and FINRA with the CRSP data, and we keep only stocks for which we have Compustat anomalies data as well. On an average day, there are 2979 unique firms in our sample.

Table 3 provides summary statistics of the variables which capture retail trading activity in the pooled sample from 2010 to 2018. We report total retail trading volume and its percentage of total trading volume by all market participants (reported both on- and off-exchange in TAQ). The percentage of daily trading in each stock that is executed by retail investors aims to capture their participation in the market. Next, we compute retail order imbalance (OI) as the difference between retail buy and sell volumes divided by the sum of retail buy and sell volumes. We also report separately retail buy volume, retail sell volume

[4] https://www.finra.org/filing-reporting/trf/trf-regulation-sho-2020

(which closes existing positions), and retail short sell volume, as well as their percentage of total retail volume and total market volume. For each variable, we report the number of stock-day observations, mean, standard deviation, minimum, median, and maximum across the pooled sample of stock-day observations. Panel A reports summary statistics in the full sample of all stocks. On average, retail participation constitutes 9.3% of the overall market trading volume, and the median retail participation is 5.6%. Retail investors are slightly contrarian on average, consistent with the results of Boehmer et al. (2019). Panels B and C report statistics in the subsamples of small-capitalization stocks and large-capitalization stocks, respectively. We calculate firm size or market capitalization as the product of stock price and shares outstanding. Each day, we sort stocks by the lag of their market capitalization to split the sample into five quintiles based on firm size. The small-cap (large-cap) subsample contains the stocks in the bottom (top) quintile. Although the mean daily retail volume in small stocks is considerably smaller than in the full sample, retail participation is much larger. On average, retail volume constitutes 21.4% of overall market trading volume in small caps, and the median retail participation in small-cap stocks is 18.5%. The opposite is true of the large-cap stocks, where retail participation is only about 5% on average. Therefore, if retail investors have an impact on market prices, it is likely to be more substantial in the small-cap subsample.

[Table 3 about here]

4. Main empirical results

We start by examining the relation between retail trading activity and our anomaly indicators in stock-day observations. We want to see whether retail investors appear to be trading on anomalies using either the June-rebalancing strategy or the announcement-rebalancing strategy. First, we want to check whether retail investors may be trading using the traditional June-rebalancing strategy instead of the more advanced announcement-rebalancing strategy. Although rebalancing portfolios immediately after firms release their annual reports provides superior returns, it requires much higher turnover and therefore may incur substantial transaction costs. It may be more feasible and less costly for retail investors to sort portfolios only once a year, at the end of June. This strategy misses out on the returns generated by trading quickly after announcements but can still be profitable since it exploits any remaining mispricing that has not been already arbitraged. Thus, in Table 4, we show the results of Fama-MacBeth regressions of retail buy, sell, and short sell volume as a percentage of total retail volume on the composite anomaly indexes, computed following the June -rebalancing strategy. Following the traditional anomaly literature, portfolios are sorted at the end of June of each year t, based on accounting information from the fiscal year ending in calendar year t−1, held for one year after that and rebalanced annually. We examine retail activity during the portfolio formation period and during the one-year holding period after that. For the purposes of our analysis, the portfolio formation period is assumed to include the two weeks before the end of June and the two weeks at the start of July of each year. For each dependent variable and each period, the first column is a regression on the average portfolio index, and the second is a regression on the extreme portfolios index. We include control variables for momentum portfolio, PEAD (SUE portfolio), analysts' consensus recommendation, institutional order imbalance, firm size, previous-day return, previous-month mean daily return, previous-month return volatility, and share turnover. The results of the regressions show that the composite anomaly indexes are significantly negatively associated with retail buy volumes and significantly positively associated with retail sell volumes that close existing positions. This points to a tendency of retail investors to trade in the opposite direction of anomalies, buying stocks sorted into the short deciles and selling stocks sorted into the long deciles, both during portfolio formation and afterwards. In other words, retail investors tend to buy stocks which should be shorted and sell stocks which should be bought according to the anomaly indexes. The only part of retail trading which exploits anomaly information correctly is retail short selling. Retail short sale volumes are significantly negatively associated with the composite anomaly indexes. This means that retail investors prefer to short stocks in the short anomaly portfolios, thus directly correcting mispricing of overvalued stocks. This result is consistent with the literature which finds that short sellers in general trade as arbitrageurs and correct the mispricing in short anomaly portfolios (see e.g. Chen, Da, Huang (2019). Therefore, our study shows that retail short sellers act in a similar way to institutional short sellers when it comes to trading on public anomaly information. They help to correct any mispricing that has remained months after the release of accounting information to the public.

[Table 4 about here]

In Table 5, we repeat the regression analyses of Table 4, but this time our composite anomaly indexes are computed following the announcement-rebalancing strategy. Each time when a firm files its annual 10-K report, the anomaly portfolios are rebalanced on the following day. The results are largely consistent with Table 4. Once again, we find that retail buy volumes and sell volumes that close existing positions are on the opposite direction of anomalies. Hence, retail investors do not seem to trade on anomalies using the announcement-rebalancing strategy either. Since we observe significant and positive abnormal returns of anomaly portfolios during that period, retail trading appears to be contrarian in nature. This is consistent with a liquidity provision role. It is possible that retail investors provide liquidity to arbitrageurs who are trading on the anomaly portfolios. In this way, retail trading may be helping to improve market efficiency indirectly, by alleviating the mispricing associated with anomalies. Once again, retail short sellers trade in the right direction of anomalies. This means that they start shorting mispriced stocks soon after the release of accounting information.

[Table 5 about here]

Figure 2 summarizes and graphs the results observed in Table 5. It shows average retail trading volumes in decile portfolios (marked as SHORT, 2, 3, …, 8, 9, LONG from bottom to top portfolio) sorted on the composite anomaly. We use the average portfolio index as the composite anomaly in these figures, but results are similar if we use the extreme portfolios index. We calculate average retail buy, sell, and short sell volume as a percentage of total retail volume. First, we calculated the average of a variable across stocks in each anomaly decile, and then we take the time-series mean over the sample period from 2010 to 2018. It is clear that retail investors prefer to buy stocks in the short anomaly portfolios and sell stocks in the long anomaly portfolios. On the other hand, retail short sellers prefer to short overvalued stocks in the short anomaly portfolios and tend to avoid shorting stocks in the long portfolios.

[Figure 2 about here]

Our composite anomaly indexes capture the information of 25 individual anomalies. However, as we show in the Appendix, these anomalies fall within four distinct groups: Value-vs-growth anomalies, Investment anomalies, Profitability anomalies, and Intangibles anomalies. Out next analysis examines whether retail investors trade differently on the different anomaly categories. For each category, we compute an average portfolio index as the equal-weighted average of a stock's rankings for all of the individual anomalies within that category. Then, we use these four indexes to run regressions similar to the ones in Table 5. The results are reported in Table 6. They are consistent with our main inferences for all anomaly categories, except for the Profitability anomalies. Retail trading is very different in this category. Retail investors seem to buy stocks that are in the long portfolios of profitability anomalies and sell stocks in the short portfolios, hence trading correctly on these anomaly variables and exploiting the return predictability. Retail short sellers, however, appear to be trading incorrectly on the profitability anomalies.

[Table 6 about here]

The fact that the majority of retail activity tends to be in the wrong direction of anomalies, even after the accounting information has been released to the public, prompts us to examine retail trading before firms announce their accounting information. We want to answer the question of whether retail investors are responsible for creating the mispricing associated with anomalies. Interestingly, the study of McLean, Pontiff and Reilly (2019) finds that both institutional and retail investors trade in the opposite direction of anomalies (while firms through their share issuance and repurchase decisions turn out to be the "smart money"). Edelen, Ince, and Kadlec (2016) also show that, in the one-year period before portfolio formation, institutions tend to buy stocks which are later classified as overvalued and end up being in the short leg of anomalies. Hence, they suggest that institutions may cause anomalies, rather than arbitrage them away. These studies do not provide a direct causal link between trading on the wrong side of anomalies and stock mispricing. Nevertheless, we can argue that, if such patterns of trading are the cause of mispricing prior to the release of accounting information to the public, then it is more likely for institutions to have a larger impact on prices, given their significantly larger trading volumes as compared to retail investors. Retail buy and sell volumes tend to be very close in quantity, meaning that any retail order imbalances are generally a very small percentage of overall trading activity. This makes it unlikely that retail investors could have a large impact on market prices and push prices away from fundamentals in a significant way. However, we observe that retail volume as a percentage of total market volume is highest in the small-cap subsample. This means that retail investors may have an impact on the prices of small stocks. As we observed in Table 1 Panel B, small-capitalization stocks also happen to be the most mispriced stocks which generally drive anomaly returns. Considering the fact that anomaly returns tend to be larger and more persistent in small stocks, it is easy to see why retail traders are sometimes blamed for creating mispricing, although direct causal evidence of this has not been shown in the literature. This leads us to explore the question of whether high retail participation before firm announcements is the source of mispricing. Table 7 presents regressions where we use firm-year observations aligned at t=0 which is the date of a firm's annual 10-K announcement. In regressions 1-4, the dependent variable is the mean daily return of each stock in the window from 210 days to 64 days before its annual announcement. We use observations only up until 64 days before an announcement to avoid any overlap with the period of earnings announcements. The accounting variables are not public information at that time, so trading on the anomalies does not constitute a viable trading strategy. Instead, the return during the window (-210,-64) captures any mispricing created before the anomaly variables become public information at t=0. We regress this return on the composite anomaly indexes at the future time of announcement (t=0). The coefficients are negative and significant, confirming the fact that mispricing is created during the window (-210,-64). This means that stocks which later end up in the long (short) anomaly portfolios have their prices pushed down (up) and become undervalued (overvalued), resulting in negative anomaly returns. What we are mainly interested in here are the interaction effects with retail participation. Retail participation is measured as retail trading volume as a percentage of total market volume. To capture any effect of retail activity on returns, we add an interaction of the composite anomaly indexes with a dummy variable which we call *High retail participation (-210,-64)*. It is equal to one if retail participation during the period (-210,-64) is above the median retail participation computed within each firm size quintile separately, to control for the effect of firm size on retail participation. We conduct the analysis in the full sample as well as in the subsample of small-capitalization stocks, since they tend to be the most mispriced and most heavily traded by retail investors. The interaction effects are insignificant in all regressions. This means that stocks with high retail participation before 10-K announcements are not more mispriced that stocks with lower retail participation. Furthermore, we test whether there is any relation between retail participation during the period (-210,-64) and subsequent anomaly returns in the month after announcement. Hence, we regress mean daily return during the window (0,20) on the composite anomaly indexes and the same interaction effects with *High retail participation (-210,-64)*. As expected, the anomaly indexes are significantly positively associated with returns after the release of accounting information to the public. Once again, the interaction effects are insignificant. Although retail investors are often accused of pushing prices away from fundamentals, the evidence does not point to retail trading as the cause of mispricing and the subsequent return predictability associated with anomaly variables.

[Table 7 about here]

As we observed in Table 1, the profitability of anomaly strategies is highly sensitive to the speed of trading after the release of firm announcements. For most anomalies, the first month after announcement generates the highest abnormal returns. Next, we investigate the patterns of retail trading on anomalies in different windows around 10-K announcements. Table 8 shows the results of regression analysis using firm-year observations aligned at t=0, the date of a firm’s annual 10-K announcement. We calculate mean daily retail activity in each stock in three-month windows before its annual 10-K announcement, in the one-month window after the announcement, and in three-month windows afterwards. Then, for each of the time windows, we regress the dependent variable on two independent variables that measure the number of long anomaly portfolios and the number of short anomaly portfolios that the stock enters at the time of its 10-K announcement (t=0). Rather than using the average portfolio index or the extreme portfolios index, this analysis has the advantage of revealing any differences in retail activity between the long and the short anomaly portfolios. In addition, we only count the number of extreme portfolios that a stock enters at t=0 if there was a change in its portfolio ranking compared to the previous-year ranking. This controls for any momentum effect that may exist in firm fundamentals and ensures that our variables capture the arrival of new information to the market. The dependent variable in Panel A of Table 8 is the retail buy volume as a percentage of total retail volume. The results show that retail investors tend to buy stocks that enter more short portfolios, both before firm announcements and during the one month after announcement. They correct their trading afterwards and stop buying stocks in the short anomaly portfolios, but there is no evidence that they start buying stocks in the long portfolios. Panel B shows the patterns in retail sell volumes that close existing positions. Retail investors consistently sell stocks in the long anomaly portfolios during all time windows, both before and after firm announcements. Finally, Panel C looks at retail short sell volume as a percentage of total retail volume. Interestingly, we see that the negative relation between our composite anomaly indexes and retail short selling that we observed in previous analyses appears to be driven by these investors' activity in the long anomaly portfolios. They consistently avoid shorting stocks in the long deciles, but do not seem to short stocks in the short deciles.

[Table 8 about here]

The results in Table 6 showed us that retail investors trade differently on some categories of anomalies. In particular, retail short sellers were shown to trade correctly on all anomalies except for Profitability anomalies. Previous literature has shown that short sellers are informed investors and institutional short sellers short overvalued stocks in the bottom anomaly portfolios. For this reason, we want to investigate whether there is any category of anomalies where the same is true for retail short sellers as well. In Table 9, we repeat the analyses of Table 8 Panel C, but we show retail short sale activity in the four different anomaly categories separately. The categories include: Value-vs-growth anomalies, Investment anomalies, Profitability anomalies, and Intangibles anomalies (refer to the Appendix for details about the anomalies). We find that, after anomaly variables are released to the public, retail short sellers start to short stocks in the bottom portfolios of Value-versus-growth anomalies. This means that, at least for one of the anomaly categories, retail short sellers do act in a similar way to institutional short sellers. Their shorting of the overvalued stocks within those anomaly portfolios helps to correct stock mispricing and improve market efficiency.

[Table 9 about here]

5. Conclusion

In this study, we examine retail investor activity in the extreme portfolios of well-known cross-sectional anomalies. We contribute to the literature by showing how retail investors trade on publicly available information about accounting anomaly variables. We also investigate whether retail activity is the cause of stock mispricing and the subsequent return predictability associated with anomalies. In this way, the study provides important evidence on the relationship between retail trading and market efficiency.

We examine separately retail purchases, retail sales that close existing positions, and retail short sales. We examine the traditional June-rebalancing strategy, which consists of rebalancing anomaly portfolios only once a year at the end of June, as well as an announcement-rebalancing strategy, which consists of rebalancing portfolios each time when a firm releases its annual report containing the accounting anomaly variables. We find consistent evidence that retail investors tend to trade in the opposite direction of anomalies, buying stocks which should be shorted and selling stocks which should be bought according to composite anomaly indexes that combine the information from 25 individual anomalies. This is evident both before and after the anomaly variables become public information. Retail short sellers are the exception, as they trade in the correct direction of anomaly information. They consistently avoid shorting stocks in the long anomaly portfolios. In addition, after anomaly variables are released to the public, retail short sellers start to short overvalued stocks in the portfolios of value-versus-growth anomalies, thus helping to correct stock mispricing.

Retail investors are often pointed out as a group of investors susceptible to irrational biases and uninformed trading decisions, which can push stock prices away from fundamentals. Despite this widely held belief, we do not find evidence that retail trading is the cause of mispricing and subsequent return predictability. Retail order imbalances generally constitute only a small fraction of overall market activity, making it unlikely for retail trading to have a significant impact on market prices. Retail participation is considerably larger in small-capitalization stocks and retail investors may have an impact on the prices of those stocks. However, after controlling for firm size, there is no association between retail activity and mispricing of stocks within the extreme anomaly portfolios. Overall, our study shows that retail investors do not cause mispricing in anomaly portfolios, and in some situations they help to correct mispricing and therefore improve market efficiency.

Appendix

Below are details about the anomaly variables used in our analyses.

Anomaly		Reference publications
Value-versus-growth anomalies		
1.	Book-to-Market Equity (BM)	Fama, French (1993); Book equity computed following Davis, Fama, French (2000)
2.	Assets to Market Equity (AM)	Hou, Xue, Zhang (2015)
3.	Earnings to Price (EP)	Basu (1983)
4.	Cash Flow to Market Equity (CFP)	Hou, Xue, Zhang (2015)
5.	Payouts to Market Equity (POM)	Boudoukh et al. (2007)
6.	Net Payouts to Market Equity (NPOM)	Boudoukh et al. (2007)
Investment anomalies		
7.	Asset Growth (AG)	Cooper, Gulen, Schill (2008)
8.	Investment to Assets (IA)	Lyandres, Sun, Zhang (2008)
9.	Investment Growth (IG)	Xing (2008)
10.	Net Operating Assets (NOA)	Hirshleifer et al. (2004)
11.	Net Stock Issues (NSI)	Fama, French (2008)
12.	Net External Financing (NXF)	Bradshaw, Richardson, Sloan (2006)
13.	Inventory Changes (INV)	Thomas, Zhang (2002)
14.	Operating Accruals (OA)	Sloan (1996) for data before 1988; Hribar, Collins (2002) from 1988
15.	Percent Operating Accruals (POA)	Hafzalla, Lundholm, Van Winkle (2011)
Profitability anomalies		
16.	Gross Profitability (GP)	Novy-Marx (2013)
17.	Gross Margins (GM)	Novy-Marx (2013)
18.	Asset Turnover (ATRN)	Soliman (2008)
19.	Capital Turnover (CTRN)	Haugen, Baker (1996)
20.	Taxable Income to Book Income (TIBI)	Green, Hand, Zhang (2013)
Intangibles anomalies		
21.	Advertising Expenses to Market Equity (ADM)	Hou, Xue, Zhang (2015)
22.	R&D to Sales (RDS)	Hou, Xue, Zhang (2015)
23.	R&D to Market Equity (RDM)	Hou, Xue, Zhang (2015)
24.	Operating Leverage (OL)	Novy-Marx (2011)
25.	Hiring Rate (HR)	Belo, Lin, Bazdresch (2014)

References

Akbas, F., Armstrong, W. J., Sorescu, S., Subrahmanyam, A. (2015). Smart money, dumb money, and capital market anomalies. Journal of Financial Economics 118, p.355-382.

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

Barber, B. M., Odean, T., Zhu, N. (2009). Do Retail Trades Move Markets? The Review of Financial Studies 22(1), p.151-186.

Basu, S. (1983). The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence. Journal of Financial Economics 12, p.129-156.

Belo, F., Lin, X., Bazdresch, S. (2014). Labor Hiring, Investment, and Stock Return Predictability in the Cross Section. The Journal of Political Economy 122(1), p.129-177.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X. (2019). Tracking Retail Investor Activity. Working Paper.

Boudoukh, J., Michaely, R., Richardson, M., Roberts, M. R. (2007). On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. The Journal of Finance 62(2), p.877-915.

Bowles, B., Reed, A. V., Ringgenberg, M. C., Thornock, J. R. (2020). Anomaly Time. Working Paper.

Bradshaw, M. T., Richardson, S. A., Sloan, R. G. (2006). The relation between corporate financing activities, analysts' forecasts and stock returns. Journal of Accounting and Economics 42, p.53-85.

Calluzzo, P., Moneta, F., Topaloglu, S. (2019). When Anomalies Are Publicized Broadly, Do Institutions Trade Accordingly? Management Science 65 (10), p.4555-4574.

Chen, Y., Da, Z., Huang, D. (2019). Arbitrage Trading: The Long and the Short of It. The Review of Financial Studies 32(4), p.1608-1646.

Cooper, M. J., Gulen, H., Schill, M. J. (2008). Asset growth and the cross-section of stock returns. Journal of Finance 63(4), p.1609-1651.

Cui, X., Kolokolova, O., Wang, J. (2019). Hedge Funds and Quasi-Indexers: Cui Bono? Working Paper.

Daniel, K., Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. Journal of Finance 52(1), p.1-33.

Davis, J. L., Fama, E., French., K. (2000). Characteristics, Covariances, and Average Returns: 1929 to 1997. The Journal of Finance 55(1), p.389-406.

Edelen, R. M., Ince, O. S., Kadlec, G. B. (2016). Institutional investors and stock return anomalies. Journal of Financial Economics 119(3), p.472-488.

Engelberg, J., McLean, R. D., Pontiff, J. (2018). Anomalies and News. Journal of Finance 73(5), p.1971-2001.

Fama, E. F., French, K. R. (2008). Dissecting Anomalies. The Journal of Finance 63(4), p.1653-1678.

Fama, E., French, K. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33(1), p.3-56.

Gao, X., Wang, Y. (2019). Do Institutional Investors Exploit Market Anomalies? New Evidence from Alternative Mutual Funds. Working Paper.

Green, J., Hand, J. R. M, Zhang, X. F. (2013). The supraview of return predictive signals. Review of accounting studies, , Vol.18 (3), p.692-730

Hafzalla, N., Lundholm, R., Van Winkle, E. (2011). Percent accruals. The Accounting Review 86(1), p.209-236.

Haugen, R., Baker, N. (1996). Commonality in the determinants of expected stock returns. Journal of Financial Economics 41(3), p.401-439.

Hirshleifer, D., Hou, K., Teoh, S. H., Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, p.297-331.

Hirshleifer, D., Myers, J. N., Myers, L. A., Teoh, S. H. (2008). Do Individual Investors Cause Post-Earnings Announcement Drift? Direct Evidence From Personal Trades. Working Paper.

Hou, K., Xue, C., Zhang, L. (2015). Digesting Anomalies: An Investment Approach. The Review of Financial Studies 28(3), p.650-705.

Hribar, P., Collins., D. W. (2002). Errors in estimating accruals: Implications for empirical research. Journal of Accounting Research 40(1), p.105-134.

Jegadeesh, N., Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance 48 (1), p.65-91.

Kaniel, R., Liu, S., Saar, G., Titman, S. (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. The Journal of Finance 67(2), p.639-680.

Kaniel, R., Saar, G., Titman, S. (2008). Individual Investor Trading and Stock Returns. The Journal of Finance 63(1), p.273-310.

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

Kelley, E. K., Tetlock, P. C. (2017). Retail Short Selling and Stock Prices. The Review of Financial Studies 30(3), p.801-834.

Lee, C. M. C., Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. The Journal of Finance 46 (2), p.733-746.

Livnat, J., Mendenhall, R. R. (2006). Comparing the Post–Earnings Announcement Drift for Surprises Calculated from Analyst and Time Series Forecasts. Journal of Accounting Research 44(1), p.177-205.

Lyandres, E., Sun, L., Zhang, L. (2008). The New Issues Puzzle: Testing the Investment-Based Explanation. The Review of Financial Studies 21(6), p.2825-2855.

McLean, R. D., Pontiff, J., Reilly, C. (2019). The Trinity of Market Participants: Taking Sides on Return Predictability. Working Paper.

Miller, E. M. (1977). Risk, uncertainty and divergence of opinion. Journal of Finance 32(4), p.1151-1168.

Novy-Marx, R. (2011). Operating leverage. Review of Finance 15(1), p.103-134.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. Journal of Financial Economics 108(1), p.1-28.

Shleifer, A., Vishny, R. W. (1997). The Limits of Arbitrage. Journal of Finance 52(1), p.35-55.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71(3), p.289-315.

Soliman, M. T. (2008). The Use of DuPont Analysis by Market Participants. The Accounting Review 83(3), p.823-853.

Song, W. (2019). Essays in asset pricing. Dissertations and Theses Collection, Singapore Management University.

Stambaugh, R. F., Yu, J., Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, p.288-302.

Stambaugh, R., Yuan, Y. (2017). Mispricing Factors. The Review of Financial Studies 30(4), p.1270-1315.

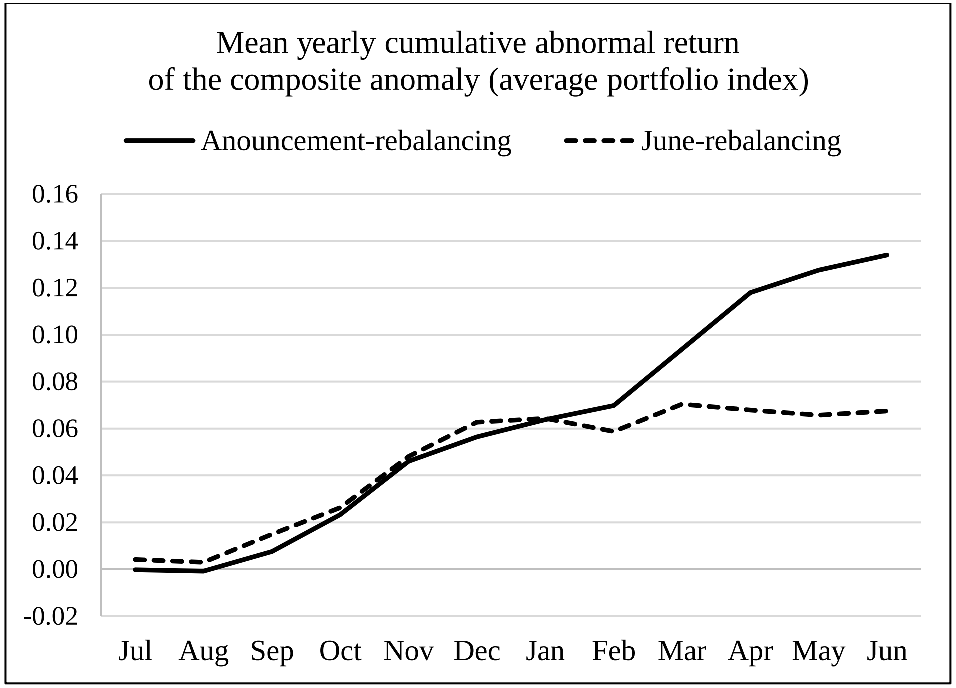
Thomas, J. K., Zhang, H. (2002). Inventory Changes and Future Returns. Review of Accounting Studies 7(2), p.163-187.

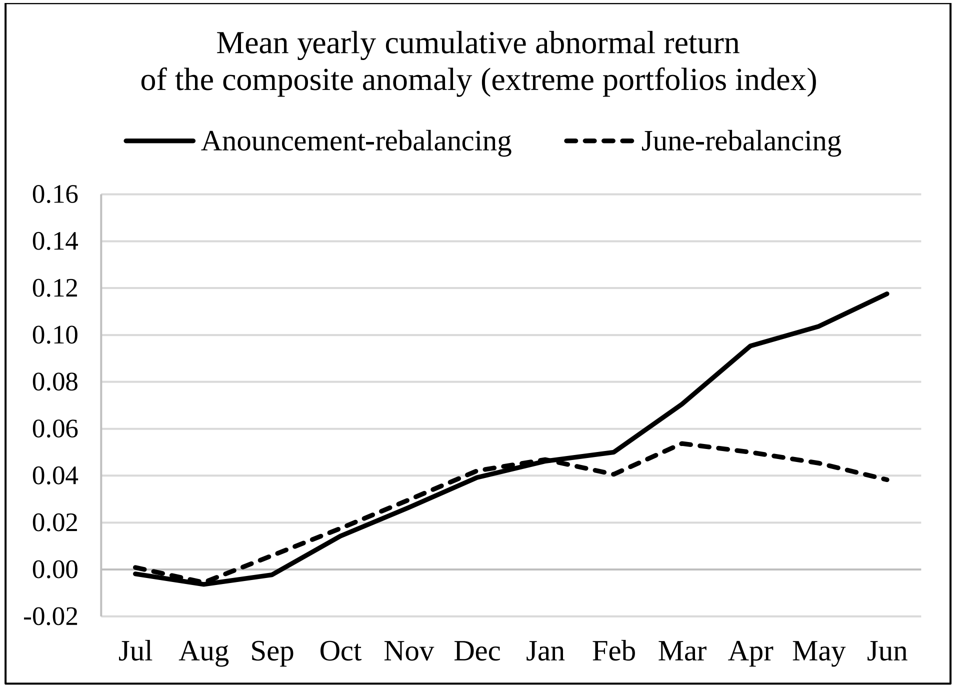
Xing, Y. (2008). Interpreting the Value Effect through the Q-Theory: An Empirical Investigation. The Review of Financial Studies 21(4), p.1767-1795.

Figure 1
Composite anomaly CARs with announcement-rebalancing versus June-rebalancing of portfolios

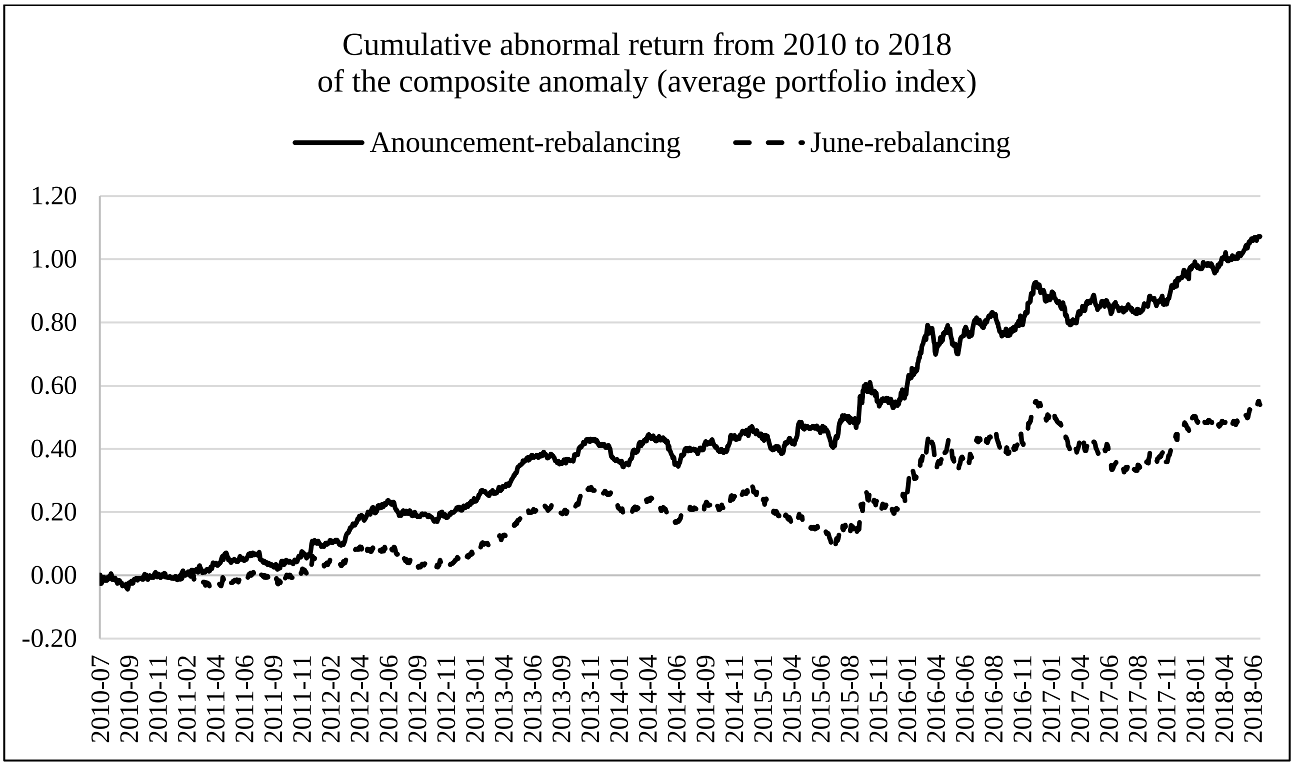
Figure 1 shows the difference in cumulative abnormal returns (CARs) when using an announcement-rebalancing strategy versus a June-rebalancing strategy to trade on the composite anomaly indexes. The announcement-rebalancing strategy consists of rebalancing the anomaly portfolios every time a firm files its annual 10-K report (usually between February and April). The June-rebalancing strategy follows the traditional anomaly literature where portfolios are sorted at the end of June of each year t based on accounting information from the fiscal year ending in calendar year t−1. The figures in Panel A plot the mean yearly CAR of the two composite anomaly indexes (abnormal returns are cumulated starting from 1 July of year t until 30 June of year t+1). The figures in Panel B plot the CAR earned with the two strategies over the whole sample period from 2010 to 2018. A stock's abnormal return is equal to its realized return minus expected return, where expected returns are calculated using the market model by regressing past one-year returns on market returns. The average portfolio index for a given stock is obtained by taking the equal-weighted average of the stock's rankings across 25 individual anomalies. The extreme portfolios index is equal to the number of long portfolios minus the number of short portfolios the stock belongs to.

Panel A: Mean yearly CARs of the composite anomaly indexes





Panel B: Total CARs of the composite anomaly indexes over the sample time period



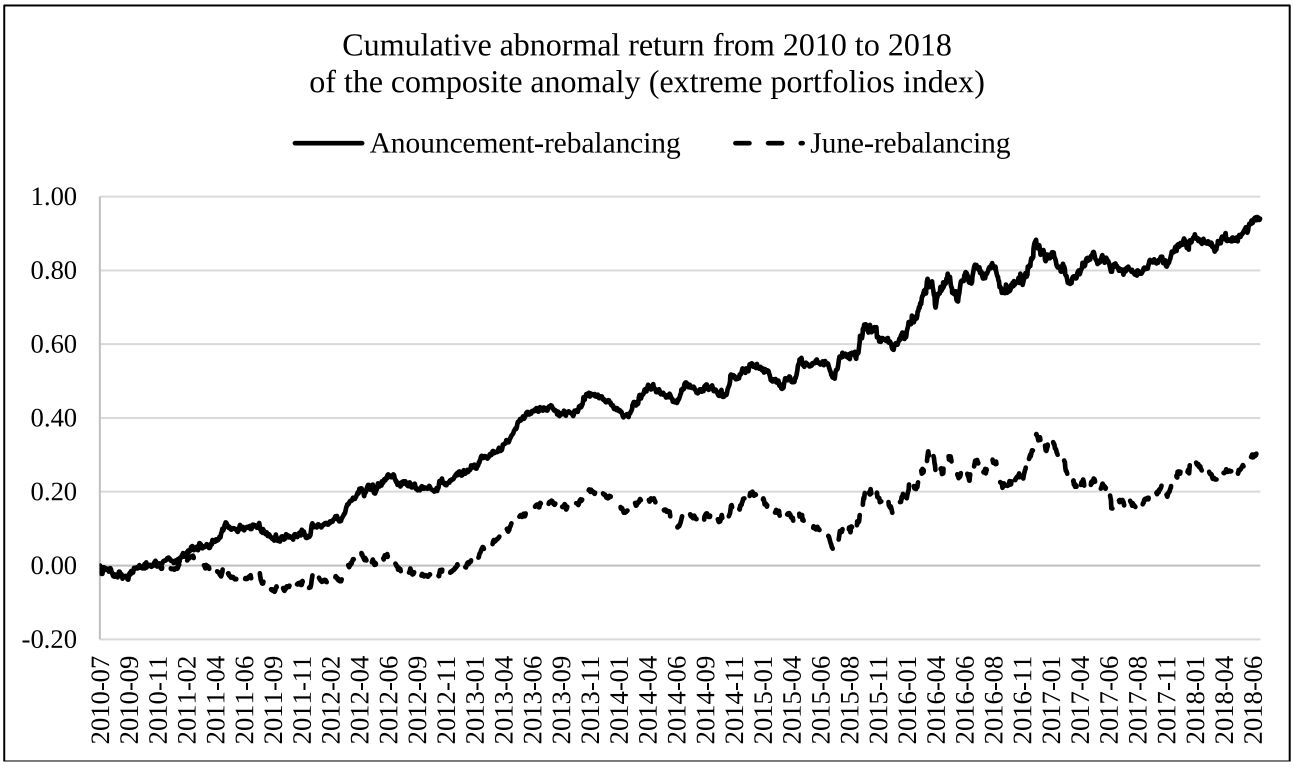


Figure 2
Retail trading activity in the composite anomaly portfolios

Figure 2 shows average retail trading volumes in decile portfolios (marked as SHORT, 2, 3, …, 8, 9, LONG from bottom to top portfolio) sorted on the composite anomaly index (here we use the average portfolio index, but results are similar if we use the extreme portfolios index). We calculate average retail buy, sell, and short sell volume as a percentage of total retail volume. First, we calculated the mean of a variable across stocks in each anomaly decile, and then we take the time-series mean over the sample period from 2010 to 2018.

