Retail Trading in Options and the Rise of the Big Three Wholesalers

Svetlana Bryzgalova,* Anna Pavlova,† and Taisiya Sikorskaya‡

May 16, 2022

Abstract

We document rapid increases in (i) retail trading in options and (ii) payment for order flow, received by the U.S. retail brokerages from the so-called wholesalers in exchange for routing orders to them. Exploiting new flags in transaction-level data, we isolate wholesaler trades and build a novel proxy for retail options trading. Often cash-constrained, retail investors prefer cheaper, lottery-like weekly options, with the average bid-ask spread of a whopping 12.6%. They lose money on average and participate in frenzies. The inflow of retail investors also coincides with an increase in call options left suboptimally unexercised. Arbitrageurs exploit these investor mistakes via so-called "dividend play" trades, producing (virtually) riskless arbitrage profits. Puzzlingly, they forgo 50% of these profits, leaving money on the table for option writers.

JEL Classification: G4, G5, G11, G12, D45

Keywords: Retail, payment for order flow, dividend play, suboptimal option exercise, market concentration, arbitrage, financial literacy, WallStreetBets

*We thank Brad Barber, James Choi, Francisco Gomes, Ahmed Guecioueur, Björn Hagströmer, Su Li, Camelia Kuhnen, Dmitriy Muravyev, Veronika Pool, Chester Spatt (discussant), Dimitri Vayanos, Ivo Welch, Ingrid Werner, seminar participants at the 2022 NBER Asset Pricing meeting, Indiana University, London Business School, Queen Mary University of London, Stockholm Business School, the University of Central Florida, the University of Notre Dame, and the University of Hong Kong for their thoughtful comments. We thank Brad Barber for providing data on Robinhood investor frenzies. We have also benefited from conversations with Ed Monrad and Shyam Chandramouli of Optiver and Parker Lim of XTX Markets.

*London Business School and CEPR, sbryzgalova@london.edu

†London Business School and CEPR, apavlova@london.edu

‡London Business School, tsikorskaya@london.edu

1 Introduction

The advent of zero-commission trading in stocks and options has revolutionized retail brokerage services in the United States. Since their market entry in 2015, smartphone investing app Robinhood and other commission-free brokerages have attracted an unprecedented inflow of retail customers, mainly young and tech-savvy, yet inexperienced, investors. By the end of 2021, Robinhood alone has amassed 22.5 million active users.

One concern frequently brought up in the context of the recent retail trading boom is related to the controversial practice of payment for order flow (PFOF). Retail brokerages route clients' orders to financial intermediaries known as wholesalers for execution and receive PFOF in return. In equities, wholesalers cross this order flow on their private trading platforms, away from national exchanges, and other market makers cannot compete for these orders. This is known as *internalization*. PFOF is a divisive practice because such order flow fragmentation may lead to wider bid-ask spreads on exchanges and because it incentivizes retail brokerages to encourage investors to trade more. U.S. policymakers are currently reviewing PFOF, but their attention has been focused solely on equities.[1] Unlike equities, all options in the U.S. trade on exchanges, which mechanically should expose them to competition from other market makers. It is therefore thought that problems arising from internalization are specific to equities.

In this paper, we argue that much of the retail order flow in options is also effectively internalized. We identify a friction that may hinder competition from other market makers on options exchanges. Specifically, wholesalers execute retail orders through so-called *price improvement mechanisms*, which, as we show, often amounts to internalization. This allows us to isolate wholesaler trades and build a proxy for retail trading in options by exploiting a recently introduced flag for price improvement mechanisms in transaction-level data. We find that our measure of retail trading grew 104% in our sample of one and a half years, in line with the growth in PFOF for options.[2] Retail traders prefer cheaper, weekly options, the average quoted bid-ask spread for which is as high as 12.6%, and lose money on average. A large fraction of retail order flow is serviced by very few wholesalers: The share in PFOF of the top three has grown to nearly 90% as of the second quarter of 2021.

We also explore how arbitrageurs react to the inflow of retail investors. As a laboratory for the study of arbitrage activity, we use so-called *dividend play* trades, which are

[1] In June 2021, Gary Gensler, chair of the U.S. Securities and Exchange Commission (SEC), announced an upcoming comprehensive review of the current microstructure rules, including the system of PFOF; see https://www.ft.com/content/83dff8fc-14ac-4e67-a969-20b358c349e8.