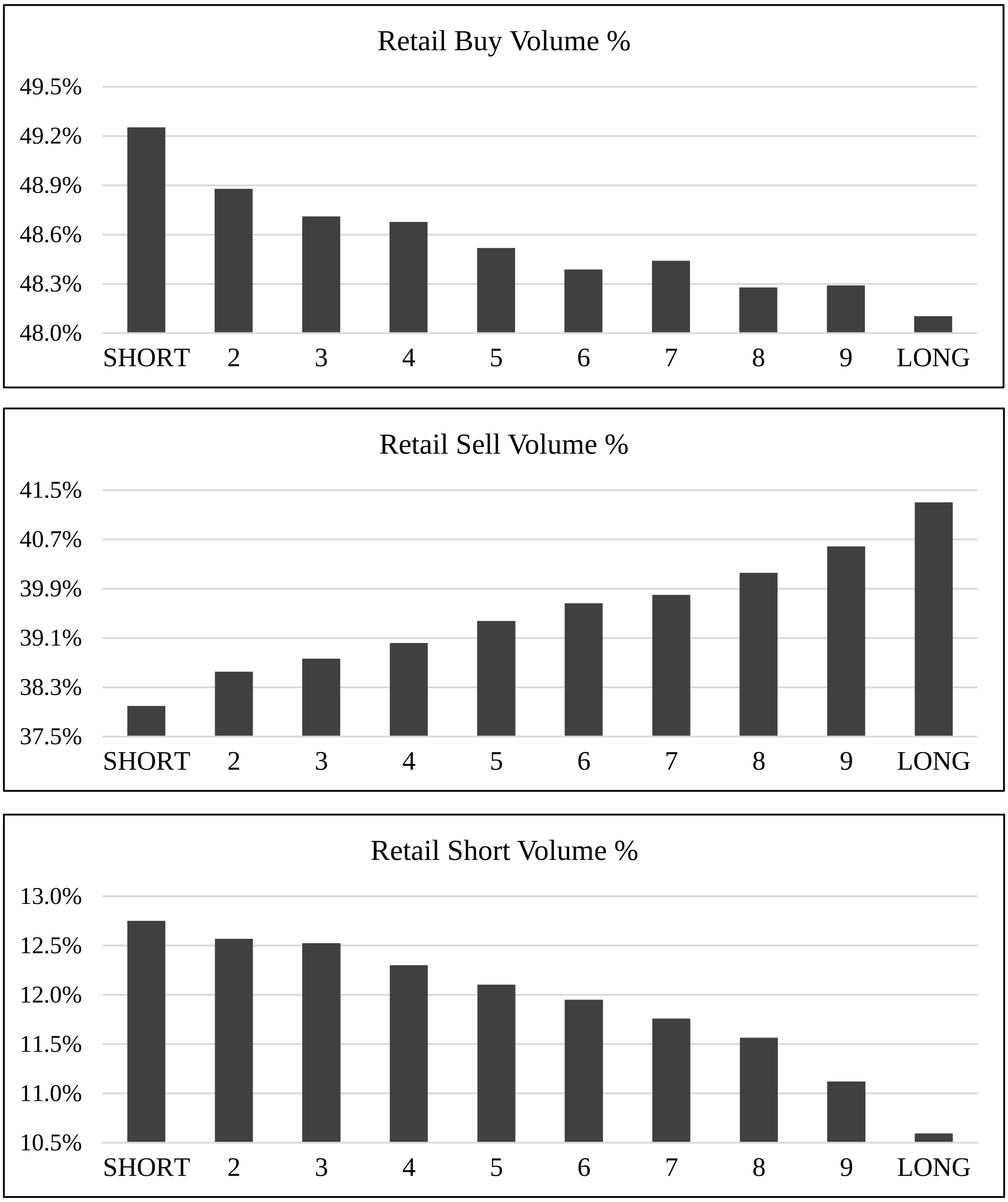


Table 1
Mean CARs of anomaly portfolios in windows around 10-K announcements

This table shows mean cumulative abnormal returns (CARs) of anomaly long-minus-short portfolios in monthly windows around firms' 10-K announcements. The anomaly portfolios are sorted in calendar time and rebalanced on the day after each announcement. We align all firms' 10-K announcement dates at t=0 and compute equal-weighted average CARs for each anomaly portfolio in each monthly window, starting from 10 months before the announcement and ending 10 months after. A stock's abnormal return is calculated as realized return minus expected return obtained using the market model with a three-year estimation window. Panel A reports portfolio CARs for the 25 individual anomalies (refer to the Appendix for details about each anomaly variable) and for the two composite anomaly indexes. The average portfolio index for a given stock is obtained by taking the equal-weighted average of the stock's rankings across the 25 anomalies. The extreme portfolios index is equal to the number of long portfolios minus the number of short portfolios the stock belongs to on a given day. Panel B reports CARs for the two composite anomaly indexes in subsamples of small-capitalization stocks and large-capitalization stocks. Panel C reports CARs in subsamples of stocks sorted by post-earnings-announcement drift (PEAD) portfolio. Quarterly standardized unexpected earnings (SUEs) are calculated using Compustat seasonal changes in earnings per share (EPS). Then, daily SUE-based portfolios are formed and held from the announcement day until the next announcement. We mark negative-SUE stocks as those in decile 1 or 2, zero-SUE stocks as those in decile 5 or 6, and positive-SUE stocks as those in decile 9 or 10. The sample time period is from 2010 to 2018. Numbers in **bold font** indicate that the result is statistically significant at the 5% level.

Panel A: Full sample

	(-210,-190)	(-189,-169)	(-168,-148)	(-147,-127)	(-126,-106)	(-105,-85)	(-84,-64)	(-63,-43)	(-42,-22)	(-21,-1)	(0,20)	(21,41)	(42,62)	(63,83)	(84,104)	(105,125)	(126,146)	(147,167)	(168,188)	(189,209)
Value-versus-growth anomalies																				
BM	**-2.82%**	**-1.44%**	**-2.55%**	**-1.40%**	**-1.58%**	**-1.02%**	**-3.86%**	**1.10%**	**2.53%**	**2.03%**	**3.49%**	**4.42%**	**1.04%**	**1.22%**	**2.22%**	**2.92%**	**2.53%**	**3.65%**	**1.60%**	**2.41%**
AM	**-2.45%**	**-2.11%**	**-1.46%**	**-1.61%**	**-1.89%**	**-1.42%**	**-3.11%**	**1.84%**	**1.93%**	**1.11%**	**4.28%**	**5.23%**	**1.34%**	0.60%	**3.17%**	**3.37%**	**2.53%**	**3.79%**	**2.38%**	**1.57%**
EP	-0.05%	-0.44%	-0.47%	**-1.44%**	**-2.06%**	-0.73%	**-2.59%**	-0.25%	**1.24%**	0.82%	**1.08%**	-0.58%	**-1.10%**	-0.63%	**-1.13%**	-0.34%	0.14%	0.53%	-0.71%	-0.29%
CFP	-0.45%	-0.62%	-0.56%	**-1.13%**	**-1.79%**	**-0.89%**	**-2.05%**	-0.45%	**1.50%**	0.72%	**1.89%**	**1.52%**	-0.22%	-0.01%	-0.27%	0.23%	0.34%	**1.14%**	0.32%	-0.17%
POM	**-1.03%**	-0.58%	**-1.29%**	**-1.96%**	-0.72%	**-2.06%**	**-1.42%**	-0.67%	-0.41%	0.01%	-0.17%	0.48%	-0.38%	0.23%	0.46%	**1.23%**	0.48%	**1.12%**	0.66%	-0.66%
NPOM	-0.51%	-0.75%	**-1.15%**	**-1.35%**	-0.06%	-0.79%	-0.91%	0.25%	0.52%	0.16%	-0.03%	0.56%	-0.39%	0.38%	0.54%	**0.99%**	0.55%	**1.84%**	0.76%	-0.60%
Investment anomalies																				
AG	**-2.02%**	**-1.42%**	**-2.27%**	-0.59%	0.76%	-0.35%	-0.55%	**2.36%**	**3.47%**	**1.96%**	**2.49%**	**2.75%**	**2.16%**	**1.72%**	**2.76%**	**2.80%**	**2.48%**	**2.86%**	**1.73%**	**2.91%**
IA	0.34%	0.46%	**1.22%**	**2.41%**	**2.44%**	**2.13%**	**1.88%**	**3.04%**	**2.58%**	**1.63%**	**4.22%**	**2.70%**	**1.89%**	**2.26%**	**1.75%**	**2.43%**	**1.76%**	**3.19%**	**2.48%**	**3.08%**
IG	0.16%	-0.09%	0.54%	**1.22%**	**1.43%**	**1.13%**	**0.97%**	**1.76%**	**2.02%**	**2.15%**	**1.70%**	**2.36%**	**1.15%**	**0.86%**	**1.46%**	**1.41%**	**1.63%**	**1.81%**	**1.64%**	**1.48%**
NOA	0.59%	0.03%	-0.34%	**1.71%**	**1.65%**	-0.79%	**0.94%**	**1.79%**	**2.53%**	**1.47%**	**1.08%**	-0.17%	**1.53%**	**1.66%**	0.72%	**1.51%**	**1.48%**	-0.28%	**0.85%**	**1.97%**
NSI	**-1.08%**	-0.85%	0.19%	-0.53%	-0.41%	**1.69%**	**1.64%**	0.03%	0.04%	**1.22%**	**1.52%**	**2.12%**	0.63%	0.67%	**1.91%**	0.20%	0.71%	**2.85%**	**1.60%**	0.01%
NXF	0.87%	**-1.04%**	0.64%	0.16%	-0.31%	**2.02%**	**1.79%**	-0.31%	-0.13%	0.94%	**2.27%**	**1.14%**	0.12%	0.27%	0.70%	-0.78%	-0.29%	**2.43%**	0.55%	0.12%
INV	-0.11%	-0.09%	-0.04%	**1.19%**	**1.31%**	0.23%	**1.86%**	**2.11%**	**2.00%**	**1.28%**	**3.32%**	**2.00%**	**1.20%**	**1.69%**	0.73%	**1.83%**	**0.95%**	**2.79%**	**0.93%**	**1.26%**
OA	0.46%	**0.76%**	**-1.04%**	**1.31%**	0.49%	-0.12%	0.24%	0.78%	**0.90%**	0.47%	**1.50%**	**1.38%**	**0.96%**	**1.16%**	0.25%	**1.14%**	0.86%	0.76%	0.07%	0.39%
POA	0.26%	-0.34%	-0.25%	0.53%	0.22%	0.33%	**0.89%**	0.09%	0.65%	0.41%	**1.71%**	**1.59%**	0.27%	**0.86%**	**0.76%**	0.01%	-0.09%	**2.03%**	0.81%	0.33%
Profitability anomalies																				
GP	-0.25%	-0.05%	0.56%	-0.47%	-0.13%	**1.68%**	0.07%	**-1.57%**	-0.82%	0.66%	**1.54%**	**1.44%**	-0.43%	0.78%	**1.24%**	-0.29%	**-1.11%**	**2.21%**	0.23%	-0.40%
GM	0.32%	0.29%	0.04%	-0.64%	0.35%	**1.41%**	0.43%	-0.57%	-0.46%	0.58%	**1.35%**	**1.42%**	-0.67%	**0.94%**	0.62%	-0.47%	0.38%	**2.09%**	-0.21%	0.20%
ATRN	-0.79%	**-1.07%**	0.42%	-0.83%	-0.37%	**1.62%**	0.54%	-0.20%	-0.41%	0.35%	**1.63%**	**1.91%**	-0.39%	-0.01%	**1.33%**	-0.41%	-0.60%	**2.24%**	**1.14%**	0.15%
CTRN	-0.46%	-0.64%	**1.31%**	-0.68%	-0.57%	**1.83%**	0.28%	-0.74%	**-1.24%**	-0.26%	0.84%	0.88%	**-1.13%**	-0.39%	0.37%	**-0.99%**	**-1.40%**	**1.03%**	0.69%	-0.42%
TIBI	**-0.91%**	-0.06%	-0.58%	-0.67%	-0.12%	-0.40%	0.10%	-0.48%	-0.67%	-0.36%	-0.36%	**0.92%**	0.30%	-0.15%	**0.95%**	0.27%	0.44%	0.29%	0.75%	-0.09%
Intangibles anomalies																				
ADM	**-1.36%**	**-2.35%**	**-1.68%**	0.18%	-0.46%	**-2.18%**	**-1.72%**	1.04%	**2.28%**	1.05%	**2.36%**	**3.68%**	0.57%	-0.09%	**2.05%**	**3.17%**	0.93%	**2.72%**	0.98%	0.95%
RDS	1.05%	**1.80%**	0.22%	0.66%	0.75%	**-2.29%**	-0.32%	0.56%	**1.59%**	1.42%	-0.94%	**-2.18%**	**1.60%**	**1.43%**	0.13%	0.84%	1.22%	**-2.25%**	0.40%	1.18%
RDM	0.22%	0.56%	-0.51%	0.33%	0.56%	**-2.93%**	-1.15%	**1.75%**	**3.92%**	**2.28%**	0.80%	0.83%	**1.83%**	**2.60%**	0.98%	**2.63%**	**2.98%**	-0.43%	1.41%	**2.52%**
OL	-0.45%	-0.28%	0.07%	-0.14%	0.46%	0.16%	0.52%	0.87%	**1.17%**	**1.61%**	**1.40%**	0.89%	0.07%	0.75%	0.74%	0.51%	0.36%	0.81%	**1.97%**	**1.31%**
HR	-0.06%	**-0.92%**	-0.59%	0.74%	**1.58%**	**1.10%**	0.65%	**3.07%**	**2.85%**	**2.09%**	**3.35%**	**3.48%**	**1.83%**	**1.53%**	**2.33%**	**1.95%**	**1.31%**	**3.44%**	**1.96%**	**3.98%**
Composite anomalies																				
Average portfolio index	**-1.30%**	**-1.25%**	**-1.33%**	0.09%	0.67%	-0.04%	-0.78%	**2.86%**	**3.46%**	**3.10%**	**4.92%**	**4.30%**	**1.11%**	**2.08%**	**2.84%**	**3.37%**	**2.49%**	**4.12%**	**2.51%**	**3.43%**
Extreme portfolios index	**-1.93%**	**-1.21%**	**-1.71%**	-0.06%	0.18%	0.66%	-0.50%	**2.33%**	**3.33%**	**2.87%**	**4.58%**	**4.08%**	**1.43%**	**2.59%**	**2.57%**	**2.70%**	**2.57%**	**4.24%**	**2.83%**	**2.88%**

Panel B: Subsamples by firm size

	(-210,-190)	(-189,-169)	(-168,-148)	(-147,-127)	(-126,-106)	(-105,-85)	(-84,-64)	(-63,-43)	(-42,-22)	(-21,-1)	(0,20)	(21,41)	(42,62)	(63,83)	(84,104)	(105,125)	(126,146)	(147,167)	(168,188)	(189,209)
Small-capitalization stocks																				
Average portfolio index	-0.09%	**-3.85%**	-1.66%	-2.34%	**3.83%**	1.67%	0.63%	1.09%	**6.90%**	**2.82%**	**7.09%**	**3.29%**	2.89%	-0.57%	**4.81%**	**3.70%**	2.25%	1.43%	**5.05%**	4.30%
Extreme portfolios index	-0.33%	-1.35%	-1.06%	-1.69%	2.04%	1.69%	0.64%	1.94%	**5.15%**	**3.60%**	**4.89%**	**3.15%**	**3.07%**	**2.60%**	**3.78%**	**3.62%**	**2.69%**	**3.38%**	**4.15%**	3.14%
Large-capitalization stocks																				
Average portfolio index	-0.40%	1.08%	-1.38%	0.87%	-0.10%	0.09%	0.19%	0.31%	**3.20%**	**4.03%**	1.44%	1.17%	0.33%	**1.75%**	1.18%	1.44%	1.08%	**3.41%**	1.73%	1.40%
Extreme portfolios index	-0.63%	0.65%	**-3.53%**	-0.27%	0.14%	0.81%	-1.21%	0.71%	**2.40%**	1.16%	**1.87%**	1.27%	0.62%	1.21%	1.11%	0.52%	1.02%	**3.80%**	**3.22%**	**1.51%**

Panel C: Subsamples by PEAD

	(-210,-190)	(-189,-169)	(-168,-148)	(-147,-127)	(-126,-106)	(-105,-85)	(-84,-64)	(-63,-43)	(-42,-22)	(-21,-1)	(0,20)	(21,41)	(42,62)	(63,83)	(84,104)	(105,125)	(126,146)	(147,167)	(168,188)	(189,209)
Negative-SUE stocks																				
Average portfolio index	-1.48%	-0.57%	**-3.47%**	-1.04%	1.51%	0.87%	0.43%	**3.86%**	1.74%	0.69%	**3.00%**	**3.06%**	2.01%	1.62%	**2.65%**	**3.76%**	**2.79%**	**5.05%**	**3.82%**	**3.80%**
Extreme portfolios index	**-2.09%**	-0.11%	**-2.01%**	-1.50%	0.63%	1.27%	0.12%	**2.12%**	1.86%	1.09%	**3.73%**	**2.54%**	1.47%	**1.85%**	1.51%	**3.56%**	1.59%	**4.32%**	**3.45%**	**3.21%**
Zero-SUE stocks																				
Average portfolio index	0.32%	-1.35%	-1.44%	-1.58%	-0.20%	1.85%	-0.80%	**2.57%**	1.64%	1.34%	**2.55%**	**2.18%**	0.29%	1.78%	1.74%	0.13%	-0.10%	**3.46%**	**3.30%**	**2.76%**
Extreme portfolios index	**-2.57%**	-1.85%	**-1.91%**	-2.43%	-0.57%	0.19%	-0.49%	0.10%	1.22%	0.71%	**3.25%**	0.90%	1.19%	**2.88%**	**2.34%**	0.93%	0.81%	**2.88%**	**2.92%**	1.30%
Positive-SUE stocks																				
Average portfolio index	-1.99%	-1.39%	-0.48%	0.04%	**2.69%**	1.24%	0.97%	1.75%	**4.51%**	**4.95%**	**5.53%**	**3.90%**	-0.41%	**2.43%**	**4.09%**	**2.37%**	**2.66%**	**3.69%**	**3.65%**	**3.77%**
Extreme portfolios index	-1.53%	-0.92%	-1.46%	0.28%	**1.83%**	1.53%	1.20%	**3.43%**	**3.15%**	**5.39%**	**3.80%**	**4.26%**	1.34%	**3.18%**	**3.75%**	1.53%	**2.18%**	**4.18%**	**3.42%**	**3.27%**

Table 2
Stock returns in windows around 10-K announcements

Table 2 reports the results of regression analysis testing the return patterns observed around 10-K announcements. We use firm-year observations aligned at t=0 which is the date of a firm's annual 10-K announcement. We calculate mean daily return of each stock in three-month windows before its annual 10-K announcement, in the one-month window after the announcement, and in three-month windows afterwards. Then, for each of the time windows, we regress the mean daily return on two independent variables that measure the number of long anomaly portfolios and the number of short anomaly portfolios that the stock enters at the time of its 10-K announcement (t=0). We include the following control variables. Momentum decile portfolios are created daily based on compounded returns over the past six months, skipping one month between the formation period and the holding period. We control for PEAD associated with quarterly earnings announcements by including a variable for the SUE portfolio that each stock belongs to. We also control for analysts' consensus (mean) recommendation from IBES, transforming the scale as follows: 2=Strong Buy, 1=Buy, 0=Hold, -1=Underperform, -2=Sell. If a stock has no analyst coverage at a given time, we set the consensus recommendation to zero. Institutional order imbalance is computed using all TAQ transactions excluding retail trades, signed with the Lee and Ready (1991) algorithm. The rest of the control variables include firm size (logarithm of lagged market capitalization), past-month mean daily return, past-month return volatility, and share turnover (total market volume divided by shares outstanding). Our sample includes the period from January 2010 to December 2018. The table contains the estimated regression coefficients and their corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Return, daily mean						
	(-210,-127)	(-126,-64)	(-63,-1)	(0,20)	(21,83)	(84,146)	(147,209)
Intercept	-0.096***	0.048	0.285***	-0.403***	-0.147***	-0.305***	0.076*
	(-2.92)	(1.37)	(7.83)	(-5.12)	(-4.30)	(-8.22)	(1.89)
Number of LONG portfolios entered at t=0	-0.022***	-0.023***	-0.005**	0.014***	0.005**	0.006***	0.007***
	(-12.47)	(-11.08)	(-2.20)	(2.98)	(2.44)	(2.85)	(2.88)
Number of SHORT portfolios entered at t=0	0.010***	0.013***	0.000	-0.010***	-0.001	-0.005***	0.000
	(6.39)	(7.19)	(0.16)	(-2.58)	(-0.68)	(-2.86)	(0.18)
Momentum portfolio, daily mean	0.023***	0.015***	0.012***	-0.004*	0.019***	0.020***	0.022***
	(22.00)	(13.00)	(10.21)	(-1.72)	(17.46)	(17.83)	(16.56)
PEAD (SUE) portfolio, daily mean	0.005***	0.008***	0.002**	0.018***	0.007***	0.009***	0.008***
	(5.49)	(7.08)	(2.16)	(8.53)	(6.84)	(7.97)	(6.23)
Consensus recommendation, daily mean	-0.006	-0.019***	-0.006	0.024**	0.005	0.001	-0.019***
	(-1.23)	(-3.69)	(-1.02)	(2.09)	(1.06)	(0.25)	(-3.09)
Institutional order imbalance, daily mean	0.000***	0.001***	0.000	0.000***	0.000***	0.001***	0.001***
	(3.32)	(5.20)	(1.04)	(3.55)	(3.51)	(7.30)	(5.32)
Log market cap, daily mean	-0.003*	-0.004**	-0.013***	0.017***	0.002	0.006***	-0.008***
	(-1.77)	(-2.46)	(-8.24)	(4.82)	(1.39)	(3.68)	(-4.29)
Past-month return	-0.059***	-0.042***	-0.037***	-0.052***	-0.027***	-0.063***	-0.002
	(-13.52)	(-9.62)	(-8.99)	(-4.97)	(-5.79)	(-11.47)	(-0.43)
Past-month return volatility	0.013***	0.005***	0.015***	0.018***	0.003	0.025***	0.001
	(7.69)	(3.03)	(8.27)	(4.92)	(1.53)	(12.82)	(0.66)
Share turnover, daily mean	0.000***	0.000***	0.000***	0.000***	0.000	0.000**	0.000***
	(3.17)	(6.94)	(-3.10)	(-2.62)	(-1.62)	(2.34)	(4.67)
N observations	15,862	16,353	16,672	17,004	18,835	18,760	18,420
Adjusted R^2	0.0605	0.0366	0.0181	0.0083	0.0228	0.0354	0.0234

Table 3
Summary statistics of retail trading activity

This table provides summary statistics of retail trading activity variables in the pooled sample of stocks from 2010 to 2018. We follow the method of Boehmer et al. (2019) to identify marketable retail buy and sell trades in the TAQ data, as well as retail short sales in the FINRA data. Retail trading volume refers to the number of lots traded by retail investors. Retail order imbalance (OI) is calculated as the difference between retail buy and sell volumes divided by the sum of retail buy and sell volumes. We compute separately retail buy volume, retail sell volume from sales that close existing positions, and retail short sell volume, as well as their percentage of total retail volume and total market volume. For each variable, we report number of stock-day observations, mean, standard deviation, minimum, median, and maximum. Panel A reports summary statistics in the full sample of all stocks. Panels B and C report statistics in the subsamples of small-capitalization stocks (bottom quintile by firm size) and large-capitalization stocks (top quintile by firm size), respectively.

Panel A: All stocks

	N stock-days	mean	std	min	median	max
Retail trading volume	5,016,004	98,132	499,889	1	15,489	116,980,418
% of total volume	5,016,004	9.3%	10.4%	0%	5.6%	100%
Retail OI	5,016,004	-0.03	0.41	-1.00	-0.02	1.00
Retail buy volume	5,016,004	49,093	255,338	0	7,264	57,667,598
% of retail volume	5,016,004	48.5%	20.7%	0%	49.1%	100%
% of total volume	5,016,004	4.5%	6.4%	0%	2.5%	100%
Retail sell volume	5,016,004	38,925	197,944	0	5,666	44,131,246
% of retail volume	5,016,004	39.8%	21.5%	0%	37.9%	100%
% of total volume	5,016,004	3.9%	6.1%	0%	2.0%	100%
Retail short sell volume	5,016,004	10,114	63,198	0	1,400	21,742,745
% of retail volume	5,016,004	11.7%	13.1%	0%	8.5%	100%
% of total volume	5,016,004	0.9%	2.2%	0.0%	0.4%	100%

Panel B: Small-capitalization stocks (bottom quintile by firm size)

	N stock-days	mean	std	min	median	max
Retail trading volume	1,002,304	41,617	263,119	1	5,372	48,312,308
% of total volume	1,002,304	21.4%	15.9%	0%	18.5%	100%
Retail OI	1,002,304	-0.04	0.61	-1.00	-0.03	1.00
Retail buy volume	1,002,304	20,888	136,062	0	2,300	21,730,003
% of retail volume	1,002,304	48.0%	30.5%	0%	48.4%	100%
% of total volume	1,002,304	10.3%	11.3%	0%	7.6%	100%
Retail sell volume	1,002,304	17,150	102,289	0	2,042	17,994,720
% of retail volume	1,002,304	44.4%	30.7%	0%	42.2%	100%
% of total volume	1,002,304	9.5%	11.0%	0%	6.6%	100%
Retail short sell volume	1,002,304	3,579	34,221	0	0	8,587,585
% of retail volume	1,002,304	7.6%	15.6%	0%	0%	100%
% of total volume	1,002,304	1.6%	4.3%	0%	0%	100%

Panel C: Large-capitalization stocks (top quintile by firm size)

	N stock-days	mean	std	min	median	max
Retail trading volume	1,004,114	279,473	947,236	1	77,143	116,980,418
% of total volume	1,004,114	5.3%	3.1%	0%	4.6%	60.8%
Retail OI	1,004,114	-0.02	0.22	-1.00	-0.01	1.00
Retail buy volume	1,004,114	139,735	483,954	0	37,244	57,667,598
% of retail volume	1,004,114	49.2%	11.0%	0%	49.5%	100%
% of total volume	1,004,114	2.6%	1.7%	0%	2.2%	55.6%
Retail sell volume	1,004,114	110,285	374,423	0	27,951	44,131,246
% of retail volume	1,004,114	37.8%	11.9%	0%	37.3%	100%
% of total volume	1,004,114	2.0%	1.4%	0%	1.7%	59.7%
Retail short sell volume	1,004,114	29,453	117,447	0	8,844	20,399,509
% of retail volume	1,004,114	13.0%	8.7%	0%	11.3%	100%
% of total volume	1,004,114	0.7%	0.7%	0.0%	0.5%	51.5%

Table 4
Do retail investors trade on anomalies using the June-rebalancing strategy?

This table shows the results of Fama-MacBeth regressions of retail buy, sell, and short sell volume as a percentage of total retail volume on the composite anomaly indexes, computed following the June-rebalancing strategy. Following the traditional anomaly literature, portfolios are sorted at the end of June of each year t, based on accounting information from the fiscal year ending in calendar year t−1, held for one year after that and rebalanced annually. We also report retail activity during the portfolio formation period, which for the purposes of this analysis is assumed to include the two weeks before the end of June and the two weeks at the start of July of each year. For each dependent variable and each period, the first column is a regression on the average portfolio index, and the second is a regression on the extreme portfolios index. The average portfolio index is equal to the equal-weighted average of the stock's rankings for the 25 individual anomalies. The extreme portfolios index is equal to the number of long portfolios minus the number of short portfolios the stock belongs to on that day. We include the following variables as controls. Momentum decile portfolios are created daily based on compounded returns over the past six months, skipping one month between the formation period and the holding period. We control for PEAD associated with quarterly earnings announcements by including a variable for the SUE portfolio that each stock belongs to. We also control for analysts' consensus (mean) recommendation from IBES, transforming the scale as follows: 2=Strong Buy, 1=Buy, 0=Hold, -1=Underperform, -2=Sell. If a stock has no analyst coverage at a given time, we set the consensus recommendation to zero. Institutional order imbalance is computed using all TAQ transactions excluding retail trades, signed with the Lee and Ready (1991) algorithm. The rest of the control variables include firm size (logarithm of lagged market capitalization), previous-day return, previous-month mean daily return, previous-month return volatility, and share turnover (total market volume divided by shares outstanding). Our sample includes the period from January 2010 to December 2018. The table contains the estimated regression coefficients and their corresponding t-statistics in parentheses. We apply Newey-West adjustment for the standard errors with 63 lags (equivalent to 3 months), or with 5 lags (equivalent to 1 week) for the regressions in the portfolio formation period. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Retail buy volume %				Retail sell volume %				Retail short sell volume %			
	Portfolio formation		Holding period		Portfolio formation		Holding period		Portfolio formation		Holding period	
Intercept	41.190***	41.397***	41.909***	42.046***	59.413***	59.678***	59.038***	59.320***	-0.603	-1.075	-0.946	-1.365
	(34.73)	(34.73)	(76.44)	(75.04)	(38.15)	(41.80)	(42.95)	(45.02)	(-0.45)	(-0.82)	(-0.78)	(-1.17)
Average portfolio index	-0.056***		-0.015		0.165***		0.122***		-0.110***		-0.107***	
	(-3.80)		(-0.91)		(6.29)		(5.61)		(-5.76)		(-8.90)	
Extreme portfolios index		-0.070***		-0.027***		0.150***		0.106***		-0.079***		-0.079***
		(-4.97)		(-2.86)		(7.85)		(5.79)		(-5.54)		(-6.54)
Momentum portfolio	-0.018	-0.016	-0.052***	-0.068***	-0.012	-0.014	-0.004	0.016	0.031*	0.030	0.056***	0.051***
	(-0.80)	(-0.70)	(-4.31)	(-4.71)	(-0.50)	(-0.54)	(-0.21)	(0.88)	(1.67)	(1.56)	(5.32)	(5.16)
PEAD (SUE) portfolio	0.049**	0.049**	0.059***	0.057***	0.011	0.012	0.014	0.012	-0.060***	-0.061***	-0.073***	-0.068***
	(2.34)	(2.28)	(5.86)	(5.46)	(0.58)	(0.65)	(1.00)	(0.83)	(-3.30)	(-3.32)	(-8.43)	(-7.24)
Consensus recommendation	0.177*	0.183*	0.330***	0.334***	-0.079	-0.133	-0.241***	-0.285***	-0.098	-0.050	-0.089*	-0.049
	(1.93)	(1.85)	(9.21)	(9.16)	(-0.50)	(-0.77)	(-3.38)	(-3.81)	(-1.03)	(-0.51)	(-1.80)	(-0.95)
Institutional order imbalance	-0.001	-0.001	-0.002	-0.003	-0.031***	-0.031***	-0.031***	-0.032***	0.032***	0.032***	0.033***	0.034***
	(-0.15)	(-0.14)	(-0.48)	(-0.59)	(-6.01)	(-6.02)	(-6.83)	(-7.29)	(12.88)	(12.95)	(12.00)	(20.26)
Log market cap (t-1)	0.324***	0.319***	0.295***	0.295***	-0.988***	-0.996***	-0.995***	-1.005***	0.663***	0.677***	0.700***	0.710***
	(6.00)	(5.84)	(11.67)	(11.93)	(-15.52)	(-16.36)	(-18.73)	(-19.5)	(11.18)	(11.50)	(14.49)	(14.87)
Return (t-1)	0.058**	0.058**	0.060***	0.060***	-0.110***	-0.110***	-0.090***	-0.089***	0.052***	0.052***	0.030***	0.029***
	(2.32)	(2.29)	(4.66)	(4.64)	(-4.66)	(-4.61)	(-7.80)	(-7.58)	(4.00)	(3.96)	(3.22)	(3.16)
Past-month return	-1.517***	-1.513***	-1.164***	-1.067***	0.562***	0.569***	0.466***	0.448***	0.955***	0.944***	0.698***	0.619***
	(-9.95)	(-9.94)	(-10.87)	(-8.59)	(2.89)	(2.92)	(3.87)	(3.46)	(11.97)	(11.72)	(7.18)	(14.87)
Past-month return volatility	0.280***	0.279***	0.145***	0.141***	0.136*	0.127	0.162***	0.143***	-0.416***	-0.406***	-0.307***	-0.285***
	(5.83)	(5.78)	(6.19)	(5.83)	(1.67)	(1.58)	(4.39)	(4.94)	(-8.41)	(-8.45)	(-10.03)	(-12.45)
Share turnover	0.000	0.000	0.012**	0.014**	0.000*	0.000*	-0.027**	-0.025**	-0.001**	-0.001**	0.015**	0.011*
	(0.32)	(0.27)	(2.26)	(2.07)	(1.73)	(1.74)	(-2.16)	(-2.22)	(-2.07)	(-2.04)	(1.96)	(1.89)
N stock-day observations	250,940	250,940	4,471,114	4,471,114	250,940	250,940	4,471,114	4,471,114	250,940	250,940	4,471,114	4,471,114
Average adjusted R^2	0.0059	0.0058	0.0079	0.0077	0.0178	0.0176	0.0203	0.0199	0.0292	0.0288	0.0347	0.0347

Table 5
Do retail investors trade on anomalies using the announcement-rebalancing strategy?