[2] We consider the combined PFOF from the largest U.S. retail brokerages reports under SEC Rule 606 (routing of orders). See Section 3.1 for the list of brokerages in our sample.

pairs trades that produce (virtually) riskless arbitrage profits. These profits derive from call options left suboptimally unexercised before the underlying pays a dividend. An inflow of retail investors has boosted potential gains from this strategy. We identify dividend play trades in transaction-level data and document that, instead of harvesting all the available arbitrage profits, market participants leave about 50% of them on the table. This puzzling behavior cannot be explained by arbitrageur costs or constraints.

We start by documenting a stylized fact that, although only a fraction of investors trade options, most of PFOF received by retail brokerages comes from options, not equities. For example, in 2021, U.S. brokerages received \$2.4 billion in PFOF for options and only \$1.3 billion for equities. The lion's share of PFOF for options came from only three wholesalers: Citadel, Susquehanna, and Wolverine.

Retail brokers in the U.S. are required to provide the best execution to their clients, and so they have an agreement with a wholesaler to provide price improvement relative to the best available bid and ask prices.[3] To do so, they would often use an options exchange process known as a price improvement auction or mechanism. Exploiting a flag for price improvement mechanisms, introduced by the Options Price Reporting Authority (OPRA) in November 2019 for transaction-level data, we are able to identify wholesaler trades and build a novel measure of retail trading in options. In our dataset, these are trades executed through a single-leg price improvement mechanism, for which we abbreviate as SLIM.[4] The monthly dollar trading volume in SLIM transactions grew by 104% from January 2020 to July 2021, alongside the PFOF in options (158%).

We show that our measure picks up recent retail investor frenzies in GameStop and other "meme" stocks, as measured by mentions in `WallStreetBets`, an investing forum popular with new retail investors. Furthermore, it is strongly correlated with an alternative retail investor trading measure – small trades in options (up to 10 contracts) – commonly used in the media and industry,[5] as well as with Robinhood user popularity provided by Robintrack, and the retail frenzies measure of Barber, Huang, Odean, and Schwartz (2022). We also construct a novel retail popularity measure based on the internalized volume in equities and exchange-traded funds (ETFs), and show that it comoves with SLIM trades in the corresponding options. We provide further validation of our measure in Section 3.3.2.

The new generation of retail investors is more tech-savvy and participates in invest-

[3]Most of order flow in options received by retail brokerages in our sample is routed to wholesalers. The fraction of orders routed directly to exchanges is small; see Table A1 in the Appendix.

[4]Specifically, we use OPRA type "SLAN," which stands for single-leg non-ISO price improvement mechanisms. See Appendix A.2 for a description.

[5]For instance, Bloomberg relies on small trades to proxy retail participation in options; see `https://www.bloomberg.com/professional/blog/gamestop-highlights-importance-of-option-related-equity-flows/`.

ment forums, but they are still financial novices. It is quite striking that they are so active in options markets, despite much higher bid-ask spreads on options relative to stocks.[6] For example, 50% of retail trades in our sample are in ultra short-term options, weeklys (i.e., options with less than a week to expiration), with an average quoted bid-ask spread of 12.6%. However, the true trading costs for options may be obfuscated by the zero commissions; an opportunity to trade options is displayed prominently on gamified investing apps used by the new generation of investors.[7] These investors may be attracted by a cheap way of achieving leverage that these options provide.[8] Moreover, on some investing platforms, for example, Robinhood, weekly options are presented as a default choice to an options trader.

Retail investors in our sample strongly prefer call options to puts: The volume share in calls is 69%. They trade mostly at-the-money (72% of trades) or slightly-out-of-the-money (24% of trades) options. The latter involve higher trading costs, with the average quoted bid-ask spread of 28%. 14% of retail trades have a "micro" size of up to $250, and their average quoted bid-ask spread is 23.6%. We document that retail investors prefer options on larger companies, those with lower share prices, and higher recent trading volume (e.g., attention-grabbing tickers). This is consistent with the literature on retail participation in equities. We view these cross-sectional relationships as evidence of speculative rather than hedging motives behind retail trades. Finally, we document significant increases in both call and put net purchases during retail investor frenzies, especially in trades of a smaller size.

Are retail options trades profitable? To answer this question, we analyze performance of SLIM trades at the one-, two-, five-, 10-day horizons, and until expiration. On aggregate, these trades lose money for all horizons. For example, assuming a holding horizon of 10 days, we estimate that the aggregate portfolio of retail investors lost $1.2 billion from November 2019 until June 2021. The losses are concentrated primarily in short-term call options. Moreover, this calculation does not include trading costs. The aggregate trading costs, measured as a distance from an actual trade price to midquote for all SLIM trades in our sample, amount to a staggering $5.2 billion. This number is much higher than direct trading costs (about $950 million), computed using commissions of retail brokerages in our sample.[9]

Given the recent surge in retail investor activity, it is important to understand its

[6] Muravyev and Pearson (2020) report that the average quoted bid-ask spread of options on stocks in the S&P 500 is as high as 17.2%. For comparison, for the S&P 500 stocks, this number is 3.55bps (as reported in Hagströmer (2021)). Higher aggregate PFOF for options relative to that for stocks (see Table A3 in the Appendix) indicates that executing order flow in options is a very lucrative business for wholesalers.