This table contains the results of Fama-MacBeth regressions of retail buy, sell, and short sell volume as a percentage of total retail volume on the composite anomaly indexes, computed following the announcement-rebalancing strategy. Each time when a firm files its annual 10-K report, the anomaly portfolios are rebalanced on the next day. For each dependent variable, the first column is a regression on the average portfolio index, and the second is a regression on the extreme portfolios index. The average portfolio index is equal to the equal-weighted average of the stock's rankings for the 25 individual anomalies. The extreme portfolios index is equal to the number of long portfolios minus the number of short portfolios the stock belongs to on that day. We include the same control variables as in Table 4. We use the full sample of stock-day observations from January 2010 to December 2018. The table contains the estimated regression coefficients and their corresponding t-statistics in parentheses. We apply Newey-West adjustment for the standard errors with 63 lags (63 trading days equates to 3 months). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Retail buy volume %		Retail sell volume %		Retail short sell volume %	
	1	2	3	4	5	6
Intercept	42.421***	42.178***	59.170***	59.836***	-1.590	-2.013*
	(47.91)	(63.99)	(40.71)	(45.65)	(-1.36)	(-1.71)
Average portfolio index	-0.079*		0.201***		-0.122***	
	(-1.96)		(4.56)		(-8.97)	
Extreme portfolios index		-0.061***		0.132***		-0.071**
		(-3.89)		(6.65)		(-2.23)
Momentum portfolio	-0.143*	-0.080***	0.098	0.014	0.046***	0.067***
	(-1.78)	(-3.97)	(1.10)	(0.82)	(2.83)	(3.30)
PEAD (SUE) portfolio	0.053***	0.053***	-0.012	0.006	-0.041	-0.059***
	(4.82)	(4.76)	(-0.37)	(0.33)	(-1.20)	(-3.25)
Consensus recommendation	0.352***	0.354***	-0.233***	-0.282***	-0.119**	-0.072
	(9.50)	(9.29)	(-3.39)	(-3.94)	(-2.47)	(-1.45)
Institutional order imbalance	-0.003	-0.002	-0.038***	-0.031***	0.041***	0.033***
	(-0.70)	(-0.45)	(-5.17)	(-6.82)	(6.29)	(14.34)
Log market cap (t-1)	0.303***	0.301***	-1.021***	-1.027***	0.717***	0.726***
	(12.27)	(11.97)	(-19.63)	(-19.67)	(15.23)	(15.41)
Return (t-1)	0.057***	0.058***	-0.086***	-0.087***	0.029***	0.029***
	(4.40)	(4.62)	(-7.29)	(-7.64)	(3.19)	(3.20)
Past-month return	-0.999***	-1.004***	0.846***	0.752***	0.153	0.252
	(-6.61)	(-6.87)	(2.59)	(3.24)	(0.36)	(0.78)
Past-month return volatility	0.146***	0.152***	0.066	0.091***	-0.212***	-0.243***
	(4.78)	(5.6)	(1.23)	(3.02)	(-3.06)	(-6.75)
Share turnover	0.016*	0.013*	-0.003	-0.022**	-0.013	0.009
	(1.73)	(1.95)	(-0.18)	(-2.28)	(-0.56)	(1.60)
N stock-day observations	4,590,012	4,590,012	4,590,012	4,590,012	4,590,012	4,590,012
Average adjusted R^2	0.0076	0.0075	0.0206	0.0204	0.0357	0.0354

Table 6
Retail trading on different categories of anomalies

This table repeats the Fama-MacBeth regression analyses of Table 5, but examines retail trading on the four different categories of anomalies: Value-vs-growth anomalies, Investment anomalies, Profitability anomalies, and Intangibles anomalies (refer to the Appendix for details about the anomalies). For each category, we compute an average portfolio index as the equal-weighted average of a stock's rankings for all of the individual anomalies within that category. All anomaly portfolios are created following the announcement-rebalancing strategy. The other variables are the same as in Table 5. The table contains the estimated regression coefficients and their corresponding t-statistics in parentheses. We apply Newey-West adjustment for the standard errors with 63 lags. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Retail buy volume %	Retail sell volume %	Retail short sell volume %
Intercept	42.060***	59.581***	-1.642
	(69.43)	(39.81)	(-1.23)
Average portfolio index (Value-vs-growth anomalies)	0.029	0.191***	-0.220***
	(1.54)	(7.22)	(-21.75)
Average portfolio index (Investment anomalies)	-0.105**	0.209***	-0.104***
	(-2.49)	(3.79)	(-4.79)
Average portfolio index (Profitability anomalies)	0.036***	-0.370***	0.334***
	(3.41)	(-29.63)	(21.32)
Average portfolio index (Intangibles anomalies)	-0.083***	0.327***	-0.245***
	(-8.50)	(20.34)	(-15.35)
Momentum portfolio	-0.066***	0.034**	0.032***
	(-6.45)	(2.16)	(3.38)
PEAD (SUE) portfolio	0.059***	0.008	-0.067***
	(5.80)	(0.61)	(-9.45)
Consensus recommendation	0.379***	-0.139**	-0.240***
	(10.70)	(-2.16)	(-5.33)
Institutional order imbalance	0.000	-0.033***	0.033***
	(-0.02)	(-7.22)	(18.89)
Log market cap (t-1)	0.312***	-1.073***	0.761***
	(12.37)	(-20.07)	(15.12)
Return (t-1)	0.061***	-0.091***	0.030***
	(4.94)	(-8.24)	(3.36)
Past-month return	-1.135***	0.597***	0.538***
	(-11.56)	(6.39)	(11.43)
Past-month return volatility	0.212***	-0.001	-0.212***
	(6.13)	(-0.02)	(-9.27)
Share turnover	0.013**	-0.024**	0.012**
	(2.07)	(-2.32)	(2.00)
N stock-day observations	4,573,285	4,573,285	4,573,285
Average adjusted R^2	0.0084	0.0242	0.0420

Table 7
Are stocks with high retail participation before firm announcements more mispriced?

Table 7 examines the relationship between retail participation (retail volume as a % of total market volume) before 10-K announcements and stock mispricing. We use firm-year observations aligned at t=0 which is the date of a firm's annual 10-K announcement. We regress mean daily return of each stock in the windows (-210,- 64) and (0,20) on the composite anomaly indexes at the time of announcement (t=0). To capture any effect of retail activity on returns, we add an interaction with the dummy variable *High retail participation (-210,-64)*, which is equal to one if retail participation during the period (-210,-64) is above the median retail participation computed within each firm size quintile separately, to control for the effect of firm size on retail participation. Refer to Table 2 for details about the control variables.

Dependent variable:	Return, daily mean (-210,-64)				Return, daily mean (0,20)			
Sample:	All stocks		Small-cap stocks		All stocks		Small-cap stocks	
	1	2	3	4	5	6	7	8
Intercept	-0.065***	-0.073***	1.536***	1.504***	-0.190*	-0.192*	-1.543**	-1.646***
	(-2.68)	(-3.06)	(9.86)	(9.74)	(-1.73)	(-1.78)	(-2.45)	(-2.64)
Average portfolio index at t=0	-0.006***		-0.006**		0.011***		0.022*	
	(-6.68)		(-2.00)		(3.24)		(1.89)	
Average portfolio index at t=0 * High retail participation (-210,-64)	-0.001		-0.004		-0.003		-0.013	
	(-0.46)		(-1.03)		(-0.76)		(-0.80)	
Extreme portfolios index at t=0		-0.005***		-0.008**		0.011***		0.026**
		(-6.11)		(-2.47)		(3.32)		(2.38)
Extreme portfolios index at t=0 * High retail participation (-210,-64)		-0.001		0.000		-0.002		-0.012
		(-0.85)		(0.06)		(-0.36)		(-0.81)
High retail participation (-210,-64)	0.041***	0.043***	0.023	-0.005	-0.008	-0.016	-0.077	-0.072
	(5.40)	(5.49)	(0.75)	(-0.15)	(-0.29)	(-0.54)	(-0.67)	(-0.62)
Retail participation, daily mean (0,20)					-0.584***	-0.588***	-0.158	-0.143
					(-4.61)	(-4.64)	(-0.47)	(-0.43)
Control variables	YES	YES	YES	YES	YES	YES	YES	YES
N observations	16,095	16,095	3,091	3,091	16,897	16,897	3,312	3,312
Adjusted R^2	0.1952	0.1952	0.1689	0.1679	0.0108	0.0111	0.0173	0.0183

Table 8
Retail trading on anomalies in windows around 10-K announcements

Table 8 shows the results of regression analysis using firm-year observations aligned at t=0 which is the date of a firm's annual 10-K announcement. We calculate mean daily retail activity in each stock in three-month windows before its annual 10-K announcement, in the one-month window after the announcement, and in three-month windows afterwards. Then, for each of the time windows, we regress the dependent variable on two independent variables that measure the number of long anomaly portfolios and the number of short anomaly portfolios that the stock enters at the time of its 10-K announcement (t=0). The dependent variable in Panel A / Panel B / Panel C is the daily mean of retail buy volume / retail sell volume that closes existing positions / retail short sell volume, as a percentage of total retail volume. We include the following control variables. Momentum decile portfolios are created daily based on compounded returns over the past six months, skipping one month between the formation period and the holding period. We control for PEAD associated with quarterly earnings announcements by including a variable for the SUE portfolio that each stock belongs to. We also control for analysts' consensus (mean) recommendation from IBES, transforming the scale as follows: 2=Strong Buy, 1=Buy, 0=Hold, -1=Underperform, -2=Sell. If a stock has no analyst coverage at a given time, we set the consensus recommendation to zero. Institutional order imbalance is computed using all TAQ transactions excluding retail trades, signed with the Lee and Ready (1991) algorithm. The rest of the control variables include firm size (logarithm of lagged market capitalization), past-month return, past-month return volatility, and share turnover (total market volume divided by shares outstanding). Our sample includes the period from January 2010 to December 2018. The table contains the estimated regression coefficients and their corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Retail buying

Dependent variable:	Retail buy volume %, daily mean						
	(-210,-127)	(-126,-64)	(-63,-1)	(0,20)	(21,83)	(84,146)	(147,209)
Intercept	40.666***	41.640***	42.554***	41.494***	40.245***	40.825***	40.937***
	(95.55)	(99.22)	(100.7)	(69.49)	(91.33)	(92.02)	(94.02)
Number of LONG portfolios entered at t=0	-0.038	-0.031	-0.064***	-0.041	-0.057**	-0.044*	-0.089***
	(-1.63)	(-1.28)	(-2.60)	(-1.16)	(-2.28)	(-1.71)	(-3.39)
Number of SHORT portfolios entered at t=0	0.107***	0.099***	0.070***	0.124***	0.035	-0.015	-0.035
	(5.24)	(4.60)	(3.24)	(4.05)	(1.60)	(-0.70)	(-1.55)
Momentum portfolio, daily mean	-0.075***	-0.031**	-0.056***	-0.086***	-0.107***	-0.105***	-0.076***
	(-5.56)	(-2.29)	(-4.11)	(-4.90)	(-7.87)	(-7.75)	(-5.41)
PEAD (SUE) portfolio, daily mean	0.058***	0.090***	0.069***	0.065***	0.066***	0.052***	0.089***
	(4.52)	(6.71)	(5.66)	(4.06)	(4.91)	(3.76)	(6.41)
Consensus recommendation, daily mean	0.134**	0.184***	0.139**	0.075	0.498***	0.561***	0.452***
	(2.26)	(2.92)	(2.20)	(0.85)	(7.83)	(8.63)	(6.77)
Institutional order imbalance, daily mean	-0.001	0.005***	-0.001	-0.002**	-0.001	-0.001	0.001
	(-0.85)	(3.99)	(-1.05)	(-2.40)	(-0.65)	(-0.45)	(0.89)
Log market cap, daily mean	0.363***	0.298***	0.269***	0.319***	0.376***	0.359***	0.336***
	(19.32)	(15.93)	(14.36)	(11.99)	(19.23)	(18.14)	(17.10)
Past-month return	-0.159***	-0.165***	-0.071	-0.481***	-0.456***	-0.497***	-0.336***
	(-2.82)	(-3.10)	(-1.48)	(-6.10)	(-7.66)	(-7.57)	(-5.61)
Past-month return volatility	0.120***	0.053***	0.097***	0.158***	0.138***	0.056**	0.021
	(5.38)	(2.67)	(4.74)	(5.57)	(6.25)	(2.38)	(1.09)
Share turnover, daily mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(-1.27)	(0.43)	(-0.48)	(-0.86)	(-0.25)	(-0.13)	(0.60)
N observations	15,862	16,353	16,672	17,004	18,835	18,760	18,420
Adjusted R^2	0.0272	0.0220	0.0154	0.0122	0.0296	0.0288	0.0257

Panel B: Retail selling that closes existing positions

Dependent variable:	Retail sell volume %, daily mean						
	(-210,-127)	(-126,-64)	(-63,-1)	(0,20)	(21,83)	(84,146)	(147,209)
Intercept	64.312***	65.933***	63.471***	65.083***	65.596***	64.688***	65.596***
	(104.3)	(110.4)	(105.3)	(85.7)	(106.0)	(105.0)	(106.8)
Number of LONG portfolios entered at t=0	0.234***	0.237***	0.270***	0.258***	0.262***	0.260***	0.305***
	(6.99)	(6.78)	(7.64)	(5.74)	(7.41)	(7.31)	(8.27)
Number of SHORT portfolios entered at t=0	-0.143***	-0.114***	-0.026	-0.091**	-0.012	-0.008	0.049
	(-4.86)	(-3.71)	(-0.85)	(-2.33)	(-0.39)	(-0.26)	(1.53)
Momentum portfolio, daily mean	-0.008	-0.033*	-0.024	0.082***	0.054***	0.035*	-0.002
	(-0.40)	(-1.70)	(-1.24)	(3.68)	(2.84)	(1.87)	(-0.09)
PEAD (SUE) portfolio, daily mean	0.040**	-0.001	0.017	-0.023	0.008	0.041**	-0.008
	(2.15)	(-0.04)	(0.99)	(-1.13)	(0.43)	(2.13)	(-0.40)
Consensus recommendation, daily mean	-0.015	-0.207**	-0.173*	-0.072	-0.408***	-0.534***	-0.455***
	(-0.17)	(-2.31)	(-1.92)	(-0.64)	(-4.58)	(-5.92)	(-4.83)
Institutional order imbalance, daily mean	-0.002**	-0.012***	-0.002	0.001	-0.003***	-0.006***	-0.008***
	(-2.01)	(-6.38)	(-1.39)	(0.66)	(-2.80)	(-3.80)	(-4.56)
Log market cap, daily mean	-1.211***	-1.255***	-1.134***	-1.227***	-1.265***	-1.239***	-1.254***
	(-44.52)	(-47.11)	(-42.39)	(-36.26)	(-46.05)	(-45.07)	(-45.21)
Past-month return	0.115	0.403***	0.094	0.440***	0.335***	0.692***	0.625***
	(1.41)	(5.33)	(1.39)	(4.39)	(4.01)	(7.59)	(7.39)
Past-month return volatility	0.138***	-0.005	-0.121***	-0.133***	0.024	0.132***	0.040
	(4.27)	(-0.18)	(-4.17)	(-3.67)	(0.78)	(4.04)	(1.48)
Share turnover, daily mean	0.000**	0.000	0.000*	0.000***	0.000	0.000	0.000
	(2.45)	(0.07)	(1.95)	(2.61)	(1.12)	(0.61)	(0.41)
N observations	15,862	16,353	16,672	17,004	18,835	18,760	18,420
Adjusted R^2	0.1658	0.1582	0.1214	0.0856	0.1382	0.1452	0.1403

Panel C: Retail short selling

Dependent variable:	Retail short sell volume %, daily mean						
	(-210,-127)	(-126,-64)	(-63,-1)	(0,20)	(21,83)	(84,146)	(147,209)
Intercept	-4.978***	-7.573***	-6.025***	-6.577***	-5.841***	-5.514***	-6.533***
	(-11.04)	(-17.47)	(-13.94)	(-13.15)	(-13.71)	(-12.99)	(-15.26)
Number of LONG portfolios entered at t=0	-0.196***	-0.206***	-0.205***	-0.217***	-0.204***	-0.216***	-0.216***
	(-8.02)	(-8.11)	(-8.12)	(-7.33)	(-8.40)	(-8.82)	(-8.42)
Number of SHORT portfolios entered at t=0	0.037*	0.015	-0.044**	-0.033	-0.023	0.023	-0.014
	(1.70)	(0.66)	(-1.99)	(-1.30)	(-1.09)	(1.10)	(-0.63)
Momentum portfolio, daily mean	0.083***	0.064***	0.080***	0.004	0.053***	0.070***	0.078***
	(5.79)	(4.56)	(5.75)	(0.27)	(4.02)	(5.38)	(5.63)
PEAD (SUE) portfolio, daily mean	-0.098***	-0.089***	-0.087***	-0.042***	-0.074***	-0.093***	-0.081***
	(-7.21)	(-6.43)	(-6.91)	(-3.13)	(-5.71)	(-7.02)	(-5.94)
Consensus recommendation, daily mean	-0.119*	0.023	0.034	-0.003	-0.090	-0.026	0.003
	(-1.90)	(0.36)	(0.52)	(-0.05)	(-1.46)	(-0.42)	(0.04)
Institutional order imbalance, daily mean	0.003***	0.007***	0.002***	0.001*	0.004***	0.007***	0.007***
	(3.55)	(4.93)	(2.96)	(1.86)	(4.73)	(5.98)	(5.63)
Log market cap, daily mean	0.848***	0.957***	0.864***	0.908***	0.889***	0.880***	0.918***
	(42.63)	(49.51)	(45.09)	(40.73)	(46.99)	(46.47)	(47.48)
Past-month return	0.044	-0.239***	-0.024	0.041	0.121**	-0.195***	-0.288***
	(0.73)	(-4.35)	(-0.49)	(0.62)	(2.10)	(-3.10)	(-4.89)
Past-month return volatility	-0.258***	-0.048**	0.025	-0.026	-0.162***	-0.189***	-0.060***
	(-10.91)	(-2.34)	(1.18)	(-1.08)	(-7.60)	(-8.35)	(-3.24)
Share turnover, daily mean	0.000**	0.000	0.000**	0.000***	0.000	0.000	0.000
	(-2.15)	(-0.52)	(-2.25)	(-2.94)	(-1.37)	(-0.75)	(-1.21)
N observations	15,862	16,353	16,672	17,004	18,835	18,760	18,420
Adjusted R^2	0.1885	0.1821	0.1500	0.1186	0.1675	0.1719	0.1613

Table 9
Retail short selling in different categories of anomalies

In this table, we repeat the analyses of Table 8 Panel C, but we show retail short sale activity in the four different anomaly categories separately. The categories include: Value-vs-growth anomalies, Investment anomalies, Profitability anomalies, and Intangibles anomalies (refer to the Appendix for details about the anomalies).

Dependent variable:	Retail short sell volume %, daily mean			
	(0,20)	(21,83)	(84,146)	(147,209)
Intercept	-6.513***	-5.860***	-5.559***	-6.564***
	(-12.9)	(-13.65)	(-13.01)	(-15.22)
Number of LONG portfolios entered at t=0 (Value-vs-growth anomalies)	-0.222***	-0.203***	-0.253***	-0.232***
	(-3.21)	(-3.47)	(-4.29)	(-3.73)
Number of LONG portfolios entered at t=0 (Investment anomalies)	-0.091**	-0.150***	-0.155***	-0.165***
	(-1.99)	(-4.04)	(-4.15)	(-4.19)
Number of LONG portfolios entered at t=0 (Profitability anomalies)	-0.221*	-0.038	0.086	0.032
	(-1.75)	(-0.37)	(0.82)	(0.29)
Number of LONG portfolios entered at t=0 (Intangibles anomalies)	-0.752***	-0.532***	-0.569***	-0.514***
	(-6.64)	(-5.66)	(-5.97)	(-5.08)
Number of SHORT portfolios entered at t=0 (Value-vs-growth anomalies)	0.087	0.159***	0.230***	0.146**
	(1.11)	(2.60)	(3.71)	(2.23)
Number of SHORT portfolios entered at t=0 (Investment anomalies)	0.002	0.007	0.046	0.031
	(0.05)	(0.22)	(1.52)	(0.99)
Number of SHORT portfolios entered at t=0 (Profitability anomalies)	-0.294***	-0.346***	-0.369***	-0.396***
	(-2.93)	(-4.26)	(-4.48)	(-4.60)
Number of SHORT portfolios entered at t=0 (Intangibles anomalies)	-0.177	-0.186*	-0.084	-0.195*
	(-1.43)	(-1.85)	(-0.83)	(-1.82)
Control variables	YES	YES	YES	YES
N observations	17,004	18,835	18,760	18,420
Adjusted R^2	0.1201	0.1692	0.1743	0.1631

Chapter II: Who Profits from Trading Options?

1. Introduction

One of the fastest growing public financial markets in recent years is the exchange-traded options market. According to data from the Options Clearing Corporation (OCC), the total options trading volume in the U.S. reached 7.47 billion contracts in 2020, worth 4.59 trillion dollars, which represents a growth rate of more than 25 times compared to the public options market size in 1996. A unique feature of the options market is the multitude of trading opportunities it allows. While stock investors only get linear payoffs and are sometimes constrained in short selling, options traders can construct various payoff structures to gain specific risk exposure to both the underlying stock price and volatility. For example, naive investors may view options as leveraged stocks and speculate on stock price movements using simple positions such as long calls or long puts. Meanwhile, more skilled investors may use options to construct hedged positions and limit risk exposure with strategies such as bull/bear spreads, or to trade on volatilities using straddles and strangles. Although conventional wisdom suggests that derivatives trading requires comprehensive financial knowledge and is more suitable to sophisticated investors such as institutions, retail participation in exchange-traded options markets is surprisingly high. For example, Goldman Sachs estimates that odd lot trades, as a measure of retail trading, account for about 10% of total options volume in 2019, compared to about 1.5% in the equity market.[1] How do retail investors perform compared to institutions in the booming options market? What factors affect the performance of the different investor classes? In this study, we aim to answer these broadly unaddressed questions that are relevant to both researchers and policy makers. We conjecture that the complexity of options strategies

[1] See Bloomberg article "Bored Day Traders Locked at Home Are Now Obsessed With Options".

used by different investors reflects their financial knowledge, trading skills and level of sophistication, which in turn affect their investment performance. Our results confirm this hypothesis. We show that the trading style, defined in terms of the options strategies used by an investor, significantly affects trading performance, in addition to the effect of investor class (retail or institution). Furthermore, these style effects cannot be explained by risk premia or known behavioral biases.

We examine one of the leading derivatives markets in the world, namely the South Korean options market. Our analysis takes advantage of a complete administrative dataset of options transactions, with anonymous account identities, executed on the Korea Exchange (KRX) between 1 January 2010 and 30 June 2014. The underlying asset is the Korea Composite Stock Price Index (KOSPI 200), similar to the S&P 500 index. At that time, the KOSPI options market is the most liquid public derivatives market according to Futures Industry Association (FIA), with more than 7.7 million option contracts traded per day. The KOSPI options market is also a global market available to foreign investors, both institutional and retail. In fact, foreign investors contribute around 35% of the aggregate trading volume in this market, indicating that our findings may be relevant beyond the Korean market. In addition to the options data, we also obtain data of futures on the same underlying index, which are used by some traders to construct combinations with options. These detailed account-level data of the whole market provide us with a unique opportunity to observe the full transaction and position records for each account, labelled as either institutional or retail investor, and to calculate the account profitability.

Given scarce evidence in the literature of how different types of accounts use options, we start by documenting some stylized facts in our sample. Firstly, consistent with conventional wisdom, we find a significant difference in the complexity of trading strategies used by institutions and retail investors. Around two thirds of institutional accounts trade both options and futures, while two thirds of retail accounts trade only options but not futures. Moreover, more than 50% of retail end-of-day positions represent simple option strategies consisting of concentrated bets in only one type of option contracts (only long calls, short calls, long puts, or short puts). This is indicative of a high tendency for speculation on the underlying price by retail investors. Institutions, on the other hand, often use more complicated strategies, such as combinations of options and futures. Secondly, volatility trading is used by investors more commonly than other classic options trading strategies. Volatility strategies account for 16.7% of retail account-day positions and 9.4% of institutional account-day positions. Thirdly, there are large account-day positions of both institutional and retail investors that appear risk neutral, having near-zero option Greeks (price sensitivities to the underlying price and volatility). Specifically, 21.2% of institutional account-days hold such Greek neutral positions, accounting for 44.4% of total institutional open positions. Similar Greek neutral positions are held by 6.7% of retail account-days, accounting for 34.9% of total retail open positions. These positions likely reflect the activity of market makers and arbitrageurs. Finally, we also find active day trading by both retail and institutional investors. About 17% of institutional accounts and 2.6% of retail accounts only trade intraday throughout our sample period. Together, these day traders contribute about 10% of total trading volume in the market each day.

To test the conjecture that the choice of trading style and its complexity reflects investors' trading skills, we first classify the dominant trading strategy of each account. We define day traders as those who always open and close option positions on the same day and never hold positions overnight. By examining the overnight positions of the remaining accounts, we extract their most frequently used strategy, and categorize them into the following trading styles: simple strategy traders, volatility traders, Greek neutral traders, and others. In order to define the dominant strategy of an account, we require their most frequent strategy to be used on at least 60% of the days when the account holds overnight positions. Among these various types of investors, we expect that volatility traders and Greek neutral traders are more sophisticated than simple strategy traders, given the differences in strategy complexity. The prediction regarding day traders is less clear. Barber et al. (2014) investigate individual day traders in the Taiwanese stock market and whether their behavior is consistent with rational Bayesian learning. The results are mixed: although there is evidence that retail day traders learn about their own trading abilities over time, their aggregate performance is negative. Evidence by Kuo and Lin (2013) and Ryu (2012) shows that individual day traders in the futures market incur substantial losses, while institutional investors tend to make profits from day trading. In general, institutional day traders can be high frequency traders that have a speed advantage over the rest of the market. Beside these four groups of investors with an identified dominant strategy, there is another group of investors who do not have a dominant strategy or mainly use less popular strategies. We label them as "others" and use this group as a benchmark in the performance analyses.

We find that about 76% of retail investors and 50% of institutional investors have a dominant strategy, indicating that it is common for investors to stick to one strategy in trading options. The evidence suggests that this is especially true for retail investors, while institutional investors are more likely to use multiple strategies and less likely to have a dominant one. Consistent with our earlier findings, simple strategies are the most popular among retail investors, with 66% of retail accounts predominantly using such strategies. Surprisingly, about 19% of institutional accounts also use simple strategies as their dominant trading style. Compared to retail investors, institutional investors are more likely to engage in day trading, as this strategy is dominant in 17% of institutional accounts and only 2.6% of retail accounts. We observe a similar pattern for Greek neutral strategies, which are dominant among 9.9% of institutional accounts and 2.4% of retail accounts. Finally, volatility trading is the dominant strategy in 4% of institutional accounts and 5% of retail accounts.

To analyse investment performance, we first calculate the profit and loss (P&L) of each account. We find that retail investors lose substantially in the KOPSI options market. The median retail investor loses 5.5 million KRW (around 5,000 USD) in our sample, equivalent to approximately 21% of annual household disposable income per capita over the same period[2]. Even at the top 30th percentile of performance ranking, retail options traders still lose money. On the other hand, the median institutional P&L is positive. Since the dollar P&L is affected by the account size and length of investment period, we also calculate the average daily return and Sharpe ratio for each account over the full sample period. Then, we use an account-level multivariate regression framework to investigate how the performance metrics are related to investor class (institution vs. retail) and trading style (dominant strategy). Our baseline results show that both investor class and trading style affect investment performance. Institutional investors significantly outperform retail investors in terms of both mean return and Sharpe ratio. Simple strategy traders have the worst performance among all trading styles. Day traders on average also perform worse than the benchmark group of investors (those without an identifiable dominant strategy). Volatility traders earn the highest average returns, and Greek neutral traders have the largest Sharpe ratios. The evidence also suggests that the style effects can be stronger than the investor class effect. The performance difference between institutional and retail investors narrows once we examine investor class and trading styles in the same model. In contrast, the significance of the trading style effects remains strong after including investor class into the model. In addition, the outperformance of institutional investors is related to systematic risk exposure. When we remove the loadings on systematic risk factors from the daily returns, we find that institutional investors on average do not earn higher returns

[2] OECD (2021), "Household disposable income" (indicator), https://doi.org/10.1787/dd50eddd-en (accessed on 16 February 2021).

than retail investors, but they still have higher Sharpe ratios. On the other hand, all the trading style effects remain the same when we use risk-adjusted returns.

Next, we examine the trading style effects for institutional and retail investors separately. Because the sample is populated with more retail investors, the style effects in this group of investors are similar to the baseline results in the full sample. Complexity of the dominant strategy is a significant predictor of individual investors' performance. Specifically, retail investors predominantly using sophisticated strategies (volatility and Greek neutral strategies) outperform the benchmark group of investors in terms of both average return and Sharpe ratio. While the retail volatility traders have the highest average returns, retail Greek neutral traders have the highest Sharpe ratios. We also find that retail investors who mainly use simple strategies or who are day traders underperform the benchmark group, with simple strategy traders performing the worst. When we turn to institutional investors, the results are different in some respects. In terms of average returns, we find that day traders, volatility traders, and even simple strategy traders all have significantly higher returns than the benchmark group. Since simple directional bets are often used by speculators and informed investors, the fact that retail traders lose pervasively from such simple strategies while some institutions profit from them, suggests that institutions tend to be better at acquiring advanced information. Greek neutral institutional investors' average return is also higher than the benchmark, but statistically insignificant. However, Greek neutral institutions have significantly higher Sharpe ratios. This superior performance in terms of Sharpe ratios also exists for institutions that are day traders or volatility traders, but not for the institutions that are simple strategy traders. The results suggest that, while investor class matters for profitability, the trading styles measure investor sophistication on a more granular level. Within each of the investor classes, especially among retail investors, the complexity of an account's dominant strategy is significantly associated with investment performance.

Furthermore, we examine the account performance attributable to the dominant versus non-dominant strategies. Rational Bayesian investors should gradually learn what is the most suitable strategy for themselves and use it as their dominant strategy. Therefore, for a rational investor, we expect the performance from positions in non-dominant strategies to be worse than the performance from positions in the dominant strategy. We find that this is generally true for institutional investors, as their performance when using non-dominant strategies tends to be worse. The same holds for retail volatility traders, whose performance from non-dominant strategies is worse than the dominant volatility trading strategy. On the other hand, retail investors who predominantly use simple strategies tend to have significantly better performance when using other strategies. These results suggest that retail investors using simple strategies are less likely to be Bayesian learners, while retail volatility traders exhibit a higher level of sophistication when choosing their dominant trading style.

The trading style effects are persistent. We perform out-of-sample tests by splitting the sample into two equal periods and identifying the dominant strategy of each account in each period separately. The dominant strategy in the first half of the sample is likely to be dominant also in the second half. Using the trading styles identified in the first period, we find similar effects of styles on trading performance in the second period, for both retail and institutional investors. This confirms that the style effects are not random and supports our interpretation that trading style reflects investor sophistication.

After observing that retail investors who use simple strategies lose to the rest of the market, we explore some possible causes of this underperformance. First, we examine whether the choice of dominant strategy by retail investors is related to behavioral patterns. The high retail participation in options trading could reflect known biases such as: lottery preference (Boyer and Vorkink, 2014); overconfidence (Barber and Odean, 2000); disposition effect (Odean, 1998); trend chasing, also known as extrapolation or positive feedback trading (De

Long et al., 1990); and self-attribution bias (Gervais and Odean, 2001; Hoffmann and Post, 2014). Underperformance due to behavioral biases could be amplified in the options market due to the high leverage and large tail risk embedded in option contracts. Using each account's activity measures, we calculate proxies for these biases at the account level. Then, we perform logistic regressions of the dominant strategy dummies on the behavioural bias proxies. We find that the likelihood of a retail investor using simple strategies as the dominant trading style is positively and significantly associated with investor overconfidence as measured by the average daily trading volume, disposition effect as measured by proportion of gains realized minus proportion of losses realized, and lottery preference as measured by the percentage of positions in long out-of-the-money options. The same investors strongly prefer to provide rather than take liquidity and they do not tend to chase trends. In contrast, we find the opposite results for retail investors whose dominant trading style consists of volatility trading or Greek neutral strategies. One exception is the fact that Greek neutral traders have high average daily trading volume, but this is a feature of the strategy itself rather than a sign of overconfidence. We do not find self-attribution bias to be present among any of the retail investors who have a dominant strategy. Our analysis also reveals that foreign retail investors are less likely to use simple strategies. Overall, the behavioral bias measures for overconfidence, disposition effect, and lottery preference are significantly associated with the level of complexity of each account's dominant trading style. Therefore, in our next analysis, we investigate whether the behavioral biases drive the previously identified effects of trading style on investment performance. We include the behavioral bias proxies as control variables in the performance analysis regressions and find that the dominant strategy dummies behave largely the same as before. Although we do find that many of the behavioral biases adversely affect investor performance, the trading style effects remain strongly significant and do not appear to be driven by the biases. This suggests a separate venue for investor sophistication to impact performance.

Our contributions to the finance literature are mainly twofold. We are the first to document detailed account-level activity in the options market. The only study from the US market that provides stylized facts about the options trading activity of several types of investors is by Lakonishok et al. (2007). However, their account-level analysis is restricted by data availability and quality. The holdings data they analyse cover only a small sample of retail investors and do not have detailed characteristics of each option contract. Given the complexity of options strategies, it is challenging to identify trading motivations and actual strategies in such data.[3] While Lakonishok et al. infer that volatility trading accounts for less than 3% of options market activity, we show that, in our complete set of account-level transactions data, volatility trading is more common, with a lower bound estimate of 16.5% at the account-day level. On the other hand, the most common strategies identified by Lakonishok et al., covered calls and protective puts, account for less than 1% of observations in our data. While our results are derived from the Korean market, we do find similar patterns in the trading activity of foreign investors in our sample. It is also possible that the development of global derivatives markets in the last two decades has changed trading patterns fundamentally since the end of Lakonishok et al.'s sample.

Apart from describing the options market in detail, our goal is to provide an ex-ante measure of investor sophistication and trading skill. Some studies measure skill based on ex-post performance (see e.g., Li, Subrahmanyam, and Yang, 2021, who use account-level data to examine trading patterns during the Chinese warrants market bubble). In contrast, we use the

[3] Several other studies examine options trading in international markets (see Bauer, Cosemans and Eichholtz, 2009; Chaput and Ederington, 2003; Fahlenbrach and Sandås, 2010; Flint, Lepone and Yang, 2014). These studies are undermined by various data issues. They either have only a small sample of options, or lack account-level transactions and positions data, or are not able to compare institutional and retail investors. In contrast, our study uses a comprehensive dataset of account-level transactions and positions in Korean index options and futures, which allows us to examine in detail the option trading strategies of different types of investors.

complexity of investors' trading styles, defined by their dominant strategy, to measure skill ex-ante and then show how it is connected to performance. The finance literature has traditionally evaluated trading skill and sophistication based on investor class. Institutions are generally regarded as sophisticated informed investors, while retail traders are believed to be uninformed biased traders who commit systematic mistakes. Several studies show that individual investors incur losses in the stock and futures markets (e.g., Odean, 1999; Barber and Odean, 2000; Barber et al., 2009; Kuo, Lin, and Zhao, 2015, 2018). On the other hand, several recent papers suggest that retail equity traders may be informed, as their aggregate trading can predict future stock returns (see Kaniel, Saar and Titman, 2008; Kaniel et al. (2012); Kelley and Tetlock, 2013, 2017; Boehmer et al., 2019). In addition, it is important to study retail investors' activity due to their role in liquidity provision. For example, Barrot, Kaniel and Sraer (2016) find that retail investors provide liquidity especially in times of market stress. Recently, retail investors are also believed to support and capitalize the quick market rebound after the market crash due to the COVID-19 pandemic.[4] Welch (2020) shows that Robinhood investors collectively increased stock holdings quickly after the COVID-related market crash and were rewarded by the subsequent bull market. In our study, we show that, although institutions outperform retail investors on average, there is a subset of highly profitable retail investors who are skilled at trading volatilities or using complicated and well hedged strategies. Evidence of such sophisticated use of derivatives by retail traders is novel to the literature. Related to our findings, Bauer, Cosemans, and Eichholtz (2009) also uncover a small subgroup of retail option traders who outperform their peers. However, they provide little insight into the source of such outperformance, while we show that volatility and Greek neutral strategies are the key contributors to superior profitability. While our comparison of the different investor classes

[4]For example, see "Young investors pile into stocks, seeing 'generational-buying moment' instead of risk" and "Robinhood traders cash in on the market comeback that billionaire investors missed" by CNBC.

adds to the literature on the characteristics of institutional versus retail traders, we further classify investors based on the complexity of their trading style. We show that this allows for a more granular identification of trading skills, both in the full universe of investors and within each investor class separately.

Our study is also related to the literature on product complexity and investment performance. There is evidence that retail demand for complex structured products is difficult to rationalize and is likely driven by behavioral factors, since these products are designed to be more profitable for the issuers (Henderson and Pearson, 2011; Li, Subrahmanyam and Yang, 2018; Hens and Rieger, 2011; Célérier and Vallée, 2016). Our study shows that, in a market where an investor can freely choose what strategies to use, the selection of a particular strategy and its level of complexity reflects the investor's knowledge and ability. Accordingly, investors that self-select into complex strategies have better performance than those using naive simple strategies. Our results indicate that access to complicated financial products can be valuable to certain investors although such access should not be automatically granted to all investors in the market.

The rest of the paper proceeds as follows. Section 2 provides a description of the KOSPI 200 derivatives markets and our data. Section 3 examines the various trading styles used by option investors. Section 4 analyses the profitability of the different types of investors. Section 5 concludes the paper.

2. Background and data description

We use a unique transaction-level dataset from Korea's main index derivatives markets. The data contains all options and futures trades at the Korea Exchange (KRX) from 1 January 2010 to 30 June 2014 and detailed account information, including anonymous account identity, investor class (institution or retail), country of domicile, etc. The options and futures contracts have the same underlying asset, the KOSPI 200 index. The index consists of the 200 largest companies listed on KRX, thus representing Korea's overall stock market. The futures contract size is equal to the KOSPI 200 index value times the multiplier of KRW 500,000. During the sample period, the average contract size is approximately KRW 126 million, equivalent to about USD 113,569. This is larger compared to the average contract size of the most popular index derivatives in the U.S. during the same period, the E-mini S&P 500 index futures (approximately USD 71,000 per contract). KOSPI 200 options have a smaller multiplier of KRW 100,000 for contracts that mature in or before June 2012, but the multiplier changes to KRW 500,000 for contracts that mature after June 2012.

Both KOSPI 200 options and futures markets are order-driven and do not have designated market makers for the provision of liquidity. Orders submitted by investors are collected in a central electronic limit order book and are executed according to price and time priority rules. The regular continuous trading session starts at 9:00 and ends at 15:05 local time. There are two auctions that determine the open and close prices, one from 8:00 to 9:00 (open) and the other from 15:05 to 15:15 (close). The transaction data contain a millisecond time stamp, account IDs for both counterparties, and bid and ask order submission times, which allow us to determine each account's long/short position as well as the order flow. We classify as the trade initiator the account that submits their order later than the counterparty in a transaction. In order to accurately describe market activities and investor performance, we focus on both options and futures contracts introduced since 2010, for which our data contain complete information about investors' positions.

Table 1 reports aggregate summary statistics of the KOSPI 200 options market including the number of transactions, trading volume (number of contracts traded), and total options premium. In Panel A, we first calculate these statistics for each day, and then we calculate their time-series mean, standard deviation, minimum, median, maximum, and total sum. During our sample period, the total trading volume in KOSPI 200 options is approximately 8.6 billion contracts, which corresponds to total options premium of KRW 1,323,552 billion (approximately USD 1,143 billion). The rest of the panels in Table 1 contain summary statistics for different subsamples of the data. Moneyness of a call (put) option is defined by the ratio of the underlying spot price (strike price) to the strike price (underlying spot price). An option is out of the money (OTM) / at the money (ATM) / in the money (ITM) if its price ratio is less than 0.95 / between 0.95 and 1.05 / greater than 1.05. Panel B shows that ATM options are the most actively traded, followed by OTM options, while ITM options attract little trading volume. In Panel C, we observe that trading volume in call options is slightly higher than in put options. Panel D shows that, as expected, most trading activity takes place during normal trading hours, which refer to the daily continuous trading session from 9:00 to 15:05. Finally, in Panel E, we can see that contracts which are closer to maturity are more actively traded.

[Table 1 about here]

In total, there are 187,323 investor accounts in our data, of which 161,010 are option traders. There is high retail participation in the Korean derivatives markets. Domestic retail investors constitute the largest number of accounts (95.5% of all accounts), followed by domestic institutions, foreign institutions, and a small number of foreign retail investors. About two thirds of retail accounts trade only options, and the remaining one third trade both options and futures. This pattern reverses for institutional accounts – about one third of them trade only options, and two thirds trade both options and futures. Although most of the accounts in the data are domestic investors, foreign institutions generate a large portion of options trading volumes. Table 2 provides summary statistics of the options trading activity of the different investor classes. For each account and each day, we calculate the number of transactions, trading volume, and premium. Reported are daily averages for the aggregate trading activity of all accounts in each investor class, as well as the average account activity for that investor class. We report statistics for liquidity takers and liquidity providers separately. We compare the bid and ask order submission times and mark the investor who submitted their order first as the liquidity provider and the investor whose order matched the first one as the liquidity taker. There are some transactions where the two orders cross at the same time, so the liquidity taker and liquidity provider cannot be identified. Those observations are not included in the results in Table 2. The table shows that, on an average day, institutional investors tend to be liquidity takers in aggregate, while retail investors tend to provide more liquidity than they take. As can be expected, an average institutional account generates substantially larger trading volumes and premium, compared to an average retail account.

[Table 2 about here]

3. How do investors use options?

3.1. Strategies in account-days

We start by extracting the trading strategy of each account on each trading day. First, we separate accounts into day traders and position holders. Day traders are the accounts that only trade intraday and never hold a position overnight throughout our sample. The remaining accounts which hold option contracts overnight are position holders. We use the transactions data to construct the end-of-day position held by each investor in each different contract. For each account, day and contract, the end-of-day position is equal to the previous day's position plus any purchased lots minus any sold lots. Then, we examine the combinations of different contracts that each account holds at the end of each day and we extract the corresponding strategies.

Table 3 examines the popularity of different options strategies used by position holders. We group strategies into five main categories: combinations of options and futures, simple strategies, volatility trading strategies, spreads, and other options strategies. (The supplementary Table I in the appendix provides additional details about the break-down of strategies within each of these five main categories.) We report the number of account-day positions in each strategy as a percentage of all account-days with non-zero end-of-day position in options, as well as the number of option contracts held in each strategy as a percentage of total end-of-day open positions. The table contains results for the whole market and for each investor class separately.

[Table 3 about here]

First, combinations of options and futures include covered calls, protective puts, and any other combinations. We extract the strategies by examining the type of options and futures exposure that each account-day holds: long calls, short calls, long puts, short puts, long futures, and/or short futures. For example, long (short) covered calls consist of long (short) calls and short (long) futures, while long (short) protective puts consist of long (short) puts and long (short) futures. Combinations of options and futures account for 6.24% of all account-days and 37.6% of all open positions. They are more popular among institutional investors, constituting 34.4% (55.4%) of all institutional account-days (open positions). In comparison, only 5.42% (10.5%) of retail account-days (open positions) hold combinations of options and futures. We can think of futures positions as exposure to the underlying, and options can be used to hedge that exposure. However, we observe that well-known hedging strategies such as covered calls and protective puts are rarely used, as they constitute less than 16% of all options-and-futures combinations (refer to the appendix for details). Therefore, options do not appear to be widely used for hedging futures positions.

Second, simple strategies are defined as concentrated positions in only one type of option contract: only long calls, only short calls, only long puts, or only short puts. Such options positions provide directional exposure to the underlying and may be thought of as naked options in the sense that they are not combined with any position in futures contracts. Hence, they do not hedge any market exposure from futures contracts. Although index options may also be used to hedge equity portfolios, it seems unlikely in our data. ETF trading volumes in Korea are very low during our sample period, so these options positions are not likely to be used for hedging ETFs. They could be used for hedging portfolios of stocks, but this is more likely to be the case for individual stock options rather than index options. Although we do not have the necessary data of investors' equity holdings to conclude with certainty, our argument that these are not hedging positions is supported by two observations. First, it is consistent with the results of Lakonishok et al. (2007) and Bauer, Cosemans, Eichholtz (2009), who find that hedging underlying price changes drives a relatively small part of the option market activity of retail accounts. In addition, we find that very few of the institutions likely to hold stock and hence likely to hedge (pension funds, insurance companies, and government institutions) are simple strategy traders. For these reasons, we can be confident in assuming that these simple strategy positions are most likely used for speculative bets rather than hedges of the underlying index. Table 3 shows that a surprisingly large percentage of account-days hold positions in simple strategies. Simple strategies account for 55.7% of the account-days in the data, although that corresponds to only 7.14% of all open positions. When we examine the breakdown of our results by investor class, we observe a widespread use of options for directional speculative trading by retail investors. Simple strategies are used by 56.8% of retail account-days. It seems that some institutions also engage in such speculative trading, although less often than retail investors. Simple strategies are used by 18.8% of institutional account-days, but their holdings account for a very small percentage of open positions. Additionally, Table I in the appendix shows that long calls and long puts comprise the majority of retail positions in simple strategies, while institutions tend to hold more short calls.

Third, we describe volatility trading strategies, which include straddles, strangles, and butterflies. These strategies are used by investors who want to trade on information about underlying volatility or to hedge volatility risk. First, we extract all account-days that use only two different option contracts to create straddles or strangles. Long (short) straddles are created by combining long (short) calls and long (short) puts with the same strike price and maturity date. Long (short) strangles are created with long (short) calls and long (short) puts with the same maturity date but different strike prices. Next, we would like to extract account-days that use more than two different option contracts to create combinations of straddles and strangles. Regardless of how many different option contracts an account holds on a given day, we can check their exposure. If the account holds long (short) calls and long (short) puts only, we can be sure that the corresponding strategy is a combination of long (short) straddles and/or long (short) strangles (in the appendix, we report these strategies as "long combinations" and "short combinations"). Finally, we extract butterflies created with three different option contracts. A call (put) butterfly spread is a strategy that combines three call (put) contracts with different strike prices, such that the option contract with the middle strike price has twice the number of lots invested in it, compared to the number of lots invested in the other two option contracts. For example, a long call butterfly can be created by buying one lot in a call option contract with the lowest strike price, selling two lots in a call option contract with the middle strike price, and buying one lot in a call option contract with the highest strike price. We must note that we are only able to identify butterfly spreads created using three different option contracts, but we are unable to extract any combinations of butterflies created using six, nine, or more different option contracts, because we cannot be sure that they do not correspond to other strategies. Similarly, some combinations of long and short straddles and strangles may be left unidentified and remain in the category of other strategies. In general, we cannot conclude with certainty about the types of strategies used by an account that holds a very large number of different option contracts. Consequently, the numbers in Table 3 represent a lower bound and the use of options for volatility trading may in fact be larger.[5] The table shows that at least 16.5% of all account-days engage in volatility trading, and their positions make up 7.24% of all the open positions in the market. This indicates that options are used for volatility trading more widely than Lakonishok et al. (2007) suggest. Therefore, options are an important instrument for trading on or hedging underlying volatility and are not solely used for speculating on underlying price changes. Interestingly, volatility trading is more popular among retail investors than institutions. Volatility strategies constitute 16.7% (13.9%) of retail account-days (open positions), compared to 9.4% (2.9%) of institutional account-days (open positions). According to the results in the appendix, strangles are the most commonly used volatility strategy. Combinations of strangles and/or straddles are popular as well, while butterflies are used very infrequently.

Fourth, spreads include strategies such as bull spreads, bear spreads, synthetic stocks, and calendar spreads. Spreads do not seem to be widely used by option investors. Only 3.7% of account-days hold spreads, and their positions account for less than 3% of all open positions. For this reason, we do not examine these strategies further.

Finally, the category of other options strategies consists of any combinations of option contracts which do not fall into the categories discussed above. Almost 18% of all account-

[5] To be sure that we have accurately identified volatility trading strategies in our data, we calculate each account-day position's exposure to the Greeks. We focus on delta, which measures the exposure of an option position to changes in the underlying price, and vega, which measures the sensitivity to changes in the underlying volatility. We scale end-of-day delta and vega exposure by the number of lots held by the account on that day. As expected for volatility strategies, they have a low average scaled delta and a high average scaled vega in absolute terms. The long volatility strategies have an average scaled delta of 0.01 and an average scaled vega of 0.11, while the short volatility strategies have an average scaled delta of -0.02 and an average scaled vega of -0.16.

days fall within this remaining category, and their holdings account for 45% of all open positions in the market. Such large positions in complex combinations of multiple different contracts which do not correspond to any of the other trading strategies, are possibly used by accounts that act as arbitrageurs or market makers (since there are no designated market makers in the KOSPI 200 derivatives markets). We create a separate category for such accounts by identifying Greek neutral positions. Unlike volatility traders, market makers and arbitrageurs are expected to have end-of-day positions with low exposure to both delta and vega. This is because, whenever they are unable to close their positions before the end of the trading day, they would need to hedge the risks associated with carrying inventory overnight. We start by collecting account-days which have remained without an assigned strategy, including those holding combinations of options and futures other than covered calls and protective puts, and those holding only options whose strategies do not correspond to simple strategies, volatility trading, or spreads. From this pool, we exclude any combinations of options with different maturities, which may be used to achieve a calendar exposure. Then, we impose artificial cut-offs on the Greeks exposure of these account-day positions. Those that are below the median in terms of both absolute scaled delta and absolute scaled vega exposure are categorized as Greek neutral positions.[6] In total, Greek neutral strategies are used by about 7% of all account-days and comprise 40.6% of all open positions in the market (refer to the appendix for details). They are more popular among institutions but appear to be used by some retail investors too.

3.2. Dominant strategy by account

Before we extract the dominant strategy of each account, we would like to exclude any accounts that are relatively inactive. An account is marked as inactive if it appears in the options

[6] These Greek neutral positions have an average absolute delta of 0.01 and an average absolute vega of 0.01. In comparison, the rest of the positions which are above the medians, have an average absolute delta of 0.14 and an average absolute vega of 0.09.

data for only a month or less, or if the account exhibits infrequent activity. We exclude day traders who trade options on less than two days per month on average, as well as position holders who trade options and hold option positions on less than two days per month on average (in the months when those investors appear in the data). We also exclude accounts whose total trading volume throughout the sample period is in the bottom fifth percentile (<=22 lots traded) or whose total number of active days in the options market is in the bottom fifth percentile (<=2 active days). Almost 24% of accounts are inactive and we exclude them from all further analyses. Most of them (16.5%) are excluded because they appear in the options data for less than a month, and the rest (7.4%) are excluded due to infrequent activity. We filter the sample in this way in order to have cleaner results with less outliers whose performance can be attributed to luck rather than strategy. In addition, our focus is only on active options traders, meaning that we exclude accounts that are active in the futures market but not in the options market. We are left with a total of 122,620 active option trader accounts.

We define the trading style of each account by extracting the account's dominant strategy. We group investors into the following five categories: day traders, simple strategy traders, volatility traders, Greek neutral traders, and others. As defined earlier, day traders are accounts that only trade intraday and never hold positions overnight. The rest of the accounts which are position holders may hold different strategies on different days throughout the sample period. We use the following rule to identify the dominant strategy of each of these accounts. If at least 60% of an account's end-of-day positions are positions in simple strategies, then we define that account as a simple strategy trader. We use the same definition to mark volatility traders and Greek neutral traders.[7] The remaining category "others" consists of accounts whose trading style is unidentifiable as any of the strategies, as well as accounts that

[7] Our inferences are robust to alternative cut-off points, such as 70% or 80%.

switch between different strategies without having a dominant one. Table 4 reports the number of accounts and their average trading volume by trading style and investor class.

[Table 4 about here]

Day traders comprise about 3% of all accounts in our sample, and this corresponds to 17% of institutional accounts and 2.6% of retail investors. Simple strategy traders are the largest group, with 64.5% of accounts falling into that category. Most of them are retail accounts, but around 19% of institutions are also simple strategy traders. Volatility traders represent 5% of all accounts, with 4% of institutions and 5% of retail investors using volatility strategies as their dominant trading style. Only 2.6% of accounts are identified as Greek neutral traders, which is not surprising given the complex skills and capacity required to use such strategies. Greek neutral strategies are more popular among institutional investors, with almost 10% of institutions being Greek neutral traders. Nevertheless, there is a small group of retail accounts that are sophisticated enough to use such well hedged positions as well. Finally, the remaining investors in the category "others" constitute about 25% of all accounts, which corresponds to 50% of institutions and 24% of retail traders. These are accounts whose trading patterns do not match any of the preceding strategies which we have defined. Some of them are likely to be volatility traders, since we are unable to uncover absolutely all volatility strategies in the data. Another small portion of them may be using spreads and calendar exposures as their dominant strategy. The rest perhaps use complicated proprietary strategies or other uncommon strategies. It is also possible that some are unsophisticated traders who do not follow any predefined trading strategy. In addition, this category also includes investors who frequently switch between any of the strategies described above and therefore do not have a single dominant trading style. In terms of trading volumes, an average account within the category of Greek neutral traders generates the largest mean daily trading volume, especially if it is an institutional account. This is consistent with our conjecture that these Greek neutral traders are likely acting as market makers or arbitrageurs. An average day trader also generates significantly large mean daily volumes. In comparison, the rest of the traders have lower trading volumes on average, because their strategies likely consist of holding positions for longer periods of time, rather than trading frequently. In terms of aggregate volumes, the two groups of simple strategy traders and other traders generate respectively 37.1% and 30.4% of total trading volume in the market, since they consist of the largest number of accounts.

4. Who profits from trading options?

After describing the most commonly used options trading styles, we proceed to analyze account performance. We calculate several different profitability measures. First, we start by calculating each account's total dollar profits generated over the sample period. For each investor account, we use the transactions records to calculate the total cumulative dollar profit and loss ($P&L) over the whole sample period. In other words, we measure the dollar amount generated by all trades that the account has executed from 1 January 2010 to 30 June 2014. If there are any positions that are not closed before the end of our sample period, we mark them to market based on the last best bid and offer (BBO) quotes midpoint on the last trading day in our sample. We use the contract multiplier to calculate profits in Korean won (KRW). However, profitability in dollar terms is not an ideal measure of investor performance because it can be contaminated by account size and capital constraints. To remove such effects, we use the following method to calculate profitability per dollar invested. We start by separating all transactions into intraday transactions and transactions that create positions held at the end of the day. Round-trip trades which buy and sell the same contracts on the same day are marked as intraday transactions. Trades which open a new position that is held at the end of the day and trades which close an existing position from the previous day are marked as transactions which create end-of-day positions. If a single transaction accomplishes both at the same time, we split it into two parts. Opening and closing transactions are matched consistently based on the "first-in, first-out" method, where inventory assets acquired first are sold first. For example, if an investor has in their inventory five lots of the same option contract and we observe a sale of one lot, it means that they are selling the oldest inventory item. For each account and each trading day, we calculate $P&L from positions and $P&L from intraday trades separately. End-of-day positions are marked to market based on the last daily best bid and offer (BBO) quotes midpoint on non-expiration dates. If a position is held until maturity, P&L is calculated using the final settlement price based on the underlying index value. Then, we calculate daily capital requirements using the margin requirements of the Korea Exchange (KRX). Such capital requirements arise from the fact that investors pay the full contract prices to open long positions and deposit margins to open short positions in the KOSPI 200 options and futures markets. The margin rate for short trades ranges between 10.5% and 15% of the contract value throughout our sample period. When the value of the margin account falls short of predetermined thresholds (maintenance margin), the exchange issues a margin call to the investor requesting top up in the margin account. Because the margin settlement is daily, we exclude intraday trades from our calculations of end-of-day capital requirements. Hence, end-of-day capital requirements are based only on the value of end-of-day position holdings. In addition to that, investors who trade intraday are subject to intraday capital requirements. We loop through each account's intraday transactions and update the required capital at each point in time. We take the maximum daily value as the account's intraday capital requirement for that day. It reflects the amount of capital the investor needed to have in their account on that day in order to execute all the observed intraday transactions. Finally, we sum each account's daily P&L from positions and P&L from intraday trades, and scale it by the sum of capital requirement at the end of the previous day and intraday capital requirement. The scaled P&L is therefore a rate of return to each dollar of capital invested in KOSPI derivatives. The formula below summarizes our method for calculating daily returns for each account:

$$\text{Daily return} = \text{Total daily \$P\&L} / \text{Total daily capital requirement}$$
$$= \frac{(\text{ \$P\&L from end-of-day positions} + \text{\$P\&L from intraday trades })}{(\text{ Lag of end-of-day capital requirement} + \text{Intraday capital requirement })}.$$

The feature of capital requirement sets options trading different from stock trading because options sellers do not receive the sales proceeds and usually put down a sizable margin instead. In other words, selling options is analogous to short selling borrowed stocks. Compared to the simple option return calculated based on price changes, our scaled P&L is a more practical and implementable return measure because it takes into account the unique feature of margins in derivatives trading. To have one single return measure for each account, we take the mean of daily returns of the same account during the whole sample period. Moreover, we calculate Sharpe ratio of each account as the mean daily return divided by the standard deviation of daily returns.

Figure 1 shows the average profitability of the 122,620 active option trader accounts. It plots daily mean returns and Sharpe ratios, averaged across accounts in each category of investors. Panel A shows results by investor class. Unsurprisingly, institutional investors outperform retail investors on average. Panel B shows results by trading style, as defined by the dominant strategy of each account. Simple strategy traders are the only group which incurs losses on average. Volatility traders generate the highest average returns. Since the profits of volatility trading strategies depend on the market conditions during a particular time period, it is important to note that our sample can be considered representative, as it contains both periods of low volatility and periods of market stress with spikes in volatility. Although volatility traders appear to be very profitable in terms of daily mean return, their average Sharpe ratio is lower due to the large downside risks they are subject to. In comparison, Greek neutral traders have lower daily mean returns because they generate a smaller profit per trade. However, they also incur significantly lower risks, since their positions are well hedged. For this reason, Greek neutral traders outperform everyone else in terms of Sharpe ratio which measures risk-adjusted returns. Panel C plots average profitability by trading style in the different investor class subsamples. Institutions seem to have positive average profits regardless of the dominant strategy they adopt. In terms of daily mean returns, institutional volatility traders outperform all other institutions. In terms of Sharpe ratio, institutional day traders and Greek neutral traders are the most successful groups. Similarly, among the retail investors, volatility traders outperform in terms of daily mean returns, while Greek neutral traders outperform in terms of Sharpe ratio. Unlike institutional day traders, retail day traders do not appear to be profitable on average. The only category of investors that loses to the rest of the market is retail simple strategy traders. Since simple directional bets are often used by speculators and informed investors, the fact that retail traders lose pervasively from such simple strategies while some institutions profit from them, suggests that institutions tend to be better at acquiring advanced information. Overall, while investor class matters for profitability, we also observe that within both the institutional and retail subsamples, the complexity of an account's dominant strategy is significantly associated with average performance.

[Figure 1 about here]

Table 5 provides more details about account performance by reporting the percentiles of three different measures: total account $P&L in millions of KRW[8], mean daily return, and Sharpe ratio. Panel A shows the performance distribution of institutional versus retail accounts. The median institutional account gains KRW 8.8 million over the whole sample period. Mean

[8]The total dollar profits generated by all accounts in the data, including inactive accounts, accounts that trade only futures but not options, and accounts excluded due to missing information about their investor class, sum to zero (derivatives trading is a zero-sum game).

daily return and Sharpe ratio is negative only for institutions at the 20th percentile or below. In contrast, the median retail investor loses KRW 5.5 million. Retail accounts start to be profitable only at the 80th percentile in terms of both total $P&L and mean daily return. Next, Panel B shows performance distributions by trading style. At the median, only Greek neutral traders gain a positive total $P&L, equal to KRW 1.8 million over the sample period. The median volatility trader and median Greek neutral trader generate positive returns, with the median Greek neutral trader having a Sharpe ratio twice as large as that of the median volatility trader. On the other hand, simple strategy traders are the only group with a negative return and Sharpe ratio at the median. They lose a significant amount of money and start to be profitable only at the 90th percentile.

[Table 5 about here]

Next, we test the observed performance patterns in multivariate regression analyses. Table 6 presents regression results for the relation between account profitability and investor type. First, we perform three account-level regressions using mean daily return as the dependent variable. The independent variable in the first regression is a dummy variable equal to one if the account is an institutional investor, or zero if the account is a retail investor (hence, the intercept represents the retail investors). The second regression uses as independent variables the trading style dummies for day traders, simple strategy traders, volatility traders, and Greek neutral traders. The intercept represents the remaining accounts in the category "others" (those without an identifiable dominant strategy), which become our benchmark group. The third regression combines the investor class dummy and the trading style dummies. After that, we repeat the same three regressions, but using Sharpe ratio as the dependent variable. In general, the regression results are consistent with the univariate results that we observed previously. The dummy variable for institutional accounts is positive and significant in all regressions, confirming that institutions outperform retail investors. Simple strategy traders exhibit significantly negative performance. Day traders also seem to underperform the benchmark group. Volatility traders strongly outperform the other accounts in terms of average return, and Greek neutral traders are the top performers in terms of Sharpe ratio. The evidence also suggests that the style effects can be stronger than the investor class effect. The performance difference between institutional and retail investors narrows once we examine investor class and trading styles in the same model. In contrast, the significance of the trading style effects remains strong after including investor class into the model.

[Table 6 about here]

In Table 7, we present extensive robustness checks. We repeat the profitability regressions of Table 6 using different definitions of an account's dominant strategy, and different measures of profitability. Regressions 1 and 2 use risk-adjusted returns and Sharpe ratio of risk-adjusted returns as the dependent variables. For each account, risk-adjusted return is equal to the intercept from time-series regressions of the account's daily return on the KOSPI 200 index return and implied volatility on the same day. Sharpe ratio is equal to the t-statistic of the intercept from the same regressions. The daily implied volatility is calculated as the price-weighted average implied volatility of all ATM options, derived from the last daily transactions executed in the 15 minutes before market close. To prevent the effects of outliers, we winsorize the dependent variables cross-sectionally at 5% and 95%. When we remove the loadings on systematic risk factors from the daily returns, we find that institutional investors on average do not earn higher returns than retail investors, but they still have higher Sharpe ratios. On the other hand, all the trading style effects remain the same when we use risk-adjusted returns. In regressions 3 and 4, the dependent variables are mean daily return from positions and Sharpe ratio of returns from positions. We calculate these by excluding position holders' intraday trading profits and only considering their profits from overnight positions. The results regarding position holders' performance are consistent with our main analysis in

Table 6. For regressions 5 and 6, we define dominant strategies and calculate mean daily returns by excluding any simple positions in short call and long put contracts. We do this to exclude any possible hedges of unobserved equity positions. Once again, the observed patterns are consistent with Table 6. Regressions 7-10 show the results when dominant strategies are defined using different cut-off points. More specifically, to define the dominant strategy of an account, we require their most frequent strategy to be used on at least 70% (or 80%) of the days when the account holds overnight positions. The results regarding all position holders' performance remain consistent and strongly significant. There is however some evidence that day traders' Sharpe ratio is now higher than that of the benchmark group. Finally, regression 11 reports the results of a panel regression using account-day observations. These results support the main conclusions from the account-level analyses. The only exception once again are day traders whose account-day returns do not appear to be significantly different from those of the benchmark group.

[Table 7 about here]

In our next analysis in Table 8, we examine the trading style effects for institutional and retail investors separately. Because the sample is populated with more retail investors, the style effects in this group of investors are similar to the baseline results in the full sample. Complexity of the dominant strategy is a significant predictor of individual investors' performance. In particular, retail investors predominantly using volatility or Greek neutral strategies significantly outperform the benchmark group of investors. Retail volatility traders have the highest mean daily returns, while retail Greek neutral traders have the highest Sharpe ratios. Thus, there appears to be a subset of sophisticated retail investors who are skilled at using complicated strategies to generate superior profits. On the other hand, retail investors who are simple strategy traders significantly underperform their peers. Most losses are concentrated in this category of unsophisticated investors and other market participants gain from trading with them. Retail day traders also underperform on average. The results are different in the subsample of institutional investors. In terms of average returns, we find that institutional day traders, volatility traders, and even simple strategy traders all have better returns than the benchmark group. Greek neutral institutions also have a mean daily return that is higher than the benchmark, although not statistically significant. Significantly superior performance in terms of Sharpe ratio exists for institutions that are day traders, volatility traders, or Greek neutral traders, but not for the institutions that are simple strategy traders. Overall, the results support our conjecture that trading style complexity measures investor sophistication and is significantly associated with investment performance. We observe this in both the full sample and in the subsamples by investor class.

[Table 8 about here]

Furthermore, we examine the account performance attributable to the dominant versus non-dominant strategies. For a given account, our definition of dominant strategy has been based on that investor's most frequently used strategy. However, from time to time, the account may engage in strategies other than their dominant one. Rational Bayesian investors should gradually learn what is the most suitable strategy for themselves and use it as their dominant strategy. Therefore, for a rational investor, we expect the performance from positions in non-dominant strategies to be worse than the performance from positions in the dominant strategy. Table 9 shows the results of panel regressions using account-day returns of institutions and retail investors as the dependent variable. We regress the returns on the same dummies for dominant trading style of each account that were used in previous analyses. However, since an account may use strategies other than their dominant one on some days, we also add non-dominant strategy dummies which are equal to one on such days. For example, "Simple strategy trader using non-dominant strategy" is a dummy variable equal to one on days when a simple strategy trader holds positions in volatility, Greek-neutral, or other strategies, and equal to zero on days when the account holds simple positions. We find that the performance of institutional investors generally tends to be worse when they are using non-dominant strategies. (The results for Greek neutral traders seem to be an exception, but this is likely driven by the fact that returns are not adjusted for risk in this analysis.) The performance of retail volatility traders is also worse from non-dominant strategies, compared to the dominant volatility trading strategy. On the other hand, retail investors who predominantly use simple strategies tend to have significantly better performance when using other strategies. These results suggest that retail investors using simple strategies are less likely to be Bayesian learners, while institutional investors and retail volatility traders exhibit a higher level of sophistication when choosing their dominant trading style.

[Table 9 about here]

As an additional robustness check, we conduct out-of-sample tests to show that the trading style effects are persistent. To do this, we split our sample into two equal subperiods. The first half of the sample, used for in-sample analysis, covers the period from 1 January 2010 to 31 Mar 2012. The second half of the sample, used for out-of-sample analysis, covers the remaining period from 1 Apr 2012 until 30 June 2014. Following the same method as in Table 4, we use 60% cut-offs to extract the dominant strategy of each account in each subperiod separately. Therefore, in-sample (out-of-sample) trading style of each account is its dominant strategy in the first (second) half of the sample. Table 10 presents transition matrixes containing the percentage of accounts by trading style in-sample and out-of-sample. We only keep accounts that trade in both subperiods. Panel A shows the transition matrix for the trading styles of institutions, and Panel B for retail investors. The highlighted diagonals reflect trading style persistence. The large percentage values in those diagonals indicate that the in-sample trading style is highly likely to be the dominant trading style also out-of-sample.

[Table 10 about here]

In Table 11, we test the relation between in-sample trading styles and out-of-sample account performance. For each account, we calculate mean daily return and Sharpe ratio in the second half of the sample. Using the dominant strategies identified in-sample, we find similar effects of trading styles on investment performance out-of-sample, for both retail and institutional investors. This confirms that the style effects are not random and supports our interpretation that trading style reflects investor sophistication.

[Table 11 about here]

After observing that retail investors who use simple strategies lose to the rest of the market, we explore some possible causes of this underperformance. We start by examining whether the choice of dominant strategy by retail investors is related to known behavioral biases. Biases could amplify retail underperformance in the options market due to the high leverage and large tail risk embedded in options. Using each account's activity and performance measures, we calculate account-level proxies for some of the most well-known behavioral biases. First, following Barber and Odean (2000) who argue that overconfidence causes excessive trading by retail investors, we calculate the logarithm of average daily options trading volume. Second, we measure disposition effect (Odean, 1998) as an account's proportion of gains realized (PGR) minus proportion of losses realized (PLR). On a given day, PGR is the total number of contracts with realized gains (positions closed at a gain) divided by that number plus the total number of contracts with paper gains (open positions which have increased in value from the previous day). PLR is calculated similarly. Then, we sum the daily values over the sample period, and we take the difference of PGR and PLR. Disposition effect is equal to zero for day traders. Third, some investors may exhibit demand for options with lottery-like payoffs (see e.g., Boyer and Vorkink, 2014). Hence, we measure an account's lottery preference as the out-of-the-money (OTM) option contracts bought. This refers to the percentage of positions in long OTM contracts (for position holders) or the percentage of opening intraday trades buying OTM contracts (for day traders). Fourth, we capture trend chasing behavior, also known as extrapolation or positive feedback trading (see De Long et al., 1990) by calculating sensitivity of delta to KOSPI 200 index return. This is derived by regressing an account's end-of-day position delta scaled by position size (for position holders), or the volume-weighted average delta of opening intraday trades (for day traders), on past five-day average index return. Fifth, we measure self-attribution bias (Gervais and Odean, 2001; Hoffmann and Post, 2014) as the sensitivity of investor activity to past performance. It is calculated by regressing position size (for position holders) or intraday trading volume (for day traders) on past five-day average return of the investor. Sixth, we use the percentage of trading volume initiated as a proxy for transaction costs. Finally, we also include a dummy equal to one if the account is a foreign investor, given some literature on the outperformance of foreign investors (e.g., Grinblatt and Keloharju, 2000). In Table 12, we perform logistic regressions for the probability of retail investors to use a particular trading style given the behavioural bias proxies. We find that the likelihood of retail investors to use simple strategies as their dominant trading style is positively and significantly associated with the proxies for investor overconfidence, disposition effect, and lottery preference. The same investors strongly prefer to provide rather than take liquidity and they do not tend to chase trends. In contrast, we find the opposite results for retail investors whose dominant trading style consists of volatility trading or Greek neutral strategies. One exception is the fact that Greek neutral traders have high average daily trading volume, but this is a feature of the strategy itself rather than a sign of overconfidence. We do not find self-attribution bias to be present among any of the retail investors using one of the three dominant strategies that we focus on. Our analysis also reveals that foreign retail investors are less likely to use simple strategies. Overall, we find that behavioral bias measures for overconfidence, disposition effect, and lottery preference are significantly associated with the level of complexity of an account's trading style.