[7] Chapkovski, Khapko, and Zoican (2021) show that gamification induces risk-taking in novice traders, while Kalda, Loos, Previtero, and Hackethal (2021) find that trading on smartphones induces investors to purchase riskier and lottery-type assets.

[8] Weekly options may have embedded leverage of over 50 (see Table A4 in the Appendix). See also `https://www.nasdaq.com/articles/you-should-be-trading-weekly-options-and-heres-why-2021-01-20`.

[9] Robinhood does not charge commissions for options trades, but many other brokerages still do.

implications for behavior of arbitrageurs in the options market. We focus on one specific mistake that option investors make, for which we can cleanly identify the trading patterns of arbitrageurs who exploit it. This mistake is a failure to exercise in-the-money call options before the underlying stock goes ex-dividend when it is optimal to do so.[10] To benefit from it, arbitrageurs engage in a "dividend play," an arbitrage strategy that diverts windfall gains from the writer of the option that was suboptimally left unexercised. The strategy is normally executed on a physical exchange floor,[11] hence available to market makers and other participants on the floor. Due to the dividend play, the daily trading volume on last cum-dividend dates in in-the-money call options for which early exercise is optimal often exceeds trading volume on the remaining dates by several orders of magnitude.[12]

We exploit the new OPRA trade flags to accurately identify dividend play trades and study the behavior of arbitrageurs. Arbitrageurs' expected profits have been boosted by the recent retail investor boom. Most of them derive from the sheer increase in open interest due to investor inflow, coupled with a higher fraction of options that are left unexercised on cum-dividend dates. Overall, traders engaging in the dividend play behave like unconstrained arbitrageurs in harvesting the windfall gain from failures to exercise options.

There is, however, one striking pattern that emerges from our examination of dividend play trades: Arbitrageurs exploit less than 50% of available arbitrage profits, leaving the rest on the table.[13] They often exploit profitable opportunities in one contract on a particular stock, while leaving another very similar one aside. This is extremely puzzling. Arbitrageurs' daily fee on dividend play trades is capped at a ticker level. Furthermore, other trading costs are very low because such transactions are typically pre-arranged by pairs of arbitrageurs. We discuss the role of transaction costs in detail in Section 5.4.

We rule out further potential explanations of our money-left-on-the-table puzzle, such as capital/margin constraints. Arbitrageur's risk exposure is usually computed at a ticker level, and so the large long and short positions in contracts on the same ticker, forming a dividend play trade, are effectively netted to zero. It is possible that the reluctance of some firms to engage in the dividend play arbitrage could be explained by the operational risk of

[10] We note that sometimes call options may be purchased as part of any strategy that involves holding multiple option contracts. In those circumstances, or whenever transactions costs overweigh profits from early exercise, exercising an option may not be optimal.

[11] Some exchanges facilitate dividend play and other strategies by imposing daily fee caps for floor market makers and other floor traders engaging in such strategies. See e.g., https://listingcenter.nasdaq.com/rulebook/phlx/rules/phlx-options-7, accessed January 12, 2022, for the dividend strategy fee caps imposed by PHLX. Over 2/3 of dividend play transactions in our sample are executed on PHLX.

[12] Even for SPY, the ticker with the most actively traded options in 2021, cum-dividend day trading volume in calls is up to 50 times larger than on any other day.

[13] Table A5 in the Appendix quantifies forgone profits of market makers in the top 40 most popular underlying stocks and ETFs for the dividend play strategy in our sample.

the trade or a stigma associated with it, since this strategy is frowned upon by the SEC.

The only other explanation that has some potential in our context is that this trade is dominated by very few arbitrageurs. We find that market participants avoid engaging in a dividend play strategy in call option contracts that had experienced higher buying pressure from retail investors (as measured by *SLIM order imbalances*) in the week preceding the cum-dividend date. This effect is especially large for tickers that have a large share of volume executed by the Big Three PFOF providers—Citadel, Susquehanna, and Wolverine—in the preceding week. This points to the conclusion that the Big Three wholesalers are the writers of call options purchased by retail investors and hence they are set to receive the windfall gain if retail investors leave their options suboptimally unexercised. It is therefore suboptimal for them to engage in the dividend play trade in those contracts. Intriguingly, other market participants appear to avoid those contracts too, effectively leaving windfall gains in those contracts to the option writers, who are likely to be the Big Three wholesalers.