[Table 12 about here]

In our next analysis in Table 13, we investigate whether the behavioral biases drive the previously identified effects of trading style on investor profitability. Hence, we include the behavioral bias proxies as control variables in the performance analysis regressions. We find that the dominant strategy dummies behave largely the same as before. Although the results show that many of the behavioral biases adversely affect retail investor performance, the trading style effects remain strongly significant and do not appear to be driven by the biases. This suggests a separate venue for investor sophistication to impact performance.

[Table 13 about here]

5. Conclusion

Despite growing trading volumes in derivatives markets worldwide, little is known about the trading activities of option investors. In this study, we analyse a comprehensive dataset of account-level transactions in KOSPI 200 index options and futures in order to investigate the profitability of option trading strategies used by different types of investors. Our study contributes to the literature by providing a detailed description of the options market and uncovering a measure of investor sophistication which goes beyond the traditional classification by investor class. Specifically, we propose the complexity of an investor's dominant trading strategy as an ex-ante measure of trading skill and show that it significantly affects investment performance.

We start by investigating how common option trading strategies are used. We find that simple option strategies, which provide directional exposure to the underlying, are widely used by retail investors. Some institutional investors also use simple strategies to take directional bets, but a majority of institutions are skilled at using various complicated strategies. We show that volatility trading is used more often than the other classic options strategies, and all investor classes in our sample trade on volatilities. We also uncover a small number of accounts, both institutional and retail, which generate large trading volumes using sophisticated and well hedged positions.

After uncovering the most popular options strategies used by different investors, we proceed to examine account performance. While we find that investor class matters for profitability and institutions outperform retail investors in general, we show that trading style complexity measures investor skills on a more granular level. Within each of the investor classes, the complexity of an account's dominant strategy is significantly associated with investment performance. For both retail and institutional investors, those using volatility and Greek neutral strategies outperform their peers. In particular, volatility trading is the most profitable strategy in terms of average returns, but it is subject to large downside risk. After calculating Sharpe ratios to adjust for risk exposure, we see that Greek neutral strategies deliver superior performance. On the other hand, we find that retail investors who use simple strategies lose to the rest of the market. Finally, we provide evidence that these trading style effects are persistent and cannot be explained by systematic risk exposure or known behavioral biases.

References

Barber, B. M., Lee, Y., Liu, Y., Odean, T. (2009). Just How Much Do Individual Investors Lose by Trading? Review of Financial Studies 22, 609—632.

Barber, B. M., Lee, Y., Liu, Y., Odean, T. (2014). Do Day Traders Rationally Learn About Their Ability? Working Paper.

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

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

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

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy 81, 637—654.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X. (2019). Tracking Retail Investor Activity. Working Paper.

Bollerslev, T., Tauchen, G., Zhou, H. (2009). Expected stock returns and variance risk premia. Review of Financial Studies 22, 4463—4492.

Boyer, B. H., Vorkink, K. (2014). Stock Options as Lotteries. Journal of Finance 69, 1485—1527.

Carr, P., Wu, L. (2009). Variance risk premiums. Review of Financial Studies 22, 1311—1341.

Célérier, C., Vallée, B. (2016). Catering to Investors Through Product Complexity. Working Paper.

Chaput, J. S., Ederington, L. H. (2003). Option Spread and Combination Trading. Journal of Derivatives 10, 70—88.

Fahlenbrach, R., Sandås, P. (2010). Does information drive trading in option strategies? Journal of Banking and Finance 34, 2370—2385.

Flint, A., Lepone, A., Yang, J. Y. (2014). Do option strategy traders have a disadvantage? Evidence from the Australian options market. Journal of Futures Markets 34, 838—852.

Gervais, S., Odean, T. (2001). Learning to Be Overconfident. Review of Financial Studies 14, 1—27.

Goyal, A., Saretto, A. (2009). Cross-section of option returns and volatility. Journal of Financial Economics 94, 310—326.

Grinblatt, M., Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland's unique data set. Journal of Financial Economics 55, 43—67.

Henderson, B. J., Pearson, N. D. (2011). The dark side of financial innovation: A case study of the pricing of a retail financial product. Journal of Financial Economics 100, 227—247.

Hens, T., Rieger, M. O. (2011). Why do Investors Buy Structured Products? Working Paper.

Hoffmann, A. O.I., Post, T. (2014). Self-attribution bias in consumer financial decision-making: How investment returns affect individuals' belief in skill. Journal of Behavioral and Experimental Economics 52, 23—28.

Hull, J. C. (2018). Options, Futures, and Other Derivatives, 10th Edition. Pearson.

Kaniel, R., Liu, S., Saar, G., Titman, S. (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. Journal of Finance 67, 639—680.

Kaniel, R., Saar, G., Titman, S. (2008). Individual Investor Trading and Stock Returns. The Journal of Finance 63, 273—310.

Kelley, E. K., Tetlock, P. C. (2013). How Wise Are Crowds? Insights from Retail Orders and Stock Returns. Journal of Finance 68, 1229—1265.

Kelley, E. K., Tetlock, P. C. (2017). Retail Short Selling and Stock Prices. Review of Financial Studies 30, 801—834.

Kuo, W., Lin, T. (2013). Overconfident individual day traders: Evidence from the Taiwan futures market. Journal of Banking & Finance 37, 3548—3561.

Kuo, W., Lin, T., Zhao, J. (2015). Cognitive Limitation and Investment Performance: Evidence from Limit Order Clustering. Review of Financial Studies 28, 838—875.

Kuo, W., Lin, T., Zhao, J. (2018). The Correlated Trading and Investment Performance of Individual Investors. Working Paper.

Lakonishok, J., Lee, I., Pearson, N. D., Poteshman, A. M. (2007). Option Market Activity. Review of Financial Studies 20, 813—857.

Li, X., Subrahmanyam, A., Yang, X. (2018). Can financial innovation succeed by catering to behavioral preferences? Evidence from a callable options market. Journal of Financial Economics 128, 38—65.

Li, X., Subrahmanyam, A., Yang, X. (2021). Winners, Losers, and Regulators in a Derivatives Market Bubble. Review of Financial Studies 34, 313—350.

Merton, R. (1973). Theory of rational option pricing. Bell Journal of Economics 4, 141—183.

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? Journal of Finance 53, 1775—1798.

Odean, T. (1999). Do investors trade too much? American Economic Review 89, 1279—1298.

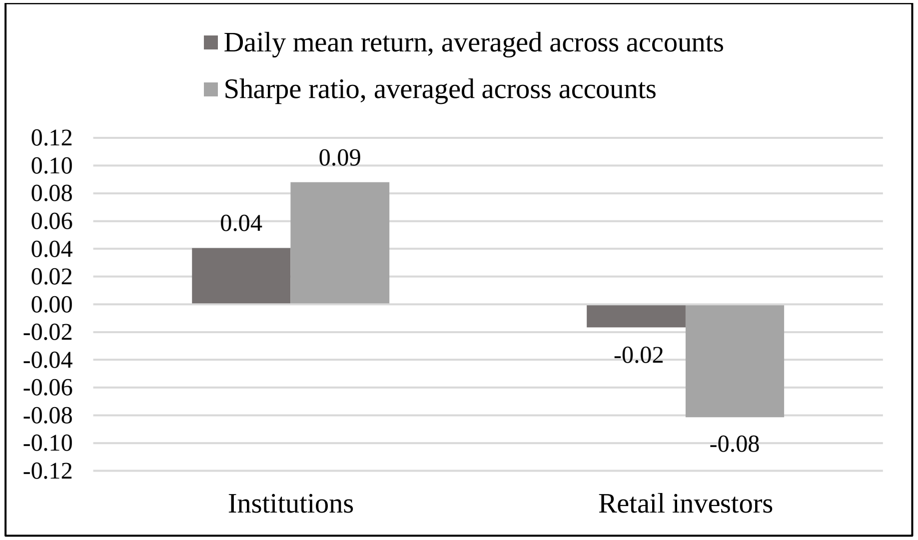
Ryu, D. (2012). The profitability of day trading: An empirical study using high-quality data. Investment Analysts Journal 41, 43—54.

Welch, I. (2020). The Wisdom of the Robinhood Crowd. Working Paper.

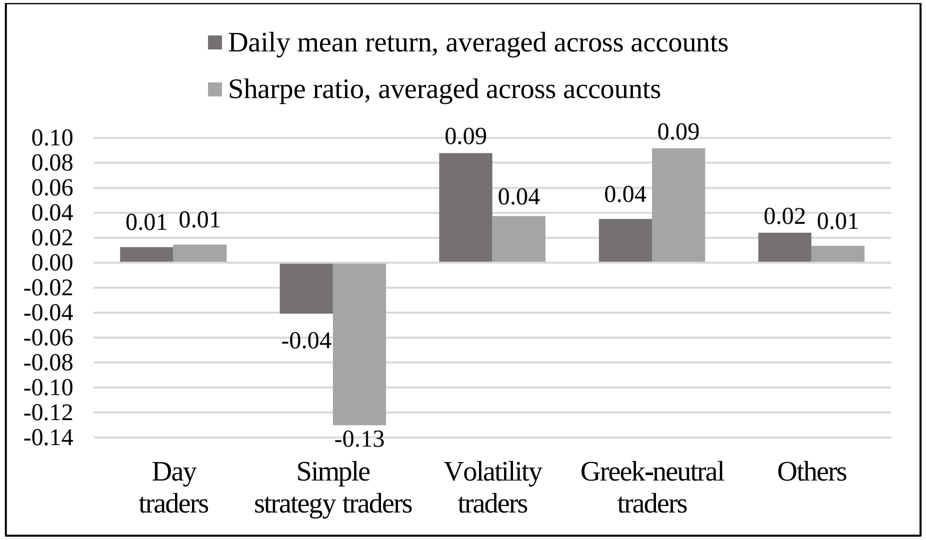
Figure 1
Profitability of the different investor accounts

This figure plots the average profitability of 122,620 active accounts in the KOSPI 200 options market from 1 January 2010 to 30 June 2014. It shows daily mean return and Sharpe ratio, averaged across accounts in each category of investors. Refer to Section 4 for details about our method for calculating daily returns. Panel A shows results by investor class; Panel B shows results by trading style; and Panel C shows results by trading style in the different investor class subsamples.

Panel A: Results by investor class

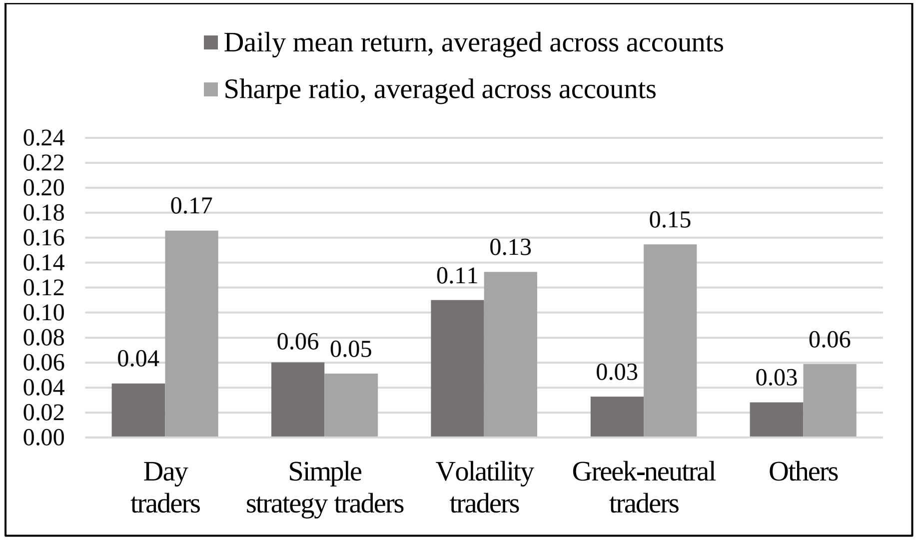


Panel B: Results by trading style



Panel C: Results by trading style in the different investor class subsamples

C.1. Institutions



C.2. Retail investors

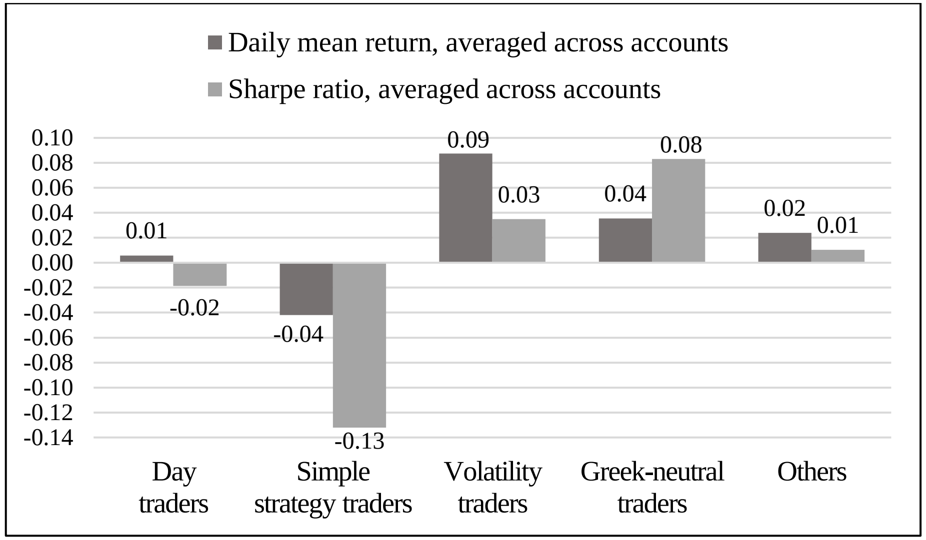


Table 1
Summary statistics

This table reports aggregate summary statistics of KOSPI 200 options for contracts introduced after 1 January 2010 and expire before 30 June 2014. Reported are time-series statistics for daily total number of transactions, total trading volume (number of contracts traded), and total premium in billions of KRW (option premium multiplied by number of contracts traded and multiplier). The subpanels contain average daily statistics of total activity for different subsamples. We classify a call option as out of the money (OTM) / at the money (ATM) / in the money (ITM) if its spot-to-strike price ratio is less than 0.95 / between 0.95 and 1.05 / greater than 1.05. We define put options to be ITM, ATM, and OTM in the same spot-to-strike price ration regions, correspondingly. Normal trading hours refer to the daily continuous trading session from 9:00 to 15:05.

Panel A: Aggregate options data

	N transactions	**Trading volume**	**Premium (billion KRW)**
daily mean	905,026	7,783,991	1,197
std	505,100	7,664,950	597
min	984	4,342	1.61
median	804,878	5,054,738	1,100
max	3,360,618	42,188,606	6,277
total over the sample period	1,000,958,323	8,609,093,878	1,323,552

	N transactions daily mean	**Trading volume** daily mean	**Premium (billion KRW)** daily mean
Panel B: By option moneyness			
OTM	183,092	2,251,366	163
ATM	716,845	5,519,158	1,014
ITM	5,108	13,516	19.3
Panel C: By option contract type			
Call	454,010	4,059,280	584
Put	451,016	3,724,711	613
Panel D: By trading hours			
Normal trading hours	897,867	7,717,283	1,187
Outside normal trading hours	7,158	66,708	9.6
Panel E: By time to maturity			
0-40 days to maturity	918,363	8,001,045	1,213
41-70 days to maturity	18,749	66,504	21.3
>70 days to maturity	2,488	10,665	7.6

Table 2
Average daily options trading activity of different investor types

This table describes options trading activity of different investor types in the KOSPI 200 options market from 1 January 2010 to 30 June 2014. For each account and each day, we calculate the number of transactions, trading volume, and premium. Reported are daily averages for the aggregate trading activity of all accounts in each category of investors, as well as the average account activity for that category of investors. We report statistics for the trading activity of liquidity takers and liquidity providers separately. We compare the bid and ask order submission times and mark the investor who submitted their order first as the liquidity provider and the investor whose order matched the first one as the liquidity taker.

	N transactions of average account	Aggregate N transactions	Trading volume of average account	Aggregate trading volume	Premium of average account (million KRW)	Aggregate premium (million KRW)
Liquidity takers						
Institutions	1,123	626,690	10,085	5,915,011	1,688	930,483
Retail investors	18.6	278,203	108.9	1,868,697	18.5	266,045
Liquidity providers						
Institutions	811.7	451,009	8,230	4,816,273	1,244	682,001
Retail investors	27.0	453,883	157.1	2,967,435	31.9	514,527

Table 3
Strategies of position holders

This table examines the popularity of different options strategies extracted from end-of-day positions in the KOSPI 200 options market from 1 January 2010 to 30 June 2014. For the whole market, as well as for the two investor classes separately, we report the frequency of account-day positions in each options strategy as a percentage of all account-days with non-zero end-of-day positions. We also report the number of option contracts held in each strategy as a percentage of total end-of-day open positions.

	All		**Institutions**		**Retail investors**	
N account-days with a position in options	19,135,324		539,680		18,595,644	
N option contracts held in open positions	4,822,669,794		2,916,207,458		1,906,462,336	
	% of all account-days	% of all open positions	% of instit. account-days	% of instit. open positions	% of retail account-days	% of retail open positions
Combinations of options and futures	6.24%	37.6%	34.4%	55.4%	5.42%	10.5%
Simple strategies	55.7%	7.14%	18.8%	1.35%	56.8%	16.0%
Volatility trading strategies	16.5%	7.24%	9.43%	2.90%	16.7%	13.9%
Spreads	3.68%	2.89%	9.54%	2.65%	3.51%	3.25%
Other options strategies	17.9%	45.1%	27.8%	37.7%	17.6%	56.4%

Table 4
Number of accounts and trading volume by trading style and investor class

In this table, we report the dominant strategy of 122,620 active accounts in the KOSPI 200 options market from 1 January 2010 to 30 June 2014. Day traders are defined as accounts that only trade intraday and never hold positions overnight throughout our sample. For the remaining investors, the dominant strategy of an account is the strategy used in at least 60% of the end-of-day positions held by that account. The trading strategies are the same as defined in Table 3. Accounts with more than 60% of end-of-day positions unidentifiable as any of the strategies, as well as accounts that switch between different strategies without having a dominant one, are labelled as "others". We report the number of accounts for each strategy in each investor class category, and their percentage of all accounts in the respective category. We also report the cross-sectional averages of mean daily trading volume of different accounts, and the percentage of each strategy's aggregate volume in total volume of the same investor class.

	All		Institutions		Retail investors	
	N accounts % of all accounts	Mean daily volume % of all volume	N accounts % of instit. accounts	Mean daily volume % of instit. volume	N accounts % of retail accounts	Mean daily volume % of retail volume
Day traders	3,747 3.1%	2,352 10.3%	677 17.0%	10,828 13.8%	3,070 2.6%	483.2 3.3%
Simple strategy traders	79,037 64.5%	323.8 37.1%	760 19.1%	8,115 26.2%	78,277 66.0%	248.2 58.7%
Volatility traders	6,106 5.0%	466.4 4.3%	158 4.0%	7,751 2.1%	5,948 5.0%	272.9 8.6%
Greek neutral traders	3,249 2.6%	2,761 17.9%	392 9.9%	20,532 24.7%	2,857 2.4%	322.8 4.4%
Others	30,481 24.9%	521.1 30.4%	1,985 50.0%	4,944 33.2%	28,496 24.0%	213.0 25.0%

Table 5
Summary statistics of account profitability

This table reports the percentiles of three different profitability measures. We look at 122,620 active accounts that trade options. First, for each account, we use the transactions records to calculate the total cumulative profit and loss ($P&L) over the whole sample period from 1 January 2010 to 30 June 2014. We use the contract multiplier to calculate profits in Korean won (KRW), and we report total account $P&L in millions of KRW. Then, for each account and each trading day, we calculate profitability per dollar as:

$$\text{Daily return} = \text{Total daily \$P\&L} \,/\, \text{Total daily capital requirement}$$
$$= \frac{(\text{ \$P\&L from end-of-day positions} + \text{\$P\&L from intraday trades })}{(\text{ Lag of end-of-day capital requirement} + \text{Intraday capital requirement })}.$$

Refer to Section 4 for more details about our method for calculating daily returns. This scaled P&L measure is a rate of return to each dollar of capital invested in KOSPI derivatives. To have one single return measure for each account, we take the mean of daily returns of the same account during the whole sample period. We also calculate Sharpe ratio of each account as the mean daily return divided by the standard deviation of daily returns. Panel A shows account profitability of institutions versus retail investors, and Panel B shows the profitability by trading style.

Panel A: By investor class

Profitability measure		**Percentiles**										
		p0	p10	p20	p30	p40	p50	p60	p70	p80	p90	p100
Total account $P&L (million KRW)	Institutions	-207,524	-215.6	-46.7	-10.5	-0.3	8.8	29.8	88.5	290.2	971.5	234,201
	Retail investors	-61,538	-74.4	-34.1	-18.9	-10.7	-5.5	-2.4	-0.8	0.2	8.0	68,055
Mean daily return	Institutions	-2.00	-0.04	-0.01	0.00	0.00	0.01	0.02	0.04	0.08	0.16	2.20
	Retail investors	-24.1	-0.11	-0.07	-0.05	-0.03	-0.02	-0.01	0.00	0.02	0.07	13.7
Sharpe ratio	Institutions	-1.72	-0.14	-0.05	0.00	0.03	0.07	0.10	0.15	0.22	0.34	1.48
	Retail investors	-9.19	-0.31	-0.22	-0.16	-0.11	-0.07	-0.03	0.01	0.06	0.14	29.0

Panel B: By trading style

Profitability measure		Percentiles										
		p0	p10	p20	p30	p40	p50	p60	p70	p80	p90	p100
Total account $P&L (million KRW)	Day traders	-3,631	-20.9	-9.3	-4.1	-1.8	-0.7	-0.2	0.1	3.4	40.4	46,640
	Simple strategy traders	-38,459	-73.2	-34.5	-19.7	-11.6	-6.4	-3.3	-1.4	-0.4	0.8	70,224
	Volatility traders	-61,538	-48.7	-18.4	-8.5	-3.6	-1.0	0.4	3.5	11.8	42.2	41,895
	Greek neutral traders	-9,139	-93.8	-26.5	-7.7	-0.9	1.8	7.6	18.3	43.9	157.7	45,701
	Others	-207,524	-97.3	-41.9	-22.1	-11.7	-5.0	-1.0	2.3	11.2	48.4	234,201
Mean daily return	Day traders	-0.71	-0.05	-0.03	-0.01	0.00	0.00	0.01	0.02	0.03	0.08	1.33
	Simple strategy traders	-24.1	-0.12	-0.09	-0.07	-0.05	-0.04	-0.02	-0.01	0.00	0.03	13.7
	Volatility traders	-3.53	-0.06	-0.03	-0.01	0.00	0.02	0.06	0.11	0.19	0.35	3.59
	Greek neutral traders	-1.68	-0.01	0.00	0.00	0.01	0.02	0.04	0.05	0.08	0.11	0.52
	Others	-5.81	-0.05	-0.02	-0.01	0.00	0.00	0.01	0.03	0.06	0.13	2.10
Sharpe ratio	Day traders	-3.52	-0.34	-0.21	-0.13	-0.06	0.01	0.07	0.13	0.22	0.37	11.6
	Simple strategy traders	-9.19	-0.35	-0.26	-0.20	-0.16	-0.12	-0.08	-0.04	0.00	0.08	29.0
	Volatility traders	-1.37	-0.21	-0.11	-0.04	0.01	0.05	0.09	0.13	0.19	0.27	2.84
	Greek neutral traders	-1.07	-0.09	-0.01	0.03	0.06	0.10	0.13	0.16	0.19	0.25	1.21
	Others	-1.72	-0.17	-0.09	-0.05	-0.01	0.02	0.05	0.08	0.12	0.19	2.59

Table 6
Profitability of the different investor accounts: Regression analyses

This table presents multivariate regression analyses of the relation between account profitability and investor types. We include the 122,620 active accounts that trade options. We perform six account-level regressions. Regressions 1, 2, 3 use mean daily return as the dependent variable, and regressions 4, 5, 6 use Sharpe ratio. The independent variable in regressions 1 and 4 is a dummy variable for institutional investors, and the intercept represents retail investors. The independent variables in regressions 2 and 5 are the trading style dummies for day traders, simple strategy traders, volatility traders and Greek neutral traders, and the intercept represents the remaining accounts in the category "others". Regressions 3 and 6 use the investor class dummy and the trading style dummies at the same time. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Dependent variable:	Mean daily return			Sharpe ratio		
	1	2	3	4	5	6
Intercept	-0.017***	0.024***	0.022***	-0.081***	0.014***	0.007***
	(-34.83)	(26.42)	(23.90)	(-122.4)	(10.77)	(5.44)
Institution	0.057***		0.030***	0.169***		0.101***
	(21.67)		(11.24)	(45.84)		(28.04)
Day trader		-0.012***	-0.015***		0.001	-0.011***
		(-4.17)	(-5.38)		(0.28)	(-2.79)
Simple strategy trader		-0.065***	-0.063***		-0.144***	-0.138***
		(-60.52)	(-58.48)		(-97.52)	(-93.16)
Volatility trader		0.064***	0.065***		0.024***	0.028***
		(28.59)	(29.09)		(7.80)	(9.11)
Greek neutral trader		0.011***	0.009***		0.078***	0.073***
		(3.75)	(3.19)		(19.37)	(18.02)
N observations	122,620	122,620	122,620	122,617	122,617	122,617
Adjusted R^2	0.0038	0.0533	0.0543	0.0168	0.1036	0.1093

Table 7
Robustness checks

Here, we repeat the profitability regressions of Table 6 using different definitions of an account's dominant strategy, and different measures of profitability. Regressions 1 and 2 use risk-adjusted returns and Sharpe ratio of risk-adjusted returns as the dependent variables. For each account, risk-adjusted return is equal to the intercept from time-series regressions of the account's daily return on the KOSPI 200 index return and implied volatility on the same day. Sharpe ratio is equal to the t-statistic of the intercept from the same regressions. The daily implied volatility is calculated as the price-weighted average implied volatility of all ATM options, derived from the last daily transactions executed in the 15 minutes before market close. To prevent the effects of outliers, we winsorize the dependent variables cross-sectionally at 5% and 95%. The dependent variables in regressions 3 and 4 are mean daily return from positions and Sharpe ratio of returns from positions. Here, we exclude position holders' intraday trading profits when calculating daily returns. For regressions 5 and 6, we define dominant strategies and calculate mean daily returns excluding any simple positions in short call and long put contracts, in order to exclude any possible hedges of unobserved equity positions. Regressions 7-10 show the results when dominant strategies are defined using different cut-off points. Finally, regression 11 reports the results of a panel regression using account-day observations instead of account-level. The independent variables in all regressions are the same as in Table 6. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Definition of dominant strategy:	60% cut-off dummy				Exclude simple positions in short call and long put		70% cut-off dummy		80% cut-off dummy		None (panel regression)
Dependent variable:	Risk-adjusted return	Sharpe ratio of risk-adj. returns	Mean daily return from positions	Sharpe ratio of position returns	Mean daily return	Sharpe ratio	Mean daily return	Sharpe ratio	Mean daily return	Sharpe ratio	Return
	1	2	3	4	5	6	7	8	9	10	11
Intercept	0.027***	0.100***	0.008***	-0.028***	0.018***	0.006***	0.018***	-0.004***	0.013***	-0.018***	0.021***
	(14.54)	(14.66)	(3.80)	(-6.41)	(17.97)	(4.57)	(22.1)	(-3.93)	(17.12)	(-17.85)	(64.78)
Institution	0.000	0.158***	0.034***	0.095***	0.027***	0.096***	0.031***	0.107***	0.034***	0.115***	0.024***
	(-0.09)	(8.13)	(5.93)	(7.46)	(8.26)	(23.07)	(11.85)	(29.59)	(13.01)	(31.95)	(20.68)
Day trader	-0.033***	-0.172***	-0.001	0.026*	-0.010***	-0.009**	-0.011***	0.000	-0.006**	0.012***	0.000
	(-5.85)	(-8.41)	(-0.21)	(1.96)	(-3.19)	(-2.05)	(-4.09)	(-0.07)	(-2.24)	(3.12)	(0.17)
Simple strategy trader	-0.041***	-0.362***	-0.082***	-0.175***	-0.048***	-0.104***	-0.061***	-0.130***	-0.057***	-0.121***	-0.076***
	(-18.87)	(-45.36)	(-34.85)	(-34.1)	(-38.41)	(-64.56)	(-60.5)	(-94.2)	(-59.64)	(-91.81)	(-164.0)
Volatility trader	0.082***	0.187***	0.068***	0.001	0.038***	0.000	0.082***	0.052***	0.100***	0.078***	0.054***
	(18.26)	(11.36)	(13.91)	(0.09)	(17.89)	(0.06)	(32.0)	(14.71)	(32.52)	(18.43)	(79.86)
Greek neutral trader	0.017***	0.356***	0.023***	0.092***	0.013***	0.075***	0.014***	0.091***	0.019***	0.110***	0.036***
	(2.94)	(16.43)	(3.64)	(6.60)	(3.91)	(16.99)	(3.84)	(18.6)	(3.82)	(16.33)	(37.02)
N observations	122,620	122,431	121,873	121,780	118,849	118,817	122,620	122,617	122,620	122,617	25,771,249
Adjusted R^2	0.0090	0.0314	0.0190	0.0153	0.0251	0.0576	0.0526	0.1048	0.0474	0.0956	0.0020

Table 8
Profitability of the different investor accounts: Regression analyses in subsamples

This table repeats the profitability regressions of Table 6 in the subsamples of institutions and retail investors separately. All variables are the same as in Table 6. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Subsample:	Institutions		Retail investors	
Dependent variable:	Mean daily return	Sharpe ratio	Mean daily return	Sharpe ratio
	1	2	3	4
Intercept	0.028***	0.059***	0.024***	0.010***
	(8.65)	(11.2)	(25.2)	(8.04)
Day trader	0.015**	0.107***	-0.018***	-0.029***
	(2.33)	(10.25)	(-5.94)	(-7.05)
Simple strategy trader	0.032***	-0.008	-0.066***	-0.142***
	(5.14)	(-0.77)	(-59.53)	(-94.78)
Volatility trader	0.082***	0.074***	0.064***	0.025***
	(6.8)	(3.82)	(27.95)	(7.92)
Greek neutral trader	0.005	0.096***	0.012***	0.073***
	(0.57)	(7.4)	(3.72)	(17.06)
N observations	3,972	3,971	118,648	118,646
Adjusted R^2	0.0152	0.0382	0.0533	0.0983

Table 9
Profitability from the dominant strategy versus non-dominant strategies

This table shows the results of panel regressions using account-day returns of institutions and retail investors as the dependent variable. We regress those on the same dummies for dominant trading style of each account. However, an account may use strategies other than their dominant one on some days. Hence, we add non-dominant strategy dummies equal to 1 on such days. For example, *Simple strategy trader using non-dominant strategy* is a dummy variable equal to 1 on days when a simple strategy trader holds positions in volatility, Greek neutral, or other strategies, and equal to 0 on days when the account holds simple positions. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Subsample:	Institutions	Retail investors
Dependent variable:	Return	Return
	1	2
Intercept	0.027***	0.034***
	(19.48)	(88.47)
Day trader	0.030***	-0.018***
	(8.72)	(-7.87)
Simple strategy trader	0.055***	-0.097***
	(13.17)	(-186.3)
Simple strategy trader using non-dominant strategy	-0.023***	0.074***
	(-4.64)	(133.5)
Volatility trader	0.207***	0.126***
	(29.49)	(107.9)
Volatility trader using non-dominant strategy	-0.173***	-0.108***
	(-15.47)	(-57.83)
Greek neutral trader	-0.006	0.000
	(-1.44)	(-0.05)
Greek neutral trader using non-dominant strategy	0.056***	0.038***
	(7.67)	(13.67)
N observations	881,443	24,889,806
Adjusted R^2	0.0012	0.0025

Table 10
Out-of-sample trading style persistence

This table presents transition matrixes containing the percentage of accounts by trading style in-sample and out-of-sample. We divide the sample into two equal subperiods and follow the same method as before using 60% cut-offs to extract the dominant strategy of each account in the two subperiods. Therefore, in-sample (out-of-sample) trading style of each account is its dominant strategy in the first (second) half of the sample. We only keep accounts that trade in both subperiods. Panel A shows the transition matrix for the trading styles of institutions, and Panel B for retail investors. The diagonals reflect trading style persistence. The sample time period is from 1 January 2010 to 30 June 2014.