Our paper offers several policy implications. Unlike reporting required by Financial Industry Regulatory Authority (FINRA) in equities, there is little transparency on wholesaler activities in the options market. Current price improvement mechanisms on U.S. options exchanges appear to favor leading wholesalers and call into question the extent of price improvement of retail orders. We also highlight the difficulties of devising effective regulation in the options market. Concerned by the impact of dividend play trades on the orderly functioning of the market, in 2014 the SEC approved a new rule designed to make the strategy impractical,[14] which resulted in much lower trading volumes on cum-dividend dates. However, the recent dramatic increase in options trading by inexperienced retail investors has led to a resurgence of the strategy, with arbitrageurs discovering a way to circumvent the barriers created by the SEC.

2 Closely Related Literature

Our paper is related to the emerging literature exploring retail investor trading in the age of Robinhood. Welch (2022), Barber, Huang, Odean, and Schwartz (2022), Boehmer, Jones, Zhang, and Zhang (2021), Eaton, Green, Roseman, and Wu (2021), and Fedyk (2021) focus on retail investor equity holdings and trading and argue that the new generation of investors differs from retail investors previously examined in the literature (most notably, by Barber and Odean (2001)) along several important dimensions. Although the counts of retail investor equity positions are available from Robintrack, data on their trading in options is

[14]The new rule proposed by the Options Clearing Corporation (OCC) and approved by the SEC could be found here: `https://www.sec.gov/rules/sro/occ/2014/34-73438.pdf`.

not available to researchers. To our knowledge, we are the first to document retail investor preferences and market participation in options, which we infer from transaction-level data that includes newly introduced OPRA trade types.

We are aware of the following papers on retail trading in options. Using account-level data from a brokerage, Bauer, Cosemans, and Eichholtz (2009) document that retail investors' motives for trading appear to be gambling and entertainment and that they incur substantial losses on their options investments. Lakonishok, Lee, Pearson, and Poteshman (2006) argue that speculation is the key driver of retail investors' trading in options and that during the dot-com bubble they favored options on growth stocks. Our paper uses transaction-level data for the entire U.S. options market to document trading patterns of the new generation of retail investors. We show that these investors also have preferences for lotteries and opt for ultra short-term (weekly) options (consistent with preferences for skewness discussed in Barberis and Huang (2008) and Boyer and Vorkink (2014)), participate in trading frenzies, and incur large trading costs (possibly masked by zero-commission offers).

Also related to our work are papers on options market structure and liquidity, for example, Battalio, Griffith, and Van Ness (2021), Ramachandran and Tayal (2021), Muravyev and Pearson (2020), Christoffersen, Goyenko, Jacobs, and Karoui (2018), Battalio, Shkilko, and Van Ness (2016), Muravyev (2016), and Mayhew (2002). None of these papers, however, constructs measures of retail investor trading and, more generally, examines retail investors. The closest to our paper is the contemporaneous work of Ernst and Spatt (2022), who use the same method as ours to identify wholesaler trades in the options market. Their main focus is on the comparison of price improvement (relative to the best prevailing quotes) achieved by wholesalers in equities versus options. Our focus is on the behavior of retail investors in the options market and their performance during the recent retail trading boom, as well as on the behavior of arbitrageurs who exploit retail investor mistakes.

It has been previously documented that not all American options are exercised rationally (e.g., Poteshman and Serbin (2003)). Battalio, Figlewski, and Neal (2020), Cosma, Galluccio, Pederzoli, and Scaillet (2020), Jensen and Pedersen (2016), and Barraclough and Whaley (2012) focus on early exercise decisions and show in more recent data that a fraction of investors still fail to exercise their options optimally. Hao, Kalay, and Mayhew (2010) and Pool, Stoll, and Whaley (2008) show how market makers exploit these mistakes by engaging in dividend play trades. Our measure of arbitrageur activity for the dividend play, based on the new OPRA codes, is more accurate and it allows us to document a surprising reluctance of arbitrageurs to harvest arbitrage profits in certain contracts.

Our findings are related to the literature on investor protection (e.g., Barbu (2022), Bhattacharya, Illanes, and Padi (2019), Egan (2019), and Campbell, Jackson, Madrian, and Tufano (2011)). We show how retail brokers and wholesalers benefit from the growth of retail trading in the options market, potentially more so than from retail trading in equities. Furthermore, retail investors' tendencies to trade options contracts with relatively larger spreads and to forgo profits from early exercise directly translate into larger gains for intermediaries and arbitrageurs. The complexity of options contracts from the viewpoint of an average retail investor and the potentially misaligned incentives of intermediaries call for more research and potential enhancements to investor protection on trading platforms.

Finally, there are related studies highlighting potential limits to competition among market makers, in particular, in equities. Christie and Schultz (1994) show that NASDAQ market makers collude so as to maintain higher bid-ask spreads. This behavior has stopped after publication of that paper. Our paper uncovers a specific friction due to which market markers affiliated with wholesalers may face less competition from other market makers on options exchanges – the fee structure of price improvement mechanisms. We also document a puzzling reluctance of arbitrageurs to engage in dividend play trades that are likely to divert profits from leading option writers.

3 PFOF and rise of retail trading in options market

In this section, we document novel facts about retail trading in the U.S. options market. Leveraging several granular datasets and regulatory filings, we characterize a recent increase in the concentration of retail brokerage markets. We propose a new measure of retail activity in the options market based on transaction-level data, describe its composition and performance, and show how it relates to the existing stock-level retail activity measures and other stock characteristics.

3.1 Data

We use option transaction-level data from OPRA LiveVol provided by CBOE. This data covers all trades on 16 U.S. exchanges in index, ETF, and equity options. In our analysis, we focus on ETF and equity options and exclude index options.[15] Our sample covers November 4, 2019 to June 30, 2021.

Following the literature, we remove canceled trades, trades with nonpositive size or price, with a negative spread (difference between best ask and best bid), and only keep trades for which trade price is above the best bid minus spread and below the best ask plus spread.

[15]Our sample also includes some ADRs. For brevity, we refer to underlying assets as "stocks and ETFs" in the text that follows.

We aggregate trades of the same contract with the same quote time, exchange ID, trade price, and trade condition ID into one line. We do not exclude open or close trades from our analysis, yet we confirm that excluding trades before 9:45 a.m. and after 3:50 p.m. does not change our results. We winsorize trade prices, sizes, and spreads at 99.5$^{\text{th}}$ percentile daily. To compute trade imbalances, we follow the method described in Muravyev (2016), whereby trades with prices above (below) the midpoint are classified as "buy" ("sell") trades and trades at midpoint are classified according to the quote rule on the exchange where the trade took place. We also confirm that our results hold when using a so-called quote rule, that is, when midpoint trades are excluded (shown to have strong performance for options data by Savickas and Wilson (2003)), and Lee and Ready (1991) algorithm (or tick rule to classify trades at midpoint instead of excluding them).[16]

We use daily option price, volume, and open interest data from OptionMetrics. It comes at a contract level for the period between January 4, 1996, and June 30, 2021. We lag open interest for all the data after November 28, 2000, to have a series of consistent open interest as of the end of day.[17]

All stock-level data comes from the Center for Research in Security Prices (CRSP). This includes dividend history, stock prices and returns, and outstanding shares. To link with OptionMetrics, we rely on the SecId-PERMNO crosswalk provided by WRDS.

Our data on retail investor popularity is as follows. We download all comments submitted by users to "Daily Discussion" (DD) and "What Are Your Moves Tomorrow" (MT), most popular daily threads on `WallStreetBets` subreddit of `reddit.com`. The sample spans October 1, 2019 to June 30, 2021, and is collected via PRAW, which is a Python API toolkit to access `reddit.com`. In particular, we download all the comments (original posts and reactions to them) for each daily DD or MT thread.[18] To count ticker mentions in the downloaded comments, we start from the list of unique historical tickers from CRSP and search for them in all the comments, and then simply sum by date. We search only for capitalized tickers, as it is typical for the reddit audience to use those. Since we might omit any lower-case mentions, and we do not cover other threads of the forum (such as occasional megathreads), our measure provides a lower bound for ticker popularity. For Robinhood breadth of ownership, we use Robintrack data, which is provided in intraday snapshots and covers May 5, 2018, to August 13, 2020. We use the number of users holding a stock as of the last intraday snapshot.

[16] The resulting ticker-level imbalances have a correlation over 99% between the quote and Lee-Ready (1991) methods, while the correlation of both of them with the Muravyev (2016) method is 94%.

[17] The lag is due to the change in the reporting format of OptionMetrics. This implies that end-of-day open interest is measured after option exercises.

[18] Few dates are missing due to retrieval limitations on `reddit.com`.