Panel A: Institutions

		Out-of-sample trading style			
		Simple strategy trader	**Volatility trader**	**Greek neutral trader**	**Other**
In-sample trading style	**Simple strategy trader**	75.8%	1.3%	0.6%	22.3%
	Volatility trader	7.9%	68.4%	2.6%	21.1%
	Greek neutral trader	1.3%	0.0%	68.0%	30.7%
	Other	5.8%	1.0%	10.5%	82.7%

Panel B: Retail investors

		Out-of-sample trading style			
		Simple strategy trader	**Volatility trader**	**Greek neutral trader**	**Other**
In-sample trading style	**Simple strategy trader**	89.3%	1.5%	0.1%	9.2%
	Volatility trader	14.7%	47.2%	1.0%	37.1%
	Greek neutral trader	3.2%	1.1%	67.2%	28.5%
	Other	23.2%	5.2%	4.6%	67.0%

Table 11
Out-of-sample performance persistence

In the same way as in Table 10, we split our sample into two equal subperiods and define the dominant strategy of each account in the first half of the sample. Hence, we mark Simple strategy traders, Volatility traders, and Greek neutral traders in-sample. For each account, we calculate mean daily return and Sharpe ratio in the second half of the sample. We then test the relation of in-sample trading styles with out-of-sample account performance. We only keep accounts that trade in both subperiods, and we perform separate regressions in the subsamples of institutions and retail investors. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Subsample:	Institutions		Retail investors	
Dependent variable:	Mean daily return out-of-sample	Sharpe ratio out-of-sample	Mean daily return out-of-sample	Sharpe ratio out-of-sample
	1	2	3	4
Intercept	0.029***	0.068***	0.023***	0.018**
	(3.73)	(5.88)	(14.67)	(2.03)
Day trader	0.014	0.097***	-0.012***	0.005
	(1.26)	(5.81)	(-2.62)	(0.20)
Simple strategy trader in-sample	0.037**	-0.011	-0.066***	-0.166***
	(2.36)	(-0.47)	(-33.42)	(-15.55)
Volatility trader in-sample	0.287***	0.142***	0.098***	0.042**
	(9.95)	(3.27)	(26.12)	(2.07)
Greek neutral trader in-sample	0.007	0.062*	0.018***	0.060*
	(0.34)	(1.94)	(2.87)	(1.79)
N observations	1,242	1,238	34,406	34,379
Adjusted R^2	0.0729	0.0335	0.0759	0.0098

Table 12
Effect of behavioral biases on retail investors' trading style

This table reports logistic regressions for the probability of retail investors using a particular trading style given variables which are proxies for well-known behavioral biases potentially related to retail behavior and profitability. *Trading volume* refers to the logarithm of average daily options trading volume, which serves as a proxy for overconfidence or excessive trading. *Disposition effect* is measured as an account's proportion of gains realized (PGR) minus proportion of losses realized (PLR). On a given day, PGR is the total number of contracts with realized gains (positions closed at a gain) divided by that number plus the total number of contracts with paper gains (open positions which have increased in value from the previous day). PLR is calculated similarly. Then, we sum the daily values over the sample time period, and we take the difference of PGR and PLR. (Disposition effect is 0 for day traders). *OTM contracts bought* refers to the percentage of positions in long OTM contracts (for position holders) or the percentage of opening intraday trades buying OTM contracts (for day traders). This is a measure of lottery preference. *Sensitivity of delta to KOSPI 200 index return* is derived by regressing an account's end-of-day position delta scaled by position size (for position holders), or the volume-weighted average delta of opening intraday trades (for day traders), on past 5-day average index return. This is a proxy for trend chasing (extrapolation) behavior. *Sensitivity of investor activity to performance* is derived by regressing position size (for position holders) or intraday trading volume (for day traders) on past 5-day average return of the investor. This is a proxy for self-attribution bias. The *% of volume initiated* is a proxy for transaction costs. Finally, we include a dummy equal to 1 if the account is a foreign investor. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Subsample:	Retail investors		
Dependent variable:	Simple strategy trader	Volatility trader	Greek neutral trader
	1	2	3
Intercept	0.244***	-1.995***	-5.618***
	(8.97)	(-36.21)	(-70.01)
Trading volume	0.034***	-0.124***	0.252***
	(7.79)	(-13.36)	(21.27)
Disposition effect	0.868***	-0.314***	-0.954***
	(25.8)	(-4.48)	(-10.28)
OTM contracts bought	2.267***	-3.876***	-0.416***
	(87.14)	(-48.85)	(-5.98)
Sensitivity of delta to KOSPI 200 index return	-1.265*	2.864**	3.869**
	(-1.72)	(2.11)	(2.20)
Sensitivity of investor activity to performance	-2.531*	-1.692	-3.439*
	(-1.88)	(-0.57)	(-1.78)
% of volume initiated	-0.978***	0.778***	1.937***
	(-33.06)	(13.43)	(26.79)
Foreign investor	-0.393***	0.330*	0.844***
	(-3.89)	(1.77)	(4.18)
N observations	113,915	113,915	113,915

Table 13
Do behavioral biases drive the profitability results?

This table shows regressions of retail investors' profitability on the trading style dummies with added control variables for each of the behavioral hypotheses. See Table 12 for the definitions of the behavioral proxies. The table contains the estimated regression coefficients and below them the corresponding t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample time period is from 1 January 2010 to 30 June 2014.

Subsample:	Retail investors	
Dependent variable:	Mean daily return	Sharpe ratio
	1	2
Intercept	0.045***	0.071***
	(21.16)	(29.49)
Day trader	-0.026***	-0.045***
	(-8.69)	(-13.00)
Simple strategy trader	-0.057***	-0.128***
	(-50.59)	(-99.14)
Volatility trader	0.057***	0.014***
	(25.09)	(5.42)
Greek neutral trader	0.015***	0.081***
	(4.85)	(23.01)
Trading volume	-0.002***	-0.005***
	(-7.09)	(-15.05)
Disposition effect	-0.011***	-0.045***
	(-4.62)	(-16.91)
OTM contracts bought	-0.056***	-0.087***
	(-32.47)	(-44.51)
Sensitivity of delta to KOSPI 200 index return	0.020	0.042
	(0.60)	(1.09)
Sensitivity of investor activity to performance	0.006	0.672***
	(0.06)	(6.42)
% of volume initiated	0.004**	-0.030***
	(2.07)	(-12.33)
Foreign investor	0.007	0.024***
	(0.94)	(2.76)
N observations	113,915	113,915
Adjusted R^2	0.0646	0.1558

Appendix

Table I
Strategies of position holders

This table is an extended version of Table 3. It provides a break-down of the different options strategies in each of the five main categories.

	All		Institutions		Retail investors	
Number of account-days with a position in options	19,135,324		539,680		18,595,644	
Number of option contracts held in open positions	4,822,669,794		2,916,207,458		1,906,462,336	
	% of all account-days	% of all open positions	% of instit. account-days	% of instit. open positions	% of retail account-days	% of retail open positions
Combinations of options and futures	**6.24%**	**37.6%**	**34.4%**	**55.4%**	**5.42%**	**10.5%**
● Long covered calls	0.21%	0.10%	0.44%	0.11%	0.21%	0.08%
● Short covered calls	0.28%	0.14%	1.35%	0.08%	0.25%	0.24%
● Long protective puts	0.24%	0.12%	0.93%	0.14%	0.22%	0.08%
● Short protective puts	0.25%	0.13%	0.86%	0.06%	0.24%	0.25%
● Greek neutral	0.59%	14.8%	7.20%	22.6%	0.40%	2.83%
● Other combinations	4.66%	22.4%	23.6%	32.4%	4.11%	7.02%
Simple strategies	**55.7%**	**7.14%**	**18.8%**	**1.35%**	**56.8%**	**16.0%**
● Long call	30.2%	3.68%	3.62%	0.24%	31.0%	8.94%
● Short call	2.04%	0.43%	6.28%	0.36%	1.92%	0.54%
● Long put	21.3%	2.61%	6.05%	0.66%	21.7%	5.58%
● Short put	2.17%	0.42%	2.88%	0.08%	2.15%	0.94%

(continued)

Table I (continued)

	All		Institutions		Retail investors	
	% of all account-days	% of all open positions	% of instit. account-days	% of instit. open positions	% of retail account-days	% of retail open positions
Volatility trading strategies	**16.5%**	**7.24%**	**9.43%**	**2.90%**	**16.7%**	**13.9%**
• Long volatility	8.32%	1.75%	1.84%	0.94%	8.51%	2.99%
• straddle	0.18%	0.02%	0.04%	0.00%	0.18%	0.04%
• strangle	4.67%	0.49%	0.64%	0.05%	4.78%	1.15%
• butterfly	0.03%	0.01%	0.10%	0.01%	0.03%	0.01%
• combinations	3.44%	1.23%	1.05%	0.88%	3.51%	1.78%
• Short volatility	8.14%	5.49%	7.59%	1.95%	8.16%	10.9%
• straddle	0.20%	0.02%	0.43%	0.01%	0.19%	0.04%
• strangle	2.56%	0.90%	1.52%	0.16%	2.59%	2.01%
• butterfly	0.01%	0.00%	0.01%	0.00%	0.01%	0.00%
• combinations	5.37%	4.57%	5.64%	1.78%	5.37%	8.84%
Spreads	**3.68%**	**2.89%**	**9.54%**	**2.65%**	**3.51%**	**3.25%**
• Long synthetic stock	0.46%	0.12%	1.97%	0.07%	0.41%	0.20%
• Short synthetic stock	0.42%	0.19%	1.57%	0.22%	0.39%	0.14%
• Bull / bear spread	2.78%	2.58%	5.98%	2.37%	2.68%	2.90%
• Calendar spread	0.03%	0.00%	0.01%	0.00%	0.03%	0.01%
Other options strategies	**17.9%**	**45.1%**	**27.8%**	**37.7%**	**17.6%**	**56.4%**
• Greek neutral	6.49%	25.8%	14.0%	21.8%	6.27%	32.1%
• Others	11.4%	19.3%	13.8%	16.0%	11.4%	24.3%

Chapter III: Rational Regulation Meets Irrational Investors

1. Introduction

Regulators around the world frequently introduce reforms aimed to protect financial market participants, particularly the less sophisticated retail investors. However, if regulations are designed for individuals that are assumed to be perfectly rational agents, in accordance with traditional economic theories, will they benefit or harm investors whose behavior deviates from perfect rationality? In this study, we show that reforms catered towards retail investors can harm rather than help them, if regulators do not consider these investors' biases and irrational behaviors.

In August 2011, the European debt crisis, combined with a credit rating downgrade of the United States sovereign debt, caused markets around the world to drop sharply in value. This included Korea's main equity index, the KOSPI 200, which experienced a severe crash. In the month of the market crash, many retail investors lost large amounts of money in the Korean equity index options market, and many even exited the market altogether. In response to the crisis, Korean regulators decided to increase the contract size for KOSPI 200 index options. The multiplier was raised five times in magnitude, from KRW 100,000 to KRW 500,000, for contracts that mature after June 2012. Given the index value around 250 at the end of 2011, the reform would increase the notional value of the underlying asset from about 25,000 USD to about 125,000 USD. In comparison, the notional value of the underlying asset for the most popular index options in the U.S., the E-mini S&P 500 index options, was about 62,500 USD at the same time. The Korean reform was motivated by a desire to limit retail participation and excessive speculation in the market. Up until then, the KOSPI options market attracted a significantly large number of retail investors. By increasing the capital required for trading options fivefold, the regulators meant to prevent the smallest retail accounts from participating in the market. The reform had the potential to protect unsophisticated retail investors from experiencing large losses such as those that occurred in the month of the market crash. However, we show that the reform did not drive significantly more investors out of the market. Moreover, the reform had a negative impact on the remaining investors that stayed in the market, because regulators did not anticipate their irrational reaction.

We hypothesize that retail investors are subject to certain biases which cause their behavior to deviate from perfect rationality. Firstly, we hypothesize that retail investors are subject to self-attribution bias (see e.g. Gervais, Odean (2001); Hoffmann, Post (2014)). They believe that their past success is due to skill rather than luck, which causes them to extrapolate their good performance into the future and to assume that they will continue to be profitable. For this reason, we expect that retail accounts with good performance before the crisis are more likely to remain in the market and trade more during and after the crisis, but not necessarily profiting from it. Our second hypothesis is about retail investors' reaction to the reform. In economics, a rational agent is assumed to be an individual who always chooses to perform actions which will result in the optimal expected outcome. Therefore, when the contract size rises from KRW100,000 to KRW500,000, we would expect a rational investor to decrease his trading volume fivefold. For example, if an account used to always trade 100 option contracts before the reform, the optimal number of contracts for him to trade after the reform is 20. To account for the change in contract size and to make the pre- and post-reform periods in our data comparable, we multiply post-reform trading volumes by a factor of five. Hence, if investors react rationally to the reform, we should not see any significant change in volumes after adjusting for the changes. However, if some investors are not paying attention to regulatory announcements about market changes, we would expect to see increased trading volumes after the reform is introduced. Following the literature on limited attention and cognitive processing (e.g. Hirshleifer (2015); Barber, Odean (2008); Kuo, Lin, Zhao (2015)), we hypothesize that retail investors are subject to these biases. We predict that they will react irrationally to the reform by failing to adjust their trading behavior for the new contract size. As a result, their post-reform risk exposure will rise unintentionally, and their profitability may suffer. Over time, it is likely that all market participants learn about the reform, so we would expect the negative effects of the reform to dissipate after a while. Contrary to retail investors, we hypothesize that institutions are not subject to the same behavioral biases. As more sophisticated investors, we expect them to behave rationally and pay attention to ongoing market changes.

We use transaction-level data with account IDs and investor type codes from the KOSPI 200 options market. The Korean derivatives market is one of the world's most active and liquid markets, with volumes comparable to those in the U.S. markets. The data contains all options trades of domestic and foreign institutions and retail investors from January 2010 to June 2014. Our analysis benefits from the fact that the Korean market attracts high retail participation, and we are able to track investors' trading activities at the account level. In total, there are 6,287 institutional accounts and 153,842 retail accounts in the data.

In our analyses of the periods around the two events, we separate accounts into three groups: accounts that exit the market, existing accounts which stay in the market, and new accounts that enter the market. The group which we call "existing accounts" consists of those investors who exist in the market starting from four months before the crisis and remain in the market for at least one year and a half until four months after the contract size reform. We analyze this group separately in order to examine the effects of both the crisis and the reform on long-term investors that choose to stay in the market regardless of the turbulence and changes that take place.

Overall, we find supporting evidence for our main hypotheses. First, we show that there is a sharp rise in the number of accounts that exit the market in the month of the market crash, consistent with a panic selling effect. In that month, the propensity of retail investors to stay in the market is significantly lower, while there is no significant change in institutions' propensity to stay. The retail investors that exit are the group with the lowest returns before the crisis. Both institutional and retail accounts that stay in the market during the crisis tend to be more active at that time, and existing accounts increase their trading volumes in that month. In addition, higher previous-month returns are associated with a higher likelihood to stay in the market and be active in the following month. Previous-month return is also significantly positively associated with trading volume in the next month, but only in the subsample of retail accounts. This is consistent with self-attribution bias and biased extrapolation – investors believe that their past success is due to skill rather than luck and assume that they will continue to be profitable in the future. This effect is larger during the crisis, consistent with some literature that suggests behavioral biases are amplified in times of market stress (e.g. Kallberg, Liu and Wang (2014)). This indicates that retail investors who perform well before the crisis extrapolate their good performance into the future, which causes them to trade more and experience losses during the crisis. There is no such effect in the subsample of institutions. The aggregate losses of retail investors in the month of the crash amounted to almost 1,000 billion KRW, while institutions gained more than 800 billion KRW in total, although most of the large gains went to only a few institutions. In terms of average returns, both institutions and retail investors are negatively affected in the crisis. The post-crisis effects do not appear to be significant, suggesting that the negative effects of the crisis are concentrated in the month of the market crash, but the investors who remain in the market, as opposed to exiting, recover quickly afterwards.

When the contract size reform is introduced, the number of accounts that exit the market spikes up again, but the increase is much less pronounced than in the month of the crash. Unsurprisingly, some retail investors exit. The smallest accounts which used to trade only a few lots at a time are unable to continue trading after capital requirements increase fivefold. The retail investors who exit have dollar losses and negative returns on average before the reform. However, in general, we find that investors' propensity to exit the market decreases after the reform. Moreover, we show that the reform has a negative effect on the existing retail accounts that stay in the market long-term. We multiply post-reform volumes by five to adjust for the fivefold increase in options contract size. Therefore, if investors rationally adjust their trading volumes after the reform, we should not expect to see any significant changes in their trading behavior. Indeed, we find no significant effect of the reform on the trading volumes of existing institutions. They continue trading in a similar way as they did before the reform, adjusting their volumes correctly in accordance with the contract size change. On the other hand, when we look at the existing retail accounts that stay in the market, we observe a significant rise in their trading volumes after the reform, which is inconsistent with a rational reaction to the contract size increase. Instead, it is consistent with retail inattention to the reform. Due to such inattention, retail accounts fail to adjust their trading behavior for the new contract size, which negatively impacts their profitability. Existing retail accounts have average dollar profits and returns that are positive before the reform but turn negative in the months after the reform. This adverse change is substantial and statistically significant. Existing institutions' profitability is positive both pre- and post-reform on average, although it is also negatively affected after the reform. This may be the result of worse conditions in the overall market after the reform. Several papers show that splits in index futures contracts increase adjusted trading volumes and hedging efficiency, while raising the contract size of futures results in lower liquidity (Norden (2006); Bjursell et al. (2010); Karagozoglu, Martell (1999)).

On the other hand, an analysis of the KOSPI200 options market reform by Yang, Ryu and Ryu (2018) suggests that market efficiency, measured by price disagreements and adjustments between the index options and futures markets, has increased after the reform. In our tests, we control for overall market movements and trends. Even though we find that both investor classes are negatively affected by the market reform, institutions do not react irrationally to it, and are better-off compared to retail investors.

Finally, we show that the negative consequences of the reform persist for several months after its introduction. Retail investors' inattention causes their trading volumes to rise in the two months post-reform. After that, they start to learn about the new contract size and gradually adjust their trades. However, it takes six months to fully reverse the initial increase in risk taking activity to pre-reform levels, and there is no significant decrease afterward.

Overall, we find important evidence that retail investors may respond irrationally to regulatory reforms aimed to protect them, which can result in those reforms having the opposite effect. Regulators should take into account behavioral biases when designing policies.

We contribute to two strands of the finance literature. Our main contribution is to show that the behavioral biases of retail investors may render regulatory reforms ineffective. In general, the literature often assumes that retail investors are unsophisticated noise traders whose irrational behavior leads them to commit systematic mistakes. In a series of papers using the trading records of a sample of retail investors, Barber and Odean provide evidence of biased retail behavior and underperformance in the equity market. Odean (1998) and Barber and Odean (2003) report evidence of a disposition effect among individual investors, defined as a tendency to realize capital gains more often than losses. Barber and Odean (2000) show that overconfident investors trade excessively and incur high trading costs which are not justified by superior returns. Several studies from the Taiwanese stock and futures markets also show that individual traders tend to incur losses (e.g., Barber et al. (2009), Kuo, Lin, Zhao (2015, 2018)). Kuo, Lin and Zhao (2015) attribute the lower investment performance to cognitive limitation. Furthermore, Barber, Odean, and Strahilevitz (2011) show that individual traders are reluctant to repurchase stocks previously sold for a loss and stocks whose price has increased since being sold. They interpret this as the effect of negative emotional reactions to past experiences with trading, such as disappointment and regret. Barber and Odean (2008) show that individual investors tend to buy attention-grabbing stocks, since they are unable to search across the full universe of stocks. This is consistent with limited attention and cognitive processing (see Hirshleifer (2015)). On the other hand, several papers argue that retail equity traders may be informed, since their aggregate trading can predict the cross-section of future stock returns, at least in the short run (e.g. Kaniel, Saar and Titman (2008); Kelley and Tetlock (2013); Boehmer et al. (2019)). Kaniel et al. (2012) show informed retail trading around firm earnings announcements. Boehmer et al. (2019) also show that retail investors possess firm-level information, but aggregate retail order flow does not predict future movements of the overall equity market. When it comes to trading in the options market, Bauer, Cosemans and Eichholtz (2009) show that most retail investors incur substantial losses, attributable to overreaction to past stock market returns, as well as gambling and entertainment as motivations for trading options. Nevertheless, they uncover a small subgroup of retail option traders who outperform the rest. All in all, the literature offers opposing views on the performance of retail investors. Our study shows that, on average, retail investors are less sophisticated and more prone to behavioral biases than institutional investors. They are subject to self-attribution bias and limited attention, which negatively affect their profitability. More importantly, we show that when regulators do not take into account the irrational behaviors of retail investors, even reforms that try to protect them may end up hurting them instead.

We also contribute to the literature on the consequences of financial crises. There are only a few studies of the effects of a market crisis on retail investor behavior, trading activities and profitability. Several studies show that, during the 2008 global financial crisis, retail investors did not withdraw their equity investments from the market (e.g. Kallberg, Liu, Wang (2014); Dorn and Weber (2013)). They continued to trade actively and used the depressed prices as a chance to enter the stock market (Hoffmann et al. (2013)). Barrot, Kaniel and Sraer (2016) argue that retail stock traders provide liquidity in times of market stress, and their aggregate stock exposure even increased in the crisis period. On the contrary, other studies show a panic selling effect in times of market stress (e.g. Hoopes et al. (2016); Burch et al. (2016)). Guiso, Sapienza and Zingales (2018) show that risk aversion increases after the 2008 crisis, and fear leads investors to divest more stock. Malmendier and Nagel (2011) suggest that individuals who have experienced low stock market returns are less willing to take risk in the future and to participate in the stock market at all. In contrast, Dorn and Weber (2013) argue that investors who remain in the German market after the crisis take on more idiosyncratic risk and have worse portfolio diversification. Consistent with some of the opposing views presented in the literature, we observe that some investors fully exit the Korean options market at the time of the August 2011 crash, while others remain and try to ride out the storm. Few studies observe the post-crisis period and it is unclear how a crisis impacts investors in the long run. Hoffmann et al. (2013) show that, although during the worst months of the crisis Dutch individual investors experienced large negative returns which caused their return expectations and risk tolerance to decrease, this effect was temporary, and their perceptions recovered soon afterwards. Similarly, our results show that the large losses that investors experience in the crisis are concentrated in the month of the market crash, but the investors who remain in the market, as opposed to exiting, recover quickly after that. However, we show that self-attribution and extrapolation biases are amplified during the crisis. This is consistent with

Kallberg, Liu and Wang (2014) who show evidence that investors' disposition effect becomes stronger during a crisis.

The rest of the paper proceeds as follows. Section 2 describes the data and institutional background, including details about the contract size reform. In Section 3, we develop our main hypotheses. Section 4 presents the empirical results. Finally, Section 5 concludes the paper.

2. Data and institutional background

We use data from South Korea's main index options market. The underlying security of the option contracts is the Korea Composite Stock Price Index, KOSPI 200, which comprises the 200 largest companies listed on the Korea Exchange (KRX). It represents Korea's overall stock market, similarly to the S&P 500 index in the US. The Korean derivatives markets are some of the world's most actively traded and liquid markets. They attract both domestic and foreign investors, generating trading volumes comparable to those in the US markets.

Our data contains the options trades of all accounts from 1 January 2010 to 30 June 2014. During our sample period, the total trading volume in KOSPI 200 options is approximately 8.6 billion contracts, and the total dollar volume is KRW 1,323,552 billion (equal to approximately USD 1,143 billion). The sample time period allows us to study the long-term effects of the financial turmoil surrounding the European sovereign debt crisis, as well as the subsequent market reform that the Korean regulators implemented. In August 2011, the long-lasting eurozone crisis, combined with a credit rating downgrade by Standard and Poor's of the United States sovereign debt, caused fears of contagion in the global economy, and markets around the world, including the KOSPI markets, experienced a sharp drop in value. Shortly after the market crash, regulators decided to increase the contract size (multiplier) for KOSPI 200 options. In January 2012, KRX announced that the options contract size will be raised five times in magnitude, from KRW 100,000 to KRW 500,000, for contracts that mature after June 2012.[1] The reform was implemented partly in order to match the futures contracts multiplier, which was already equal to KRW 500,000. However, it was also motivated by a desire to limit retail participation and speculation in the market. Up until then, the KOSPI options market attracted a significantly large number of retail investors. About 96% of the accounts in our data are retail accounts. By increasing the capital required for trading options fivefold, the regulators prevented the smallest retail accounts from participating in the market. Whether this reform had a beneficial effect in terms of protecting retail investors from excessive losses is an open question. Our study provides evidence on the long-term effects of the financial crisis and the following reform on the trading behavior and profitability of institutional and retail investors.

The data consists of all transactions executed in the KOSPI 200 options market, with a millisecond time stamp and detailed information about both the buyer and the seller. This includes encrypted account IDs and investor type codes, which we use to separate accounts into institutions[2] and retail investors. Table 1 report summary statistics for the pooled sample of account-month observations. In total, there are 6,287 institutional and 153,842 retail investor accounts that trade options at least once in the data. First, we report statistics for the trading volume (number of option contracts traded) of each class of investors. Since our transactions data starts from 2010, we exclude any contracts that start being traded in 2009. This means that

[1] We take June 2012 as the effective date of the reform. In reality, the reform was implemented in stages, such that each option contract with maturity after June 2012 adopts the new multiplier at the time when it is listed on the market and begins trading. The first contract that is listed with the new multiplier is the contract that matures in September 2012. It begins trading in March 2012. However, options that are far from maturity tend to attract little trading volume, and usually become liquid in the month preceding their maturity month. For this reason, we assume that the reform takes effect in June 2012 (the month preceding the maturity month of the July 2012 contract). In short, June 2012 is the first month when liquid option contracts are traded with the new multiplier.

[2] Institutional accounts include financial investment companies, banks, pension funds, insurance firms, trusts, government institutions, and other institutions.

we observe only a portion of the total trading volumes at the beginning of our sample, and for this reason we exclude the first two months of 2010. In order to be able to compare trading volumes before and after the market reform, we multiply post-reform trading volumes by a factor of five. Unsurprisingly, institutions generate significantly higher trading volumes on average compared to retail investors.

Table 1 also reports statistics for the profitability of the two classes of investors. We start by calculating each account's monthly dollar profit and loss ($P&L). In other words, we measure the amount in Korean won (KRW) generated by the trading activity of each account in a given month. We calculate position sizes using the contract size (multiplier). Any positions that are not closed before the end of the month are marked to market based on the last best bid and offer (BBO) quotes midpoint on the last trading day of the month. If a position is held until maturity, $P&L is calculated using the final settlement price based on the underlying index value at expiration. For institutions, an average account-month generates almost 40 million KRW, while an average retail account-month loses more than 1 million KRW. This result is largely consistent with the literature on underperformance of retail investors (see e.g. Barber, Odean (2000); Barber et al. (2009); Kuo, Lin, Zhao (2015, 2018); Bauer, Cosemans, Eichholtz (2009); and others).

However, profitability in dollar terms is related to account size and possible capital constraints. To control for such effects, we also report summary statistics for returns, although our measure of return differs from the conventional definition. We use the following method to calculate profitability per dollar invested. For each account and each day, we measure dollar profits as described above, but we separate $P&L from intraday trades and $P&L from positions held at the end of the day. Intraday trades are any round-trip transactions that buy and sell the same contracts on the same day. The rest of the trades that open a new position which is held at the end of the day or close an existing position from the previous day are marked as transactions which create end-of-day positions. If a single transaction does both at the same time, we split it into two parts. In order to match opening transactions with their corresponding closing transactions, we assume that inventory is handled following the "first-in, first-out" rule, where assets acquired first are sold first. End-of-day positions are marked to market on non-expiration dates, while maturing contracts are settled based on the value of the underlying index. After having calculated daily $P&Ls, we proceed to calculate daily capital requirements. We use the margin requirements of the Korea Exchange (KRX), where investors must pay the full option contract prices to open long positions and deposit margins to open short positions. Since margin settlement is daily, we calculate end-of-day capital requirement based on the value of end-of-day position holdings. We also calculate intraday capital requirements based on any intraday transactions. Then, the total capital requirement for a given day is equal to the sum of capital requirement at the end of the previous day and intraday capital requirement. We take the maximum of the daily capital requirements in a month, as it reflects the amount of capital the investor needed to have in his account on that month in order to execute all his trades. Finally, we calculate an account's monthly return as the sum of his daily $P&Ls from positions and $P&Ls from intraday trades, divided by the maximum monthly capital requirement. The formula below summarizes our method for calculating monthly returns:

$$\text{Monthly RET} = \text{Total monthly \$P\&L} \,/\, \text{Maximum monthly capital requirement}$$

$$= \frac{\text{sum (\$P\&L from end-of-day positions} + \text{\$P\&L from intraday trades)}}{\text{max (lag of end-of-day capital requirement} + \text{intraday capital requirement)}}$$

In essence, the $P&L scaled by capital requirement is a rate of return to each dollar of capital invested in KOSPI options. We calculate returns in this way in order to take into account the unique feature of margins in derivatives trading. The capital requirements in options trading are different from those in equity trading. Therefore, compared to the conventional method for calculating return simply based on price changes, ours is a more practical and implementable measure of profitability in the options market. Table 1 shows that institutions have mean (median) monthly return of 38% (8%), while retail investors have mean (median) return of –6% (–11%).[3]

[Table 1 about here]

Throughout the paper, we report investors' trading activity in the options market only. Option positions that appear to provide directional exposure to the underlying market may in fact be used to hedge any exposure from index futures, ETFs, or stock portfolios. For example, well-known hedging strategies that combine options and futures include covered calls and protective puts. In some cases, we may observe a large profit or loss in the options market, but the actual magnitude would be lower once we factor in the $P&L from the futures holdings. However, we find that such hedging strategies are rarely used in the KOSPI 200 derivatives markets. Only around one third of the accounts in our data trade futures at all. Out of all the end-of-day positions held in the sample, only around 6% are combinations of options and futures, and less than 1% are covered calls or protective puts. Index options may also be used to hedge equity portfolios, but it is unlikely to be the case in our data. They are not likely to be used for hedging ETFs because trading volumes in Korean ETFs are very low during our sample time period. When it comes to hedging portfolios of stocks, it is more likely to be done using individual stock options rather than index options. We are not able to verify this, as we

[3] Some of these numbers may seem inflated, but this is a result of the method we choose for calculating returns. We need to keep in mind that these returns are not the same as conventional returns calculated from simple price changes. First of all, options trading often consists of holding highly leveraged positions. The margin rate for short positions ranges from 10.5% to 15%, which means that any profits from writing options are scaled only by a fraction of the total contract value, resulting in large percentage returns. Naturally, high returns are associated with high risks, so we need to keep in mind that such risky option portfolios are likely to constitute only a small part of an investor's overall wealth allocation. Second, we assume that monthly capital requirement is equal to the maximum (rather than the sum) of daily capital requirements which may result in some inflated values. Even though our return calculations are imperfect approximations, they are useful for comparing profitability cross-sectionally and in the time series.

do not have data on investors' equity portfolios. Nevertheless, our assumption is consistent with Lakonishok et al. (2007) and Bauer, Cosemans, and Eichholtz (2009). They find that only a small portion of the option market activity of retail accounts is driven by hedging. As for institutions, we find in our data that very few of those that are likely to hedge (pension funds, insurance companies, government institutions) have one-directional exposure to begin with.

3. Hypotheses

We hypothesize that retail investors are subject to biases which cause their behavior to deviate from perfect rationality. These behavioral biases help to explain the negative effects of the crisis and the following regulatory reform on retail profitability.

Our first hypothesis is about retail investors' behavior in the crisis. Based on previous literature, we would expect some accounts to engage in panic selling and potentially exit the market during the month of the market crash (see e.g. Hoopes et al. (2016); Burch et al. (2016); Guiso, Sapienza, Zingales (2018); Malmendier, Nagel (2011)), while other accounts are likely to remain in the market and trade actively (e.g. Kallberg, Liu, Wang (2014); Dorn, Weber (2013); Hoffmann et al. (2013); Barrot, Kaniel, Sraer (2016)). To go one step further, we hypothesize that retail investors are subject to self-attribution bias (see e.g. Gervais, Odean (2001); Hoffmann, Post (2014)), which may be amplified during times of market stress, consistent with Kallberg, Liu and Wang (2014). Self-attribution causes them to believe that their past success is due to skill rather than luck and to extrapolate their good performance into the future. For this reason, we expect that retail accounts who are profitable before the crisis are more likely to remain in the market and trade more during and after the crisis, but not necessarily gaining from it.

Our second hypothesis is about retail investors' reaction to the reform. In economics, a rational agent is assumed to be an individual who always chooses to perform actions which will result in the optimal expected outcome. Therefore, when the contract size rises from KRW100,000 to KRW500,000, we would expect a rational investor to decrease his trading volume fivefold. For example, if an account used to always trade 100 option contracts before the reform, the optimal number of contracts for him to trade after the reform is 20. To account for the change in contract size and to make the pre- and post-reform periods in our data comparable, we multiply post-reform trading volumes by a factor of five. Hence, if investors react rationally to the reform, we should not see any significant change in volumes after adjusting for the new contract size. However, if some investors are not paying attention to regulatory announcements or if they do not fully understand the market changes, we would expect to see increased trading volumes after the reform is introduced. Following the literature on limited attention and cognitive processing (e.g. Hirshleifer (2015); Barber, Odean (2008); Kuo, Lin, Zhao (2015)), we hypothesize that retail investors are subject to these biases. We predict that they will react irrationally to the reform by failing to adjust their trading behavior for the new contract size. As a result, their post-reform trading volumes will rise unintentionally, and their profitability may suffer. Over time, it is likely that all market participants learn about the reform, so we would expect the negative effects of the reform to dissipate after a while.

Finally, we hypothesize that, contrary to retail investors, institutions are not subject to the same behavioral biases. As more sophisticated investors, they are expected to behave rationally, to pay attention to ongoing market changes and adjust to the reform quickly.

4. Results

We start by presenting some time-series trends in our data. Firstly, we are interested in examining market entries and exits. We define market entry and exit based on an account's activity during the past six months and the following six months. For each month t, we check whether the account has any market activity (trades or positions) during the windows [t-6, t-1] and [t+1, t+6]. If he was inactive in the past six month and is active in month t, we mark t as the month of market entry. Similarly, if he is active in month t and inactive in the next six months, we mark t as the month of market exit, when the account closes all of his positions. According to these definitions, an account may enter and exit the market several times over the sample period. However, if for example he trades only once every three months, we do not consider the inactive months in between his trades as an exit from the market. Instead, we consider the account as staying in the market but being inactive.

Figure 1 plots the percentage of all accounts that enter and the percentage of all accounts that exit the market each month. These are plotted against the underlying KOSPI 200 index return. The two highlighted areas in the graph correspond to the month of the market crisis (August 2011) and the effective month of the contract size reform (June 2012). The figure shows the crash in the value of the KOSPI 200 index in August 2011. This is accompanied by a sharp rise in the percentage of accounts that exit the market, consistent with a panic selling effect. The percentage of exiting accounts goes down after that and spikes up again around the time of the reform. At that time, the increase in exits is much smaller compared to the one during the crisis, and is more likely to be caused by the exiting of small accounts that are unable to meet the new higher capital requirements. The percentage of accounts that enter the market seems to be more stable over time, although we do notice a general decrease post-crisis.

[Figure 1 about here]

Next, we examine the time-series trends in trading volumes of institutions and retail investors. To account for the change in contract size and to make the pre- and post-reform periods comparable, we multiply post-reform trading volumes by a factor of five. Panel A of Figure 2 plots the monthly aggregate trading volume of all accounts in each investor class category. As expected, institutions always generate larger aggregate volumes compared to retail investors, even though the number of retail accounts trading in the market is significantly higher. The trading volumes of institutions and retail investors seem to move together in general. They start to decline before the crisis and go back up around the time of the reform, before eventually stabilizing in the last part of our sample period. Panel B plots the monthly cross-sectional average of institutional and retail accounts' trading volumes. A noticeable pattern once again is the increase in trading volumes around the time of the reform. When the contract size changes from KRW100,000 to KRW500,000, we would expect a rational investor to decrease his trading volume fivefold. For example, if he used to always trade 100 contracts before the reform, the optimal number of contracts for him to trade after the reform is 20. Since in our data we multiply post-reform trading volumes by a factor of five, we should not see any significant change in volumes at the time of the reform. Therefore, it is puzzling that volumes rise so much, and we attribute this increase to irrational trading behavior.

[Figure 2 about here]

We also show the changes in profitability of institutions and retail investors over time. Panel A of Figure 3 plots the aggregate dollar profit and loss ($P&L) of the two investor classes each month. It is clear that, in aggregate, institutions tend to profit from trading against the less sophisticated retail accounts. Whenever one category gains, the other category loses, reflecting the fact that options trading is a zero-sum game. The largest $P&Ls were realized in the month of the market crash. The aggregate losses of retail investors in that month amounted to almost 1,000 billion KRW, while institutions gained more than 800 billion KRW in total. Furthermore, Panel B plots the monthly cross-sectional average returns (calculated as $P&L scaled by capital requirements). The returns tend to be much more volatile. But once again, we see that the largest drop for retail investors occurred during the market crash.