In addition, we rely on FINRA OTC Transparency data to get stock trading volumes executed off lit exchanges, that is, automated trading system (ATS)[19] and non-ATS OTC trades. Pursuant to FINRA's Regulatory Notice 15-48, these are available from April 2016, by security and venue.[20]

Recently revised Rule 606[21] requires broker-dealers to report the aggregate data on PFOF in stocks and options, along with its composition across a number of categories. We download these forms for the largest brokers in the United States directly from their websites. We consider all the leading retail brokerages that rely on wholesalers for PFOF in servicing retail flow. The list of brokers, largest venues, and their corresponding payments for order flow is reported in Table A3 in the Appendix.

3.2 Zero commissions, PFOF, and market structure

The global retail brokerage industry has changed drastically in recent years. More platforms are offering zero-commission trading in equities, and commissions in other asset classes have been reduced as well. Elimination of commissions has fueled a retail participation boom in financial markets, rise in day trading, and gamification of investing.[22] The success of the zero-commission business model relies on PFOF received from intermediaries in exchange for routing retail orders to them for execution. In response to the changing industry landscape and to promote transparency, the SEC introduced new reporting requirements for brokers. In this section, we use the forms filed in compliance with the new rule (Rule 606 reports) to describe the market for PFOF.

Figure 1 plots monthly PFOF received by the U.S. retail brokerages in our sample since the more detailed reporting of PFOF was made compulsory by the SEC. Although only a fraction of retail investors trade options, the amount of PFOF from options exceeds that from stocks by about 100%, in each month in our sample. In 2021, the annual PFOF from options was \$2.4 billion, compared to \$1.3 billion from equities. Our results below help understand why wholesalers offer so much PFOF for options.

Despite recent growth in retail trading and the commercial success of the zero-commission model, the wholesaler market remains quite concentrated, with top five PFOF providers accounting for over 95% of the total PFOF received by U.S. brokerages (see Figure 2). Also apparent from Figure 2 is the high concentration of PFOF providers in options,

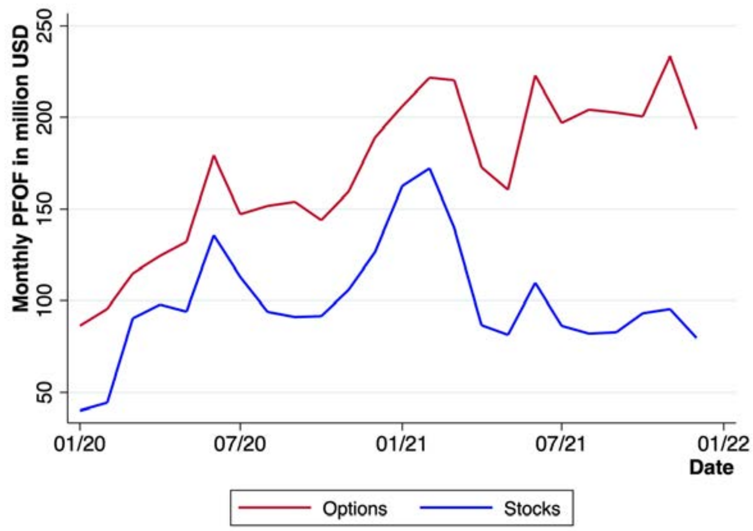
[19] ATS are typically referred to as "dark pools."

[20] Details are on the website of FINRA: https://otctransparency.finra.org/otctransparency/AtsIssueData. For details on the rule, see: https://www.finra.org/rules-guidance/notices/15-48.

[21] For details, see https://www.sec.gov/rules/final/2018/34-84528.pdf

[22] See, e.g., the interview with the SEC chair: https://www.cnbc.com/amp/2022/01/19/secs-gensler-warns-investors-about-frequent-trades-on-brokerage-apps.html.