[Figure 3 about here]

Panel B of Figure 3 shows that in the month of the market crash institutions also experienced a drop in returns on average. If their aggregate dollar profits were positive, but their average return is negative, there must be a few institutions that realized large gains at the expense of other market participants. To further understand who gained and who lost in the market crisis, we report summary statistics of $P&L and returns in the month of August 2011. Table 2 shows that, in the month of the crash, the number of institutions that realized positive dollar profits was only slightly higher than the number of institutions that lost money. The mean $P&L of institutions is approximately 710 million KRW in that month, but that seems to be driven by some outlier accounts with huge gains. The profit of the accounts in the top 80th percentile is nearly four times larger than the loss of the accounts in the bottom 20th percentile. However, when we scale $P&L by capital requirements, we get a mean return that is negative even for the institutional accounts. Still, their mean return is substantially higher than that of retail investors. An average institution has a return of –6% in the crisis, compared to –69% for an average retail investor. The average $P&L of retail investors in that month is –21.7 million KRW. There were about 2.5 times more retail accounts that realized losses compared to those that gained in the crisis, and even at the 60th percentile, retail accounts still have negative dollar profits and returns. At the extreme end of the spectrum, the largest loss is –84 billion KRW.

[Table 2 about here]

Next, we proceed to analyse the four-month windows around the crisis and around the reform, in order to compare investors’ trading activity before and after the events. First, we separate accounts into three groups: accounts that exit, existing accounts, and new accounts.

Accounts that exit are accounts that are in the market during the four months before the crisis (April – July 2011), but not in the four months after the crisis (September – December 2011). New accounts are accounts that are not in the market in the four months before the crisis, but appear and stay in the market during the four months after the crisis. The group which we call "existing accounts" consists of those investors who exist in the market starting from four months before the crisis and remain in the market for at least one year and a half until four months after the contract size reform (Apr 2011 – Sep 2012).[4] In total, there are 396 institutions and 12,218 retail investors which are categorized as existing accounts. We create this group in order to test the effects of both the crisis and the reform on long-term investors that choose to stay in the market regardless of the turbulence and changes that take place. Table 3 reports statistics for the four-month windows before and after the crisis, for institutions and retail investors in each of the three account groups. We report the number of accounts in each category, and the cross-sectional averages of their monthly mean trading volumes, their total $P&L (in millions of KRW) over the four-month periods, and their monthly mean returns. For the existing accounts, we calculate the change in volumes and profitability from the four-month period before the crisis to the four-month period after it, and we perform paired t-tests to evaluate the statistical significance of the differences in means. Both the institutions and the retail accounts which remain in the market tend to trade less after the crisis. The retail investors that exit the market are the group with the largest losses and lowest returns before the crisis. On the other hand, those that remain have positive profitability on average, both pre- and post-crisis. As expected, there are some accounts that seek to take advantage of the depressed prices and enter the market during or shortly after the crisis. The new retail accounts perform worse than the existing retail accounts post-crisis. When we focus on the institutional investors, we

[4] Staying in the market does not necessarily entail being active every month. According to our definition, we mark an account as staying in the market as long as he does not become inactive for six or more months.

see that those who exit from the market have positive pre-crisis profitability. Perhaps, in expectation of the crisis, they chose to collect their gains and exit the market before the crash. In contrast, it appears that the institutions which stay in the market experience a significant reduction in returns. Finally, the new institutions which enter the market after the crisis perform better than the existing institutional investors.

[Table 3 about here]

In Table 4, we perform the same univariate analysis as in Table 3, but for the four-month windows around the contract size reform. The reform was announced in January 2012. A step-by-step implementation process started in March 2012 and the reform was fully implemented by June 2012. We use June 2012 as the effective start of the reform because it is the first month when liquid option contracts were traded with the new multiplier. (For details about the reform implementation schedule, refer to Section 2 - footnote 1). To have a clean comparison of pre- and post-reform market activities, we exclude the month of the announcement and the transitional period when there were listed contracts with both the old and the new multiplier. Therefore, we compare the four months before the reform was announced (Sep 2011 – Dec 2011) to the four months after it was fully implemented (Jun 2012 – Sep 2012). The group of existing accounts remains the same. Accounts that exit are those that are in the market during the four months before the reform announcement, but not in the four months after the reform was implemented. New accounts are those that are not in the market in the four months before the reform announcement, but are in the market in the four months since the start of the reform. The table shows that a large number of retail investors exit the market after the reform is introduced. This group likely contains the smallest accounts which used to trade only a few lots at a time and are unable to continue trading after capital requirements increase fivefold. They have dollar losses and negative returns on average before the reform, suggesting that the regulators were in part successful at protecting unsophisticated retail investors from continuing to lose money in the options market. On the other hand, the retail accounts that stay in the market long-term have average $P&L and returns that are positive before the reform but turn negative in the four months after the reform. This adverse change in profitability is substantial and statistically significant. It is accompanied by an increase in trading volumes post-reform, which is inconsistent with a rational reaction to the reform. Instead, it is consistent with retail inattention to the reform. Similarly, the new retail accounts that enter the market right after the reform trade more than their peers and generate losses on average. As for institutions, the accounts that exit have negative pre-reform $P&L but positive returns on average. The existing institutions that stay in the market in the long run are profitable both pre- and post-reform, and do not experience any significant change in trading volume or profitability after the contract size increases. The new institutional accounts that enter the market post-reform do not trade more than the existing institutions and have similar returns on average. These results suggest that institutions do not exhibit the same irrational behavior as retail investors after the reform. Instead, they continue trading in a similar way as they did before, adjusting their trading volumes correctly in accordance with the contract size change.

[Table 4 about here]

The univariate analyses are useful for observing changes in average trading activity around the events. However, since the post-crisis period coincides with the pre-reform period, it is hard to analyze the two events separately. They are related to one another, since the reform was likely introduced as a response to the crisis. For this reason, we proceed to analyze market dynamics over the whole sample time period in a multivariate regression setting, where we can disentangle the effects of the two events better.

We start by looking at investors' propensity to stay and be active in the options market. In Table 5, we present the results of account-month logit regressions where the dependent variables are the following. The first dummy variable called *Staying in market* is equal to one in a given month t if the account: 1) has previously entered the market by trading intraday and/or holding option positions, 2) has not exited the market yet (exit defined as becoming inactive for six or more months), and 3) does not close his positions and exit in month t. According to this definition, an account can stay in the market for a certain period of time without being active every single month. Hence, we also present regressions for account activity. The second dummy variable called *Active in market* is equal to one in a given month t if the account trades and/or holds option positions, thereby generating a P&L in month t. The regressions with dependent variable *Active in market* are meant to show how active the accounts are in the months when they choose to stay in the market (observations where *Staying in market* is equal to one). We regress the dependent variables on three dummy variable for the time periods in which we are interested: *Crisis* is equal to one in the month of the market crash (August 2011), *Post-crisis* is equal to one for all months in the sample following the market crash, and *Post-reform* is equal to one from the starting date of the reform (June 2012) until the end of the sample. In all regressions, we control for trading volume of each account in the previous month, KOSPI 200 index return in the current and previous month, the implied volatility index VKOSPI at the start of the current and previous month, and any time trend in the data. We multiply post-reform trading volumes by five, and we take the logarithm of trading volume and VKOSPI variables.

Panel A of Table 5 shows the effects of the crisis and the reform on investors' propensity to stay in the options market. We conduct the regression analyses separately in the two subsamples of institutions and retail investors. The first regression suggests that there is no significant change in institutions' propensity to stay in the market during the month of the crisis itself, but they are more likely to stay after the crisis. The effect of the reform is not significant. In the second regression, we also control for the account's compounded return from the previous and the current month, as well as interactions of the return and the event dummies. We add these controls in a separate regression because they require the account to have activity (and therefore non-missing values for returns) in both months t and t-1, which slightly reduces the size of the sample. After controlling for returns, we do not observe any strongly significant effects of the events on institutions' propensity to stay in the market. The regressions in the retail subsample show a significantly lower propensity of retail investors to stay in the market in the month of the crash, consistent with our earlier observation that a large number of accounts exit in that month. The regressions also show that retail investors are more likely to stay in the market in the post-crisis and post-reform periods. Panel B shows that, among those months when investors are staying in the market, both institutional and retail accounts tend to be more active in the month of the crisis. Retail investors that remain in the market also tend to be more active after the crisis. In contrast, after the reform, both institutions and retail investors have a lower propensity to be active. These results do not change after we control for previous-month returns and their interactions with the event dummies.

Running the regressions separately in the subsamples of institutions and retail investors does not allow us to compare the magnitude of the effects on the two investor classes. To be able to compare institutions' and retail investors' propensity to stay and be active in the market, we run similar regressions in the full sample of all accounts. We add a dummy variable for the institutions in the sample, and interact it with the event dummies. Since all the remaining accounts are retail investors, the event dummies without an interaction represent that group of investors. The interactions of the event dummies with the *Institution* dummy in Panels C and D show that, compared to the retail investors, institutions are more likely to stay in the market during and after the crisis, but those that stay are less active than their retail counterparts. We observe the opposite effect after the reform: institutions are less likely to remain in the market but those that stay have higher propensity to be active, compared to the retail traders.

In general, all regressions in Table 5 show that higher previous-month returns are associated with a higher likelihood to stay in the market and be active in the following month. This is consistent with a possible self-attribution bias and biased extrapolation – investors believe that their past success is due to skill rather than luck and assume that they will continue to be profitable in the future. The effect on account activity seems to be amplified during the crisis, as Panel D shows a positive and significant coefficient on the interaction between previous-month return and the *Crisis* dummy.

[Table 5 about here]

Next, we show the effects of the crisis and the reform on investors' trading volumes, dollar profitability ($P&L) and returns. In particular, we are interested in the events' effects on the group of existing accounts which we previously identified as those investors who stay in the market in the long run, starting from at least four months before the crisis until at least four months after the reform. Table 6 presents the results of a series of account-month OLS regressions, where the independent variables are the same as in Table 5. Panel A focuses on trading volumes. Both existing institutions and existing retail investors increase their trading volumes in the month of the crash. After the crisis, institutional trading volumes do not change, but retail investors tend to trade less. Once again, consistent with self-attribution bias, we see that previous-month return is significantly positively associated with trading volume in the current month, but only in the subsample of retail accounts. In addition, this effect is amplified during and after the crisis, consistent with some literature that suggests behavioral biases are amplified in times of market stress (e.g. Kallberg, Liu and Wang (2014)). This suggests that retail investors who perform well before the crisis extrapolate their good performance into the future, which causes them to trade more and have worse performance during the crisis. There is no such effect in the subsample of institutions.

As in previous analyses, we multiply post-reform volumes by five to adjust for the fivefold increase in options contract size. Therefore, if investors rationally adjust their trading volumes after the reform, we should not expect to see any significant coefficient on the post-reform dummy. To adjust for the fact that investors may trade less often but may trade more lots at a time, we set trading volume to be equal to zero in the months when they stay in the market but are inactive. As expected, there is no significant effect on institutional trading volumes post-reform. However, we do observe a significant rise in retail trading volumes after the reform, which is consistent with inattention to the contract size increase. The regressions in Panel D confirm these results: when comparing the two investor classes directly, they have opposite signs on the post-reform dummy.

Panels B and C of Table 6 show the effects of the crisis and reform on existing investors' dollar profits and returns. In terms of $P&L, retail investors experience significant losses in the month of the crash, while institutions are not affected. In terms of returns, both institutions and retail investors are negatively affected in the crisis. The post-crisis effects are not significant after we add all the control variables, suggesting that the negative effects of the crisis are concentrated in the month of the market crash, but the investors who remain in the market, as opposed to exiting, recover quickly afterwards. In the post-reform period, both institutional and retail profitability is negatively affected. This is possibly the result of worse liquidity and hedging efficiency after the reform (see e.g. Norden (2006); Bjursell et al. (2010); Karagozoglu, Martell (1999)). Panels E and F help us to understand which investor class is more affected by the events in terms of profitability. The coefficients on the *Institution* dummy are positive and significant, reflecting the fact that institutional investors outperform retail investors in general. The interaction of the *Institution* and *Crisis* dummies is also positive and significant, verifying our earlier observation that institutions do better in the month of the crisis compared to retail investors. In addition, Panel F shows a negative association between previous-month return and return in the month of the crash for existing retail accounts, confirming that their self-attribution bias leads to losses during the crisis. We observe the opposite for institutional investors. However, there is some evidence in Panel F suggesting that institutions are worse-off after the crisis than their retail counterparts. On the contrary, after the reform is introduced, even though both investor classes are negatively affected, institutions are better-off compared to retail investors. The fact that retail investors' inattention to the reform had negative effects on their profitability further proves the failure of the regulator to protect these investors.

[Table 6 about here]

Despite the negative consequences of the reform, it is unlikely that it affects investors in the long run. We would expect that, after a certain period of time, all market participants would learn about the new contract size and adjust their trades accordingly. In Table 7, we test how persistent the effect of the reform is on the trading volumes of existing retail investors. We repeat the same analysis as the first regression in Panel A of Table 6, but we add twelve dummy variables for the twelve months after the start of the reform. By calculating the cumulative sum of the coefficient on the post-reform dummy and the coefficients on the twelve monthly post-reform dummies, we can understand the effect of the reform over time. The coefficients on the post-reform dummy itself, as well as the coefficients on the dummies for the first and second months post-reform are positive and significant. This means that retail investors' inattention caused their trading volumes to rise in the two months after the contract size increase. Starting from the third month post-reform, the monthly dummies have negative coefficients, suggesting that retail investors started to adjust their trades afterwards. However, calculating the cumulative sum of the coefficients shows that it took six months to fully reverse the initial increase in trading volumes.

[Table 7 about here]

5. Conclusion

In this study, we show that market reforms catered towards retail investors can harm rather than help them, if regulators do not consider the possible irrational reaction of these investors to the reforms. We use account-level trading data from the Korean index options market to examine the impacts of the August 2011 market crisis and the subsequent market reform. After retail investors experienced large losses in the crisis, the regulators decided to increase the contract size for KOSPI 200 options five times in magnitude, which was motivated by a desire to limit retail participation and speculation in the market. We investigate how this reform affected retail and institutional investors in view of their different behavioral patterns.

First, we show that a large number of retail investors exit the market in the month of the market crash, but many others stay and trade actively during the crisis. Retail accounts who perform well before the crisis are subject to self-attribution bias, making them more likely to remain in the market, trade more, and experience large losses in the crisis. Institutions do not exhibit the same behavior. The post-crisis effects do not appear to be significant, suggesting that the negative effects of the crisis are concentrated in the month of the market crash, but the investors who remain in the market, as opposed to exiting, recover quickly afterwards.

Second, we show that some retail investors exit the market when the reform is introduced, likely due to an inability to continue trading after capital requirements increase fivefold. The retail accounts that exit have negative profitability on average before the reform. However, in general, investors' propensity to exit the market decreases after the reform. Moreover, we show that the reform has a negative impact on the remaining retail investors that stay in the market long-term, because the regulators do not anticipate their irrational reaction. In our analyses, we multiply post-reform trading volumes by a factor of five to adjust for the fivefold increase in options contract size. Therefore, if investors rationally adjust their trading volumes after the reform, we should not observe any significant changes in their trading behavior. Indeed, we find no significant effect of the reform on the trading behavior of institutions. On the other hand, the retail accounts that stay in the market display higher risk-taking activity after the reform, which is inconsistent with a rational reaction to the contract size increase. Instead, it points to limited attention and cognitive processing of retail investors. Failing to learn about the reform and to adjust their trading behavior for the new contract size, they experience negative profitability in the months after the reform, while institutions are better-off. Finally, we show that the negative consequences of the reform persist for several months after its introduction. Eventually, all market participants learn about the new contract size and gradually adjust their trades.

Overall, we find important evidence that retail investors may respond irrationally to regulatory reforms aimed to protect them, which can result in those reforms having the opposite effect. Regulators should take into account behavioral biases when designing market reforms.

References

Barber, B. M., Lee, Y., Liu, Y., Odean, T. (2009). Just How Much Do Individual Investors Lose by Trading? The Review of Financial Studies 22(2), p.609-632.

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

Barber, B. M., Odean, T. (2003). Are individual investors tax savvy? Evidence from retail and discount brokerage accounts. Journal of Public Economics 88, p.419-442.

Barber, B. M., Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. The Review of Financial Studies 21(2), p.785-818.

Barber, B. M., Odean, T., Strahilevitz, M. A. (2011). Journal of Marketing Research 48, p.S102-S120.

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

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

Bjursell, J., Frino, A., Tse, Y., Wang, G. H.K. (2010). Volatility and trading activity following changes in the size of futures contracts. Journal of Empirical Finance 17, p.967-980.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X. (2019). Tracking Retail Investor Activity. Working Paper.

Burch, T. R., Emery, D. R., Fuerst, M. E. (2016). Who Moves Markets in a Sudden Marketwide Crisis? Evidence from 9/11. Journal of Financial and Quantitative Analysis 51(2), p.463-487.

Dorn, D., Weber, M. (2013). Individual Investors' Trading in Times of Crisis: Going It Alone or Giving Up? Working Paper.

Gervais, S., Odean, T. (2001). Learning to Be Overconfident. The Review of Financial Studies 14(1), p.1-27.

Guiso, L., Sapienza, P., Zingales, L. (2018). Time varying risk aversion. Journal of Financial Economics 128, p.403-421.

Hirshleifer, D. (2015). Behavioral Finance. Annual Review of Financial Economics 7, p.133-159.

Hoffmann, A. O.I., Post, T. (2014). Self-attribution bias in consumer financial decision-making: How investment returns affect individuals' belief in skill. Journal of Behavioral and Experimental Economics 52, p.23-28.

Hoffmann, A. O.I., Post, T., Pennings, J. M.E. (2013). Individual investor perceptions and behavior during the financial crisis. Journal of Banking & Finance 37, p.60-74.

Hoopes, J. L., Langetieg, P., Nagel, S., Reck, D. H., Slemrod, J. B., Stuart, B. (2016). Who Sold During the Crash of 2008-9? Evidence from Tax-Return Data on Daily Sales of Stock. Working Paper.

Kallberg, J., Liu, C. H., Wang, N. (2014). Individual Investors and the Financial Crisis. Working Paper.

Kaniel, R., Liu, S., Saar, G., Titman, S. (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. The Journal of Finance 67(2), p.639-680.

Kaniel, R., Saar, G., Titman, S. (2008). Individual Investor Trading and Stock Returns. The Journal of Finance 63(1), p.273-310.

Karagozoglu, A. K., Martell, T. F. (1999). Changing the Size of a Futures Contract: Liquidity and Microstructure Effects. The Financial Review 34, p.75-94.

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

Kuo, W., Lin, T., Zhao, J. (2015). Cognitive Limitation and Investment Performance: Evidence from Limit Order Clustering. The Review of Financial Studies 28(3), p. 838-875.

Kuo, W., Lin, T., Zhao, J. (2018). The Correlated Trading and Investment Performance of Individual Investors. Working Paper.

Lakonishok, J., Lee, I., Pearson, N. D., Poteshman, A. M. (2007). Option Market Activity. The Review of Financial Studies 20(3), p.813-857.

Malmendier, U., Nagel, S. (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? The Quarterly Journal of Economics 126, p.373-416.

Norden, L. (2006). Does an Index Futures Split Enhance Trading Activity and Hedging Effectiveness of the Futures Contract? The Journal of Futures Markets 26(12), p.1169-1194.

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

Yang, H., Ryu, D., Ryu, D. (2018). Market Reform and Efficiency: The Case of KOSPI200 Options. Emerging Markets Finance and Trade 54, p.2687-2697.

Figure 1
Market entries and exits

For each month in the sample, Figure 1 plots the percentage of all accounts that enter and the percentage of all accounts that exit the market (on the primary Y axis). These are plotted against the underlying KOSPI 200 index return (on the secondary Y axis). Our sample covers the period from January 2010 to June 2014, but we exclude the first and the last 6 months of the sample, because our definitions of market entry and exit are based on an account's activity during the past 6 months and the following 6 months. For each month t, we check whether the account has any market activity (trades or positions) during the windows [t-6, t-1] and [t+1, t+6]. If he was inactive in the past 6 month and is active in month t, we mark t as the month of market entry. If he is active in month t and inactive in the next 6 months, we mark t as the month of market exit, when the account closes all of his positions. The two highlighted areas in the graph correspond to the month of the market crisis (August 2011) and the effective month of the contract size reform (June 2012).

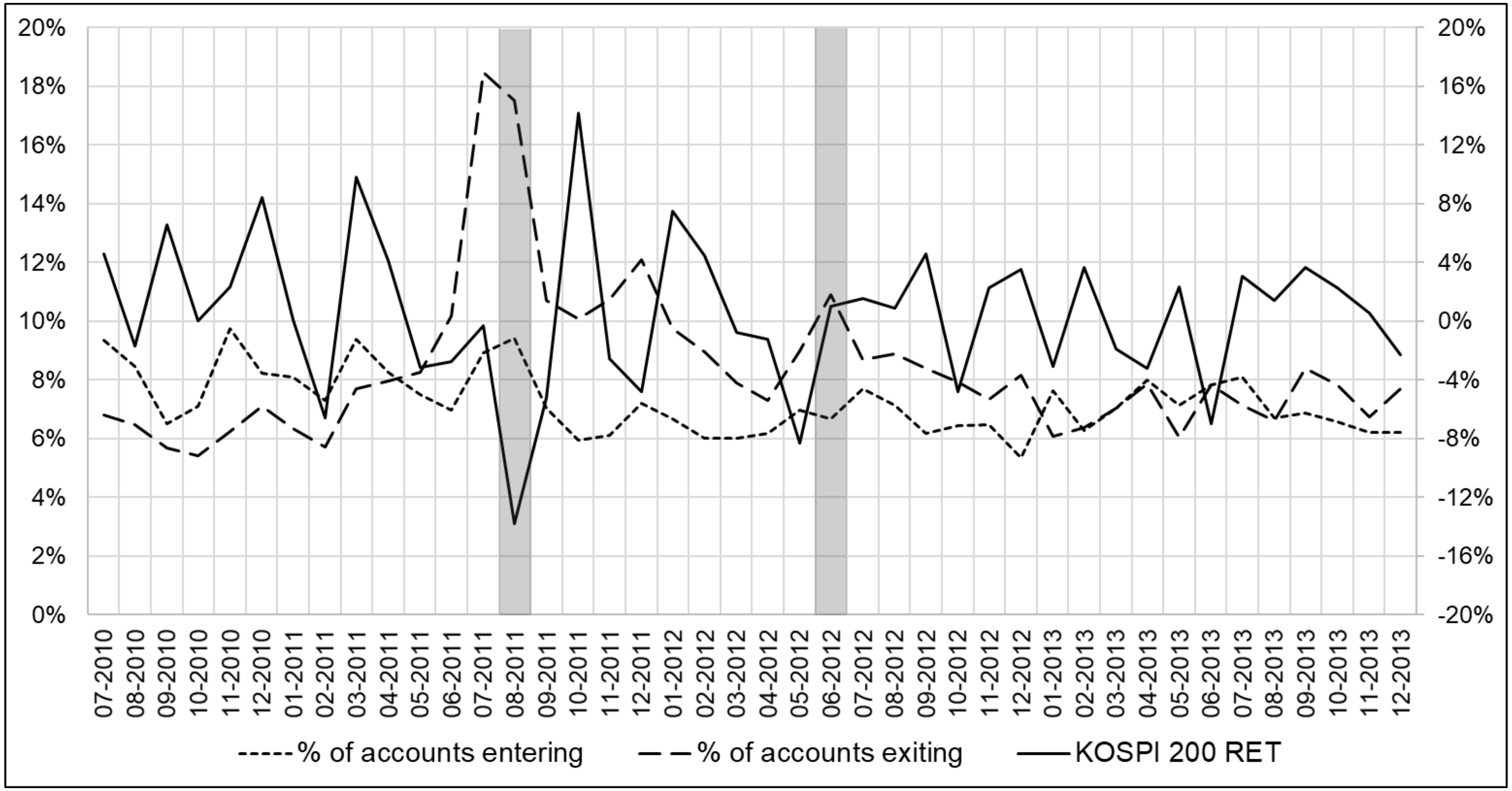
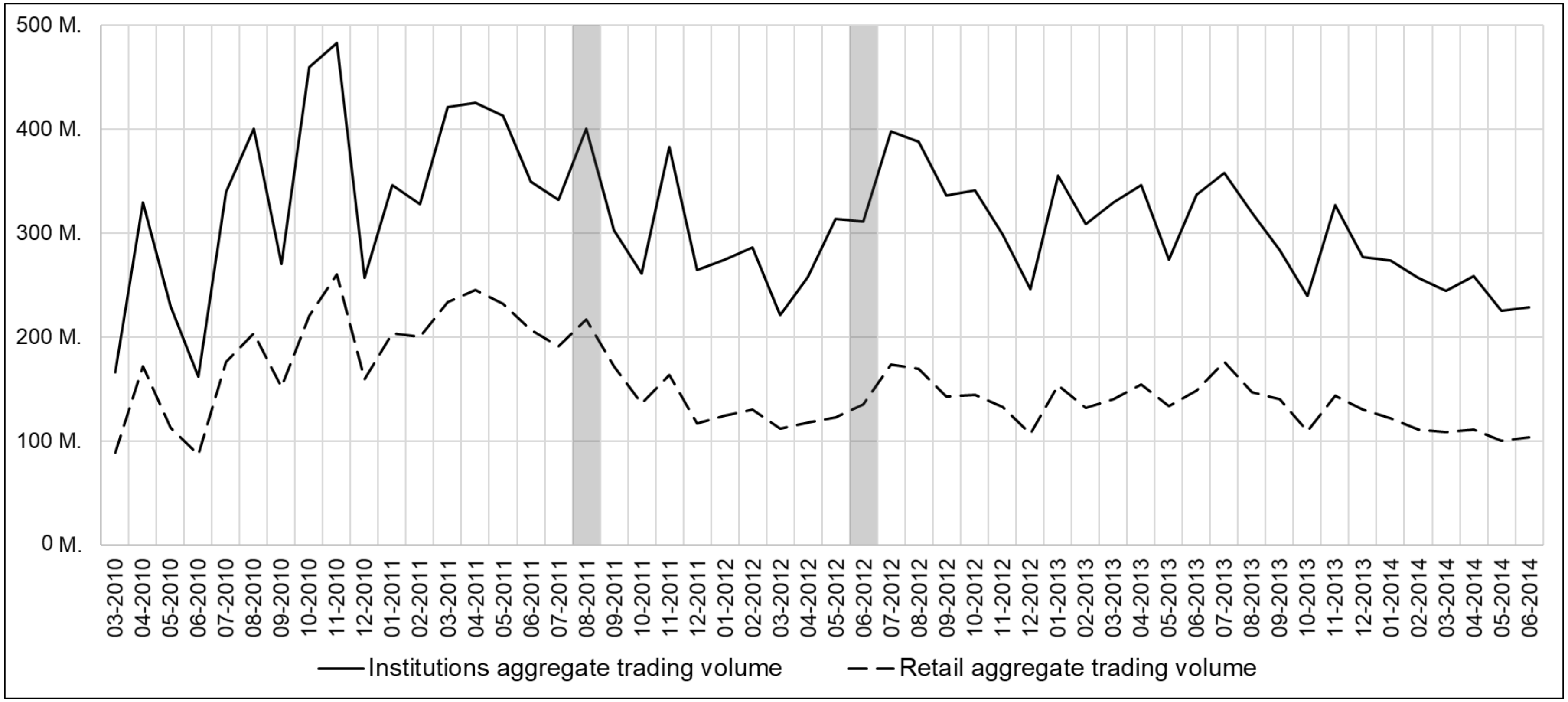


Figure 2
Trading volumes of institutions and retail investors

Figure 2 shows the time-series trends in trading volumes of institutions and retail investors. Trading volume is equal to the number of option contracts traded. The two highlighted areas in the graph correspond to the month of the market crisis (August 2011) and the effective month of the contract size reform (June 2012). To account for the change in contract size from KRW 100,000 to KRW 500,000, we multiply post-reform trading volumes by a factor of five. Panel A plots the monthly aggregate trading volume of all accounts in each investor class category. Panel B plots the monthly cross-sectional average of institutional and retail accounts' trading volumes. Our data covers the period from January 2010 to June 2014. Contracts that start being traded in 2009 are excluded. Since we observe only a portion of the total trading volumes at the beginning of our sample, we exclude the first two months.

Panel A: Aggregate trading volumes



Panel B: Cross-sectional average trading volumes

(Average trading volume of institutions is plotted on the primary Y axis, and the average trading volume of retail investors is on the secondary Y axis.)

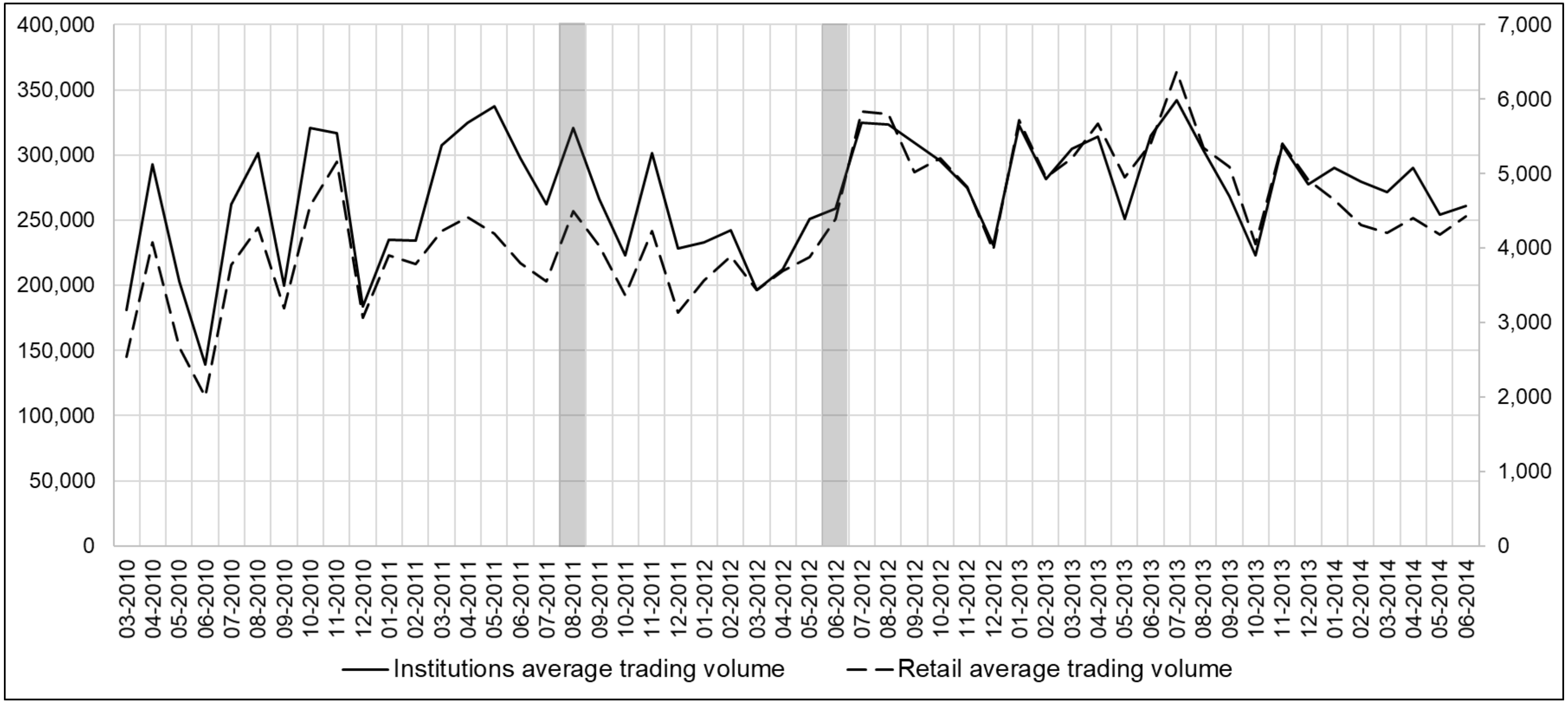
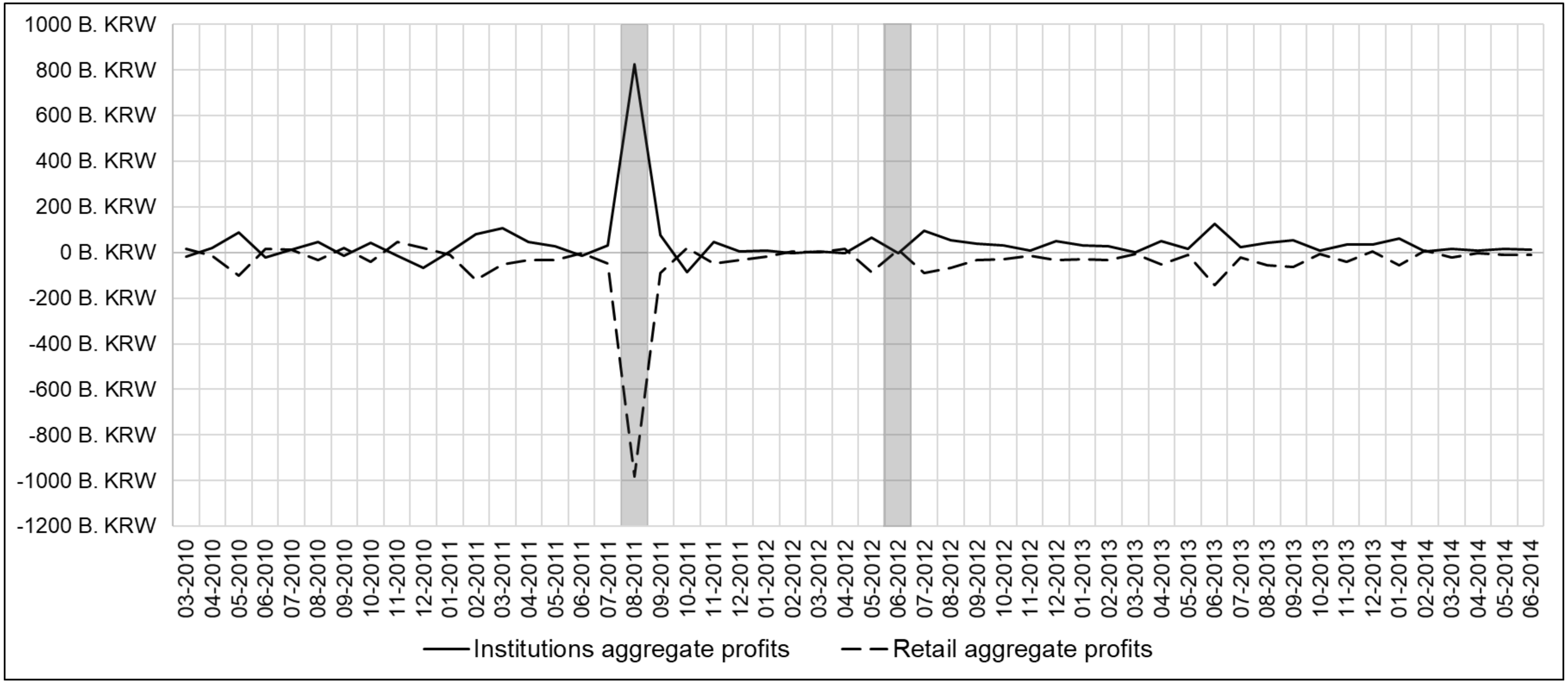


Figure 3
Profitability of institutions and retail investors

Figure 3 shows the time-series trends in profitability of institutions and retail investors. Refer to Section 2 for details about the method we use to calculate profitability. Panel A plots the aggregate dollar profit and loss ($P&L) of the two investor classes each month. Panel B plots the monthly cross-sectional average returns ($P&L scaled by capital requirements). Our data covers the period from January 2010 to June 2014. Contracts that start being traded in 2009 are excluded from the sample. Since we observe only a portion of the total trading volumes at the beginning of our sample, we exclude the first two months.