Figure 1: Payment for order flow: Options vs stocks

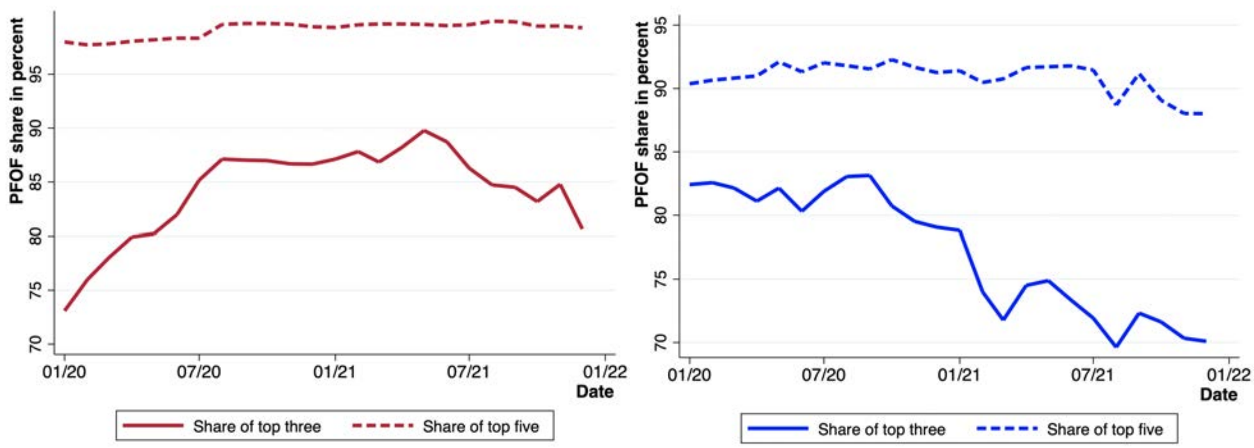


This figure plots aggregate monthly payments for order flow received by U.S. retail brokerages.

Figure 2: Market concentration in PFOF: Options vs. stocks

(a) Options

(b) Stocks



This figure plots the share of PFOF received by U.S. retail brokerages from the top three and top five wholesalers. The top three wholesalers in options are Citadel, Susquehanna, and Wolverine while the top three wholesalers in stocks are Citadel, Virtu, and Susquehanna.

with the share of the top three providers – Citadel, Susquehanna, and Wolverine – increasing from 73% at the start of our sample and reaching an average value of about 85%. It peaked at nearly 90% in the second quarter of 2021. We hereafter refer to Citadel, Susquehanna, and Wolverine as *the Big Three wholesalers* in options.

PFOF also tends to be concentrated in a handful of brokerages, as we show in Appendix A.6. This is, however, to a large extent a reflection of their preferred business model.

For example, Interactive Brokers do not rely on the wholesaler-based PFOF model and send retail orders directly to the exchanges. We therefore exclude them from our sample. On the other hand, TD Ameritrade is by far the largest receiver of payment for order flow in both stocks and options. Interestingly, Robinhood's share of PFOF in options has been steadily increasing from 15% in January 2020, likely due to the attractiveness of the zero-commission trading in options, provided by the company. It peaked at above 30% in January 2021, before stabilizing at about 25% by the end of the year.

3.3 SLIM: A measure of retail trading in options

In this section we propose a new measure of retail trading in options. While recent literature on the ongoing retail investor boom has come up with a number of new retail trading measures, all of them have been focused on equities. These stock-level measures include retail trading imbalances (Boehmer, Jones, Zhang, and Zhang (2021)), breadth of Robinhood user ownership (Welch (2022) and Eaton, Green, Roseman, and Wu (2021)), and counts of `WallStreetBets` ticker mentions (also Eaton, Green, Roseman, and Wu (2021)).[23] Even though these measures are not for options, retail investor frenzies in options and underlying equities are likely to occur at the same time, so we find it useful to include measures of retail investor activity in equities in our dataset.[24]

We add one more measure of retail equity trading to the list: internalized volume, which is the share of non-ATS OTC weekly trading volume in total volume, at a stock level, based on FINRA and CRSP data.[25] FINRA makes public the identities of the largest market makers executing non-ATS OTC transactions. Internalized trades for stocks are executed off lit exchanges, yet not in "dark pools" (which are classified as ATS transactions). The non-ATS OTC transactions consist primarily of internalized order flow from retail and institutional customers of wholesalers. Table A6 in the Appendix ranks market makers by their non-ATS OTC volume share.[26] This ranking closely resembles the one in which we sort wholesalers by their share in PFOF. To the best of our knowledge, this measure has not been used in the extant literature to date. For more details, see Appendix A.8.

The debate on the adverse consequences of internalization by both policy makers and academics so far has focused exclusively on equities. We now argue that, despite the different

[23] This list is based on the most recent measures with wider coverage, and it omits papers using proprietary data, such as NASDAQ TRF data.

[24] We include the latter two measures because we do not have Trade and Quote (TAQ) data required for constructing the measure of Boehmer, Jones, Zhang, and Zhang (2021).

[25] Not all of these trades originate from retail brokerages (FINRA defines it as "non-ATS electronic trading systems and internalized trades"). Yet, our results suggest that a significant fraction of these trades do.

[26] Our list of wholesalers/internalizers is very close to that documented in Eaton, Green, Roseman, and Wu (2021) based on NASDAQ data.

market structures, the patterns are similar, if not more acute, in options.

A highly publicized advantage to investors for having their orders routed to a wholesaler by a retail brokerage in exchange for PFOF, is that the wholesaler promises a price improvement to the customers, that is, the execution price that is at least as good as or better than the best quoted price, known as National Best Bid and Offer, or NBBO. To meet this commitment, wholesalers execute retail orders through *price improvement auctions/mechanisms,* offered by most options exchanges in the United States.