Panel A: Aggregate $P&L (billions of KRW)



Panel B: Cross-sectional average monthly returns

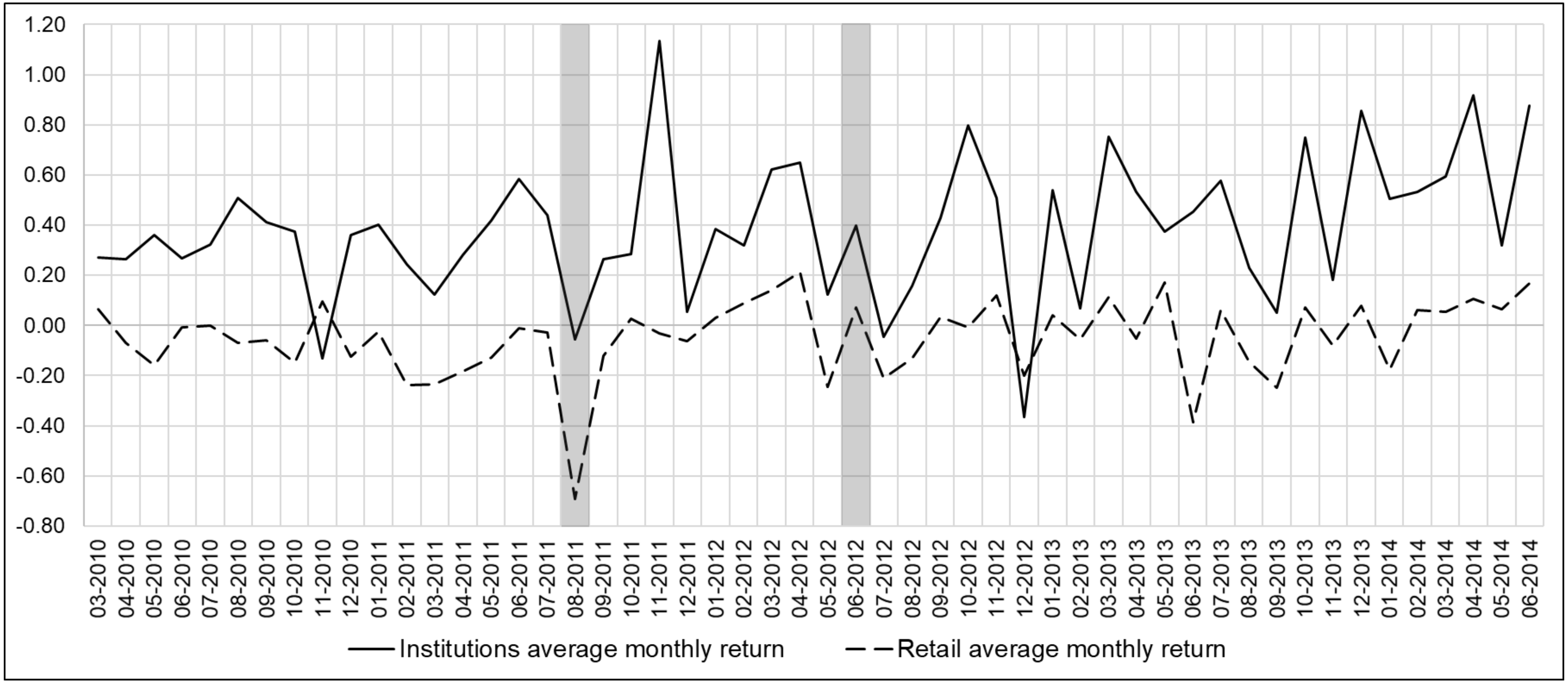


Table 1
Summary statistics

This table reports summary statistics for the pooled sample of account-month observations. At the top, we report the number of unique investor accounts that trade options at least once in the data, and the number of account-month observations for each class of investors. We use the investor type codes in our data to separate accounts into retail investors and institutions (which include financial investment companies, banks, pension funds, insurance firms, trusts, government institutions, and other institutions). All option contracts are based on the underlying KOSPI 200 index. TRDVOL refers to trading volume or number of options contracts traded. $P&L refers to dollar profit and loss (in KRW) from trading options. RET refers to monthly return, which we measure as $P&L scaled by capital requirement. Refer to Section 2 for details about the calculation method. For each variable, we report the mean, standard deviation, minimum, first quartile, median, third quartile, and maximum. Our data covers the period from 1 January 2010 to 30 June 2014. Contracts that start being traded in 2009 are excluded from the sample. Since we observe only a portion of the total trading volumes at the beginning of our sample, we exclude the first two months. The contract size for options increases from KRW 100,000 for contracts that mature in or before June 2012, to KRW 500,000 for contracts that mature after June 2012. TRDVOL post-reform is multiplied by a factor of 5, in order to be able to compare trading volumes before and after the reform.

	Institutions	**Retail investors**
N accounts	6,287	153,842
N account-months	59,596	1,879,881
TRDVOL		
mean	272,496	4,218
std	1,528,953	32,613
min	0	0
q1	205	47
median	4,760	397
q3	66,755	2,005
max	53,693,098	6,944,990
$P&L (KRW)		
mean	39,344,606	-1,380,832
std	1,695,835,613	126,473,414
min	-101,614,600,000	-84,096,739,000
q1	-3,489,500	-1,635,000
median	1,650,000	-119,500
q3	27,509,000	445,000
max	83,827,182,000	29,850,167,500
RET		
mean	0.38	-0.06
std	2.18	1.29
min	-43.93	-86.37
q1	-0.17	-0.48
median	0.08	-0.11
q3	0.52	0.14
max	208.17	120.12

Table 2
Who profited and who lost in the month of the market crash?

In this table, we report summary statistics of institutional and retail dollar profit and loss ($P&L) and returns in the month of August 2011 when the market crash occurred. We report the number of accounts that were active (traded or held positions) in that month, as well as the number of accounts that realized positive dollar profits and the number of institutions that lost money. We then report the mean, minimum, 20th, 40th, 60th, and 80th percentiles, and the maximum of $P&L (in KRW) and of returns (calculated as $P&L scaled by capital requirements). Refer to Section 2 for details about the method we use to calculate profitability.

	Institutions	**Retail investors**
N accounts	1,161	45,186
N accounts with +$	636	12,724
N accounts with –$	524	32,405
$P&L (KRW)		
mean	709,970,012	-21,712,386
min	-38,206,947,000	-84,096,739,000
pctl 20	-33,760,000	-9,316,000
pctl 40	-655,000	-1,821,000
pctl 60	8,796,000	-196,000
pctl 80	133,302,500	375,500
max	83,827,182,000	14,268,071,000
RET		
mean	-0.06	-0.69
min	-42.14	-67.52
pctl 20	-0.70	-0.97
pctl 40	-0.05	-0.53
pctl 60	0.20	-0.18
pctl 80	0.74	0.13
max	11.01	18.58

Table 3
Analysis of four-month windows around the crisis

This table compares the four-month windows before and after the market crisis. First, we separate accounts into three groups: accounts that exit, existing accounts, and new accounts. Accounts that exit are accounts that are in the market during the four months before the crisis, but not in the four months after the crisis. New accounts are accounts that are not in the market in the four months before the crisis, but enter and stay in the market during the four months after the crisis. The group which we call "existing accounts" consists of those investors who exist in the market starting from four months before the crisis and remain in the market for at least one year and a half until four months after the contract size reform (Apr 2011 – Sep 2012). We calculate statistics for each four-month window, for institutions and retail investors in each of the three account groups. We report the number of accounts in each category, and the cross-sectional averages of their monthly mean trading volumes, their total $P&L (in millions of KRW) over the four-month periods, and their monthly mean returns. Refer to Section 2 for details about the method we use to calculate profitability. For the group of existing accounts, we calculate the change in volumes and profitability from the four-month period before the crisis to the four-month period after it, and we perform paired t-tests to evaluate the statistical significance of the differences in means.

Before crisis: Apr 2011 – Jul 2011 [t-4, t-1]
After crisis: Sep 2011 – Dec 2011 [t+1, t+4]

		Number of accounts		Monthly mean TRDVOL cross-sectional average				Total $P&L (millions of KRW) cross-sectional average				Monthly mean RET cross-sectional average			
		before	after	before	after	change	t-stat	before	after	change	t-stat	before	after	change	t-stat
Accounts that exit	Institutions	136		250,142				211.5				0.33			
	Retail investors	8,022		3,311				-3.41				-0.02			
Existing accounts	Institutions	396	396	510,013	389,369	-120,644	-2.61	216.6	29.2	-187.4	-0.77	0.58	0.28	-0.30	-3.79
	Retail investors	12,218	12,218	7,015	4,865	-2,150	-6.65	3.69	2.84	-0.85	-0.70	0.05	0.06	0.02	1.74
New accounts	Institutions		84		156,493				178.2				0.35		
	Retail investors		3,206		4,619				-1.89				0.00		

Table 4
Analysis of four-month windows around the reform

This table repeats the same univariate analysis as in Table 3, but for the four-month windows around the contract size reform. The reform increased the options contract size (multiplier) from KRW 100,000 to KRW 500,000. This was announced in January 2012. A step-by-step implementation process started in March 2012 and the reform was fully implemented by June 2012. We use June 2012 as the effective start of the reform because it is the first month when liquid option contracts were traded with the new multiplier. (For details about the reform implementation schedule, refer to Section 2 - footnote 1). We exclude the month of the announcement and the transitional period when there were listed contracts with both the old and the new multiplier. Therefore, we compare the four months before the reform was announced to the four months after it was fully implemented. The group of "existing accounts" is the same as in Table 3: these investors exist in the market starting from four months before the crisis and remain in the market for at least one year and a half until four months after the reform (Apr 2011 – Sep 2012). Accounts that exit are those that are in the market during the four months before the reform announcement, but not in the four months after the reform was implemented. New accounts are those that are not in the market in the four months before the reform announcement, but are in the market in the four months since the start of the reform. We calculate statistics for each four-month window, for institutions and retail investors in each of the three account groups. We report the number of accounts in each category, and the cross-sectional averages of their monthly mean trading volumes, their total $P&L (in millions of KRW) over the four-month periods, and their monthly mean returns. Refer to Section 2 for details about the method we use to calculate profitability. To adjust for the change in contract size, we multiply trading volumes after the reform by a factor of five. For the group of existing accounts, we calculate the change in volumes and profitability from the four-month period before the crisis to the four-month period after it, and we perform paired t-tests to evaluate the statistical significance of the differences in means.

Before reform: Sep 2011 – Dec 2011 [t-9, t-6]
After reform: Jun 2012 – Sep 2012 [t, t+3]

		Number of accounts		**Monthly mean TRDVOL** cross-sectional average				**Total $P&L (millions of KRW)** cross-sectional average				**Monthly mean RET** cross-sectional average			
		before	**after**	**before**	**after x5**	**change**	**t-stat**	**before**	**after**	**change**	**t-stat**	**before**	**after**	**change**	**t-stat**
Accounts that exit	Institutions	207		162,165				-131.6				0.51			
	Retail investors	7,950		3,555				-3.78				-0.08			
Existing accounts	Institutions	396	396	389,369	379,798	-9,571	-0.31	29.2	239.7	210.5	0.66	0.28	0.20	-0.08	-0.63
	Retail investors	12,218	12,218	4,865	5,392	526	3.41	2.84	-2.60	-5.44	-4.69	0.06	-0.03	-0.10	-11.9
New accounts	Institutions		259		186,106				60.4				0.25		
	Retail investors		4,826		6,831				-10.1				-0.02		

Table 5
Propensity to stay and be active in the options market: Account-month logit regressions

This table presents account-month logit regressions for investors' propensity to stay and be active in the options market. In Panels A and C, the dependent variable is the dummy variable *Staying in market*, equal to one in month t if the account: 1) has previously entered the market by trading intraday and/or holding option positions, 2) has not exited the market yet (exit defined as becoming inactive for six or more months), and 3) does not close his positions and exit in month t. In Panels B and D, the dependent variable is the dummy variable *Active in market*, equal to one in month t if the investor trades and/or holds option positions, thereby generating a P&L in month t. The regressions with dependent variable *Active in market* are meant to show how active the accounts are in the months when they choose to stay in the market (observations where *Staying in market* = 1). We regress the dependent variables on three dummy variables for the different time periods: *Crisis* is equal to one in the month of the market crash (August 2011), *Post-crisis* is equal to one for all months in the sample following the market crash, and *Post-reform* is equal to one from the starting date of the reform (June 2012) until the end of the sample. Our sample covers the period from January 2010 to June 2014, but we exclude the last six months of the sample because we do not know whether accounts exit or stay in the market for the following six months. In all regressions, we control for trading volume of each account in the previous month, KOSPI 200 index return in the current and previous month, the implied volatility index VKOSPI at the start of the current and previous month, and any time trend in the data. We multiply post-reform trading volumes by five, and we take the logarithm of trading volume and VKOSPI variables. We re-run each regression a second time where we also control for the account’s previous-month return or the compounded return from the previous and the current month, as well as interactions of the return and the event dummies. We add these controls separately because they require the account to have activity (and therefore non-missing values for returns) in both months t and t-1, which slightly reduces the size of the sample. Panels A and B show the effects of the crisis and the reform on investors’ propensity to stay and be active in the options market, separately in the two subsamples of institutions and retail investors. Panels C and D show the effects in the full sample of all accounts. We add a dummy for the institutions and interact it with the event dummies. Since all the remaining accounts are retail investors, the event dummies without an interaction represent that group of investors. By combining institutions and retail investors into one group, these regressions allow us to compare the magnitude of the effects on the two investor classes.

Panel A: Propensity to stay in the options market

Accounts:	Institutions	Institutions	Retail investors	Retail investors
Dependent variable:	Staying in market (t)	Staying in market (t)	Staying in market (t)	Staying in market (t)
Intercept	1.122***	0.249	5.307***	6.490***
Crisis	**-0.002**	**-0.195**	**-0.484*** **	**-0.251*** **
Post-crisis	**0.241*** **	**0.137**	**0.418*** **	**0.769*** **
Post-reform	**-0.027**	**0.136* **	**0.188*** **	**0.203*** **
RET (t-1,t)		0.014***		0.029***
Crisis * RET (t-1,t)		-0.008		-0.019***
Post-crisis * RET (t-1,t)		-0.009*		-0.009***
Post-reform * RET (t-1,t)		-0.003		-0.012***
TRDVOL (t-1)	0.107***	0.136***	0.085***	0.112***
KOSPI 200 Index RET (t)	1.265***	0.751	2.553***	2.412***
KOSPI 200 Index RET (t-1)	-0.342	-2.017***	2.672***	2.779***
VKOSPI (t)	-0.610***	-0.754***	-0.602***	-0.544***
VKOSPI (t-1)	0.480***	1.099***	-0.615***	-0.884***
Time trend	-0.008**	-0.009*	-0.030***	-0.042***
N observations	53,880	40,790	1,710,521	1,322,468

Panel B: Propensity to be active in the options market

Accounts:	Institutions	Institutions	Retail investors	Retail investors
Dependent variable:	Active in market (t)	Active in market (t)	Active in market (t)	Active in market (t)
Intercept	0.851*	1.683***	-0.110	0.511***
Crisis	**0.485*** **	**0.377** **	**0.440*** **	**0.366*** **
Post-crisis	**-0.080**	**0.020**	**0.072*** **	**0.079*** **
Post-reform	**-0.163** **	**-0.189** **	**-0.264*** **	**-0.382*** **
RET (t-1)		0.039**		0.149***
Crisis * RET (t-1)		0.136		0.154***
Post-crisis * RET (t-1)		-0.039		-0.069***
Post-reform * RET (t-1)		0.032		-0.017**
TRDVOL (t-1)	0.325***	0.239***	0.444***	0.283***
KOSPI 200 Index RET (t)	1.396***	2.234***	0.248***	0.624***
KOSPI 200 Index RET (t-1)	1.231**	1.750***	0.129	0.172
VKOSPI (t)	-0.594***	-0.687***	-0.230***	-0.310***
VKOSPI (t-1)	0.356**	0.387*	0.296***	0.431***
Time trend	0.003	0.003	0.006***	0.014***
N observations	43,737	39,559	1,445,315	1,294,844

Panel C: Comparing institutions' and retail investors' propensity to stay

Accounts:	All accounts	All accounts
Dependent variable:	Staying in market (t)	Staying in market (t)
Intercept	5.194***	6.333***
Crisis	**-0.495*****	**-0.261*****
Post-crisis	**0.395*****	**0.739*****
Post-reform	**0.194*****	**0.213*****
Crisis * Institution	**0.889*****	**0.612*****
Post-crisis * Institution	**0.539*****	**0.487*****
Post-reform * Institution	**-0.367*****	**-0.371*****
Institution	**-0.643*****	**-0.568*****
RET (t-1,t)		0.027***
Crisis * RET (t-1,t)		-0.018***
Post-crisis * RET (t-1,t)		-0.008***
Post-reform * RET (t-1,t)		-0.012***
Crisis * Institution * RET (t-1,t)		-0.003
Post-crisis * Institution * RET (t-1,t)		-0.017***
Post-reform * Institution * RET (t-1,t)		0.012***
TRDVOL (t-1)	0.087***	0.114***
KOSPI 200 Index RET (t)	2.518***	2.382***
KOSPI 200 Index RET (t-1)	2.574***	2.638***
VKOSPI (t)	-0.603***	-0.553***
VKOSPI (t-1)	-0.580***	-0.828***
Time trend	-0.030***	-0.041***
N observations	1,764,401	1,363,258

Panel D: Comparing institutions' and retail investors' propensity to be active

Accounts:	All accounts	All accounts
Dependent variable:	Active in market (t)	Active in market (t)
Intercept	-0.073	0.553***
Crisis	**0.445*** **	**0.376*** **
Post-crisis	**0.075*** **	**0.088*** **
Post-reform	**-0.265*** **	**-0.382*** **
Crisis * Institution	**-0.161**	**-0.372** **
Post-crisis * Institution	**-0.317*** **	**-0.398*** **
Post-reform * Institution	**0.117** **	**0.202*** **
Institution	**-0.355*** **	**-0.421*** **
RET (t-1)		0.143***
Crisis * RET (t-1)		0.161***
Post-crisis * RET (t-1)		-0.063***
Post-reform * RET (t-1)		-0.017**
Crisis * Institution * RET (t-1)		-0.120
Post-crisis * Institution * RET (t-1)		-0.077***
Post-reform * Institution * RET (t-1)		0.044*
TRDVOL (t-1)	0.439***	0.280***
KOSPI 200 Index RET (t)	0.277***	0.667***
KOSPI 200 Index RET (t-1)	0.155*	0.221**
VKOSPI (t)	-0.240***	-0.318***
VKOSPI (t-1)	0.298***	0.429***
Time trend	0.006***	0.014***
N observations	1,489,052	1,334,403

Table 6
Trading volumes and profitability: Account-month OLS regressions

This table shows the effects of the crisis and the reform on investors' trading volumes, dollar profitability and returns. In particular, we are interested in the events' effects on the group of "existing accounts" which we previously identified in Tables 3 and 4 as those investors who stay in the market in the long run, starting from at least four months before the crisis until at least four months after the reform. Table 6 reports the results of a series of account-month OLS regressions, where the independent variables are the same as in Table 5: *Crisis* is equal to one in the month of the market crash (August 2011), *Post-crisis* is equal to one for all months in the sample following the market crash, and *Post-reform* is equal to one from the starting date of the reform (June 2012) until the end of the sample. Our sample covers the period from January 2010 to June 2014. Contracts that start being traded in 2009 are excluded. Since we observe only a portion of the total trading volumes at the beginning of our sample, we exclude the first two months. The control variables used in the regressions are the same as in Table 5. In Panels A and D, the dependent variable is trading volume (number of options contracts traded). We multiply post-reform trading volumes by five to adjust for the change in contract size, and we take the logarithm of the variable. In Panels B and E, the dependent variable is dollar profit and loss ($P&L) in KRW. We also take the logarithm of this variable, and we control for previous-month $P&L instead of return. In Panels C and F, the dependent variable is monthly return, which we measure as $P&L scaled by capital requirement. Refer to Section 2 for details about the calculation method. Panels A, B and C show the effects of the crisis and the reform in the two subsamples of institutions and retail investors separately. Panels D, E and F show the effects in the full sample of all accounts. We add a dummy for the institutions and interact it with the event dummies. Since all the remaining accounts are retail investors, the event dummies without an interaction represent that group of investors. By combining institutions and retail investors into one group, these regressions allow us to compare the magnitude of the effects on the two investor classes.

Panel A: Trading volumes

Accounts:	Existing Institutions	Existing Institutions	Existing Retail investors	Existing Retail investors
Dependent variable:	TRDVOL (t)	TRDVOL (t)	TRDVOL (t)	TRDVOL (t)
Intercept	2.115***	1.094**	1.657***	0.965***
	(4.45)	(2.57)	(23.64)	(15.01)
Crisis	**0.567*** **	**0.556*** **	**0.245*** **	**0.193*** **
	(4.07)	**(4.28)**	**(11.81)**	**(10.18)**
Post-crisis	**-0.013**	**0.051**	**-0.114*** **	**-0.110*** **
	(-0.13)	**(0.54)**	**(-7.37)**	**(-7.74)**
Post-reform	**0.074**	**0.109***	**0.132*** **	**0.101*** **
	(1.08)	**(1.78)**	**(13.02)**	**(10.85)**
RET (t-1)		0.020		0.032***
		(1.4)		(10.9)
Crisis * RET (t-1)		-0.096		0.057***
		(-1.49)		(4.29)
Post-crisis * RET (t-1)		0.000		0.020***
		(-0.02)		(4.92)
Post-reform * RET (t-1)		0.008		-0.003
		(0.42)		(-0.86)
TRDVOL (t-1)	0.839***	0.956***	0.752***	0.867***
	(192.81)	(208.09)	(817.49)	(857.85)
KOSPI 200 Index RET (t)	1.064**	0.638	0.044	0.129**
	(2.44)	(1.64)	(0.69)	(2.19)
KOSPI 200 Index RET (t-1)	0.059	0.766*	1.018***	0.897***
	(0.12)	(1.69)	(13.61)	(13.07)
VKOSPI (t)	-0.759***	-0.561***	-0.060**	-0.105***
	(-4.6)	(-3.81)	(-2.47)	(-4.67)
VKOSPI (t-1)	0.615***	0.339**	0.032	0.047**
	(3.54)	(2.18)	(1.24)	(2.00)
Time trend	-0.010**	-0.013***	-0.002***	-0.004***
	(-2.32)	(-3.37)	(-3.56)	(-6.68)
N observations	16,064	15,143	520,181	486,507
Adjusted R^2	0.7001	0.7429	0.5636	0.6037

Panel B: Profitability (in KRW)

Accounts:	Existing Institutions	Existing Institutions	Existing Retail investors	Existing Retail investors
Dependent variable:	$P&L (t)	$P&L (t)	$P&L (t)	$P&L (t)
Intercept	-0.891	-3.952	-5.093***	-4.546***
	(-0.24)	(-1.04)	(-8.97)	(-7.99)
Crisis	**-1.378**	**-0.057**	**-4.052*** **	**-4.481*** **
	(-1.25)	**(-0.05)**	**(-24.17)**	**(-26.68)**
Post-crisis	**1.069**	**0.531**	**-0.401*** **	**-0.229* **
	(1.28)	**(0.62)**	**(-3.19)**	**(-1.82)**
Post-reform	**-1.372** **	**-0.610**	**-0.781*** **	**-0.814*** **
	(-2.55)	**(-1.1)**	**(-9.51)**	**(-9.84)**
$P&L (t-1)		0.165***		0.153***
		(11.4)		(59.99)
Crisis * $P&L (t-1)		-0.256***		-0.178***
		(-4.68)		(-18.58)
Post-crisis * $P&L (t-1)		0.031		0.104***
		(1.4)		(26.09)
Post-reform * $P&L (t-1)		-0.092***		-0.087***
		(-4.38)		(-23.32)
TRDVOL (t-1)	0.982***	0.988***	-0.093***	-0.083***
	(26.34)	(22.93)	(-11.56)	(-9.16)
KOSPI 200 Index RET (t)	-5.221	-5.509	6.609***	6.709***
	(-1.51)	(-1.58)	(12.67)	(12.85)
KOSPI 200 Index RET (t-1)	8.182**	11.481***	-5.851***	-7.890***
	(2.04)	(2.84)	(-9.65)	(-12.99)
VKOSPI (t)	-0.177	1.448	0.377*	0.776***
	(-0.13)	(1.1)	(1.9)	(3.92)
VKOSPI (t-1)	-0.955	-1.989	0.998***	0.481**
	(-0.69)	(-1.43)	(4.8)	(2.31)
Time trend	-0.011	0.002	0.004	0.004
	(-0.32)	(0.07)	(0.77)	(0.81)
N observations	15,143	14,662	486,499	469,492
Adjusted R^2	0.0450	0.0700	0.0059	0.0406

Panel C: Returns

Accounts:	Existing Institutions	Existing Institutions	Existing Retail investors	Existing Retail investors
Dependent variable:	RET (t)	RET (t)	RET (t)	RET (t)
Intercept	0.387 (0.81)	0.348 (0.73)	-0.044 (-0.73)	-0.258*** (-4.25)
Crisis	**-0.902*** (-6.46)**	**-0.988*** (-6.77)**	**-0.719*** (-39.88)**	**-0.703*** (-39.11)**
Post-crisis	**-0.222** (-2.11)**	**-0.087 (-0.81)**	**0.025* (1.82)**	**-0.013 (-0.96)**
Post-reform	**-0.274*** (-4.00)**	**-0.211*** (-3.07)**	**-0.120*** (-13.62)**	**-0.083*** (-9.45)**
RET (t-1)		0.364*** (22.55)		0.237*** (85.99)
Crisis * RET (t-1)		0.053 (0.74)		-0.904*** (-73.15)
Post-crisis * RET (t-1)		-0.171*** (-7.36)		0.064*** (16.28)
Post-reform * RET (t-1)		-0.081*** (-3.98)		-0.091*** (-25.89)
TRDVOL (t-1)	0.010** (2.02)	0.002 (0.34)	-0.008*** (-9.37)	-0.012*** (-12.02)
KOSPI 200 Index RET (t)	-2.468*** (-5.63)	-2.388*** (-5.45)	1.195*** (21.29)	1.050*** (18.84)
KOSPI 200 Index RET (t-1)	1.847*** (3.63)	2.815*** (5.51)	-0.085 (-1.3)	-0.316*** (-4.87)
VKOSPI (t)	0.492*** (2.96)	0.713*** (4.28)	-0.086*** (-4.02)	0.103*** (4.88)
VKOSPI (t-1)	-0.545*** (-3.11)	-0.782*** (-4.46)	0.124*** (5.52)	0.010 (0.45)
Time trend	0.016*** (3.71)	0.014*** (3.28)	0.002*** (3.71)	0.003*** (5.19)
N observations	15,143	14,662	486,499	469,492
Adjusted R^2	0.0060	0.0549	0.0101	0.0734

Panel D: Comparing institutions' and retail investors' trading volumes

Accounts:	Existing accounts	Existing accounts
Dependent variable:	TRDVOL (t)	TRDVOL (t)
Intercept	1.648***	0.955***
	(23.71)	(14.99)
Crisis	**0.254*** **	**0.200*** **
	(12.22)	**(10.52)**
Post-crisis	**-0.107*** **	**-0.102*** **
	(-6.91)	**(-7.22)**
Post-reform	**0.134*** **	**0.103*** **
	(13.30)	**(11.11)**
Crisis * Institution	**0.068**	**0.207** **
	(0.72)	**(2.24)**
Post-crisis * Institution	**-0.050**	**-0.010**
	(-1.25)	**(-0.28)**
Post-reform * Institution	**-0.128*** **	**-0.070** **
	(-3.45)	**(-2.01)**
Institution	**0.835*** **	**0.401*** **
	(32.00)	**(16.87)**
RET (t-1)		0.032***
		(11.01)
Crisis * RET (t-1)		0.058***
		(4.35)
Post-crisis * RET (t-1)		0.021***
		(5.05)
Post-reform * RET (t-1)		-0.003
		(-0.91)
Crisis * Institution * RET (t-1)		-0.177***
		(-3.19)
Post-crisis * Institution * RET (t-1)		-0.030**
		(-2.32)
Post-reform * Institution * RET (t-1)		0.012
		(0.73)
TRDVOL (t-1)	0.757***	0.872***
	(843.43)	(886.27)
KOSPI 200 Index RET (t)	0.078	0.148**
	(1.22)	(2.53)
KOSPI 200 Index RET (t-1)	0.990***	0.892***
	(13.36)	(13.14)
VKOSPI (t)	-0.083***	-0.121***
	(-3.45)	(-5.47)
VKOSPI (t-1)	0.049*	0.057**
	(1.92)	(2.43)
Time trend	-0.003***	-0.004***
	(-4.27)	(-7.51)
N observations	536,245	501,650
Adjusted R^2	0.5871	0.6294

Panel E: Comparing institutions' and retail investors' profitability (in KRW)

Accounts:	Existing accounts	Existing accounts
Dependent variable:	$P&L (t)	$P&L (t)
Intercept	-5.179***	-4.730***
	(-9.2)	(-8.39)
Crisis	**-4.080*** **	**-4.521*** **
	(-24.31)	**(-26.89)**
Post-crisis	**-0.323*** **	**-0.175**
	(-2.59)	**(-1.39)**
Post-reform	**-0.802*** **	**-0.837*** **
	(-9.78)	**(-10.15)**
Crisis * Institution	**4.111*** **	**5.673*** **
	(5.4)	**(6.91)**
Post-crisis * Institution	**0.165**	**-0.483**
	(0.51)	**(-1.45)**
Post-reform * Institution	**0.143**	**0.850*** **
	(0.48)	**(2.69)**
Institution	**6.965*** **	**5.952*** **
	(33.16)	**(28.25)**
$P&L (t-1)		0.156***
		(61.83)
Crisis * $P&L (t-1)		-0.180***
		(-18.71)
Post-crisis * $P&L (t-1)		0.101***
		(25.45)
Post-reform * $P&L (t-1)		-0.087***
		(-23.18)
Crisis * Institution * $P&L (t-1)		-0.017
		(-0.36)
Post-crisis * Institution * $P&L (t-1)		-0.013
		(-0.85)
Post-reform * Institution * $P&L (t-1)		-0.009
		(-0.49)
TRDVOL (t-1)	-0.028***	-0.019**
	(-3.5)	(-2.17)
KOSPI 200 Index RET (t)	6.254***	6.357***
	(12.09)	(12.28)
KOSPI 200 Index RET (t-1)	-5.386***	-7.277***
	(-8.96)	(-12.08)
VKOSPI (t)	0.345*	0.781***
	(1.75)	(3.98)
VKOSPI (t-1)	0.929***	0.404*
	(4.5)	(1.96)
Time trend	0.001	0.002
	(0.27)	(0.49)
N observations	501,642	484,154
Adjusted R^2	0.0129	0.0477

Panel F: Comparing institutions' and retail investors' returns

Accounts:	Existing accounts	Existing accounts
Dependent variable:	RET (t)	RET (t)
Intercept	-0.044 (-0.72)	-0.251*** (-4.14)
Crisis	**-0.739*** (-40.66)**	**-0.721*** (-39.75)**
Post-crisis	**0.021 (1.58)**	**-0.014 (-1.04)**
Post-reform	**-0.128*** (-14.41)**	**-0.091*** (-10.31)**
Crisis * Institution	**0.489*** (5.93)**	**0.260*** (2.98)**
Post-crisis * Institution	**-0.108*** (-3.1)**	**-0.079** (-2.24)**
Post-reform * Institution	**0.095*** (2.91)**	**0.126*** (3.81)**
Institution	**0.479*** (21.06)**	**0.395*** (17.47)**
RET (t-1)		0.244*** (89.62)
Crisis * RET (t-1)		-0.911*** (-72.72)
Post-crisis * RET (t-1)		0.057*** (14.48)
Post-reform * RET (t-1)		-0.090*** (-25.48)
Crisis * Institution * RET (t-1)		1.082*** (20.73)
Post-crisis * Institution * RET (t-1)		-0.101*** (-8.08)
Post-reform * Institution * RET (t-1)		-0.002 (-0.15)
TRDVOL (t-1)	-0.007*** (-8.34)	-0.011*** (-11.39)
KOSPI 200 Index RET (t)	1.082*** (19.31)	0.944*** (16.95)
KOSPI 200 Index RET (t-1)	-0.023 (-0.36)	-0.223*** (-3.45)
VKOSPI (t)	-0.068*** (-3.19)	0.122*** (5.76)
VKOSPI (t-1)	0.102*** (4.57)	-0.014 (-0.63)
Time trend	0.002*** (4.4)	0.003*** (5.83)
N observations	501,642	484,154
Adjusted R^2	0.0118	0.0739

Table 7
How persistent is the effect of the reform on retail trading volumes?

In this table, we test how persistent the effect of the reform is on the trading volumes of existing retail investors. We repeat the same analysis as the first regression in Panel A of Table 6, but we add twelve dummy variables for the twelve months after the start of the reform. Refer to Table 6 for details about the rest of the variables.

Accounts:	Existing Retail investors
Dependent variable:	TRDVOL (t)
Intercept	1.953***
	(26.04)
Crisis	0.433***
	(19.57)
Post-crisis	-0.046**
	(-2.56)
Post-reform	**0.194*** **
	(10.05)
1m Post-reform	**0.181*** **
	(8.23)
2m Post-reform	**0.053** **
	(2.41)
3m Post-reform	**-0.047** **
	(-2.29)
4m Post-reform	**-0.175*** **
	(-8.63)
5m Post-reform	**-0.152*** **
	(-7.45)
6m Post-reform	**-0.127*** **
	(-6.27)
7m Post-reform	**-0.527*** **
	(-25.42)
8m Post-reform	**0.383*** **
	(18.77)
9m Post-reform	**-0.240*** **
	(-11.63)
10m Post-reform	**0.024**
	(1.20)
11m Post-reform	**0.091*** **
	(4.55)
12m Post-reform	**-0.265*** **
	(-13.19)
Control variables	YES
N observations	520,181
Adjusted R^2	0.5651