Here is how it works. A retail investor sends an order, which the broker routes to a wholesaler in exchange for PFOF and price improvement. Unlike a stock order, which can be internalized by a wholesaler on its own private trading platform, all options orders in the United States must be executed on exchanges. The wholesaler therefore engages its affiliated market maker to bring a paired order[27] to a price improvement auction on an exchange. Market participants ("responders") have a window of time to respond (by sending a "contra" offer) with a better price (hence, the name "price improvement mechanism"), which could lead to the wholesaler losing the trade. In practice, the fees set by exchanges are stacked against responders, and it is prohibitively expensive to break up many of these paired trades.[28] These responder fees are so high because exchanges also compete for the order flow and incentivize wholesalers to bring orders to them.[29]

Our novel measure of retail trading activity in options is based on trades that went through price improvement auctions. To construct it, we use a dataset from OPRA that includes all options transactions in the United States. We take advantage of a unique feature of our dataset: the new trade type codes introduced by OPRA on November 4, 2019. This classification is significantly more detailed than its predecessors, and hence we can construct our measure starting only from November 4, 2019. Specifically, we use the OPRA transaction code SLAN, which stands for a single-leg price improvement mechanism; we use an acronym SLIM to refer to these trades.[30] In our analysis below, we primarily focus on *SLIM Share*, which could be computed as a frequency share and as a trading volume share. We adopt the

[27]That is, with the affiliated market maker taking the other side.

[28]On most exchanges, order execution by a wholesaler-affiliated market maker gets charged the fee of just \$0.05 per contract. In contrast, it would cost another market maker \$0.50 to break up/respond to one of these already paired orders during an auction. In the latter case, the wholesaler receives a net rebate of \$0.30 per contract just for bringing the order to the exchange. Appendix A.9 contains a detailed description of the fee structure pertaining to price improvement mechanisms on U.S. options exchanges and also highlights the fee advantages enjoyed by affiliated market makers.

[29]To some extent, this is natural, since markets benefit from the presence of largely uninformed retail flow and wholesalers are therefore compensated for delivering these orders. However, the structure and size of the fees associated with servicing retail order flow, that would lead to the optimal level of competition among market makers and efficient order execution, are still an open question.

[30]See Appendix A.2 for a description.

latter definition, as it would be more relevant for assessing the influence of retail traders on the market. We compute it daily and aggregate to a ticker level using traded volumes.[31]

For comparison, we also report a measure of retail trading in options, often used in the media and industry: *Small Share*, the volume share of trades of up to 10 contracts, and the corresponding trading volume in small trades. This measure is noisier than SLIM because in addition to retail trades it contains transactions of proprietary trading firms (e.g., Simplex Trading), which were broken into smaller trades by their order execution algorithms. The frequency share of small trades is 87% in our sample, which overestimates retail investor activity in options.

In Figure 3, we plot our retail trading measure in options, SLIM Share, alongside Small Share. We also plot the total volume of SLIM and small trades. Panels (a) and (c) reveal significant growth of and comovement between SLIM and small trading volumes: Retail investor trading shows a marked increase in our sample. For example, the dollar trading volume in SLIM and small transactions has grown by 104% and 139%, respectively, from January 2020 to July 2021. This is in line with the growth of PFOF for options, which is 158% over the same period, based on monthly data. The growth in retail trading is especially high from January 2020 until March 2021. This period includes several well-publicized retail investor frenzies in equities and a meteoric rise in the number of Robinhood's active users. This increased participation is also reflected in higher average shares, especially in summer 2020, when the average SLIM Share was almost as high as 20%.

Table 1 presents various features of SLIM trades and compares them to average trades in the options market. One striking fact is that retail investors prefer to trade options with the shortest maturities: 50.3% of SLIM trades (in terms of their volume share) are in weekly options, compared with 44.0% for the entire universe of trades. This difference is highly significant, both statistically and economically. The average bid-ask spread in options with less than a week to expiration is a whopping 12.6%. (The effective spread is lower, 6.6%, but it is still orders of magnitude higher than that in equities).

Why do retail investors opt for ultra short-term options? One possible explanation is that options with the shortest maturity are listed as default on trading apps (e.g., they are a default choice on Robinhood).[32] Another explanation is investor preferences for lotteries or gambling. This explanation is consistent with preferences for skewness discussed in Barberis and Huang (2008) and Boyer and Vorkink (2014) and a number of other behavioral biases (e.g., overconfidence, sensation-seeking, and preferences for gambling), summarized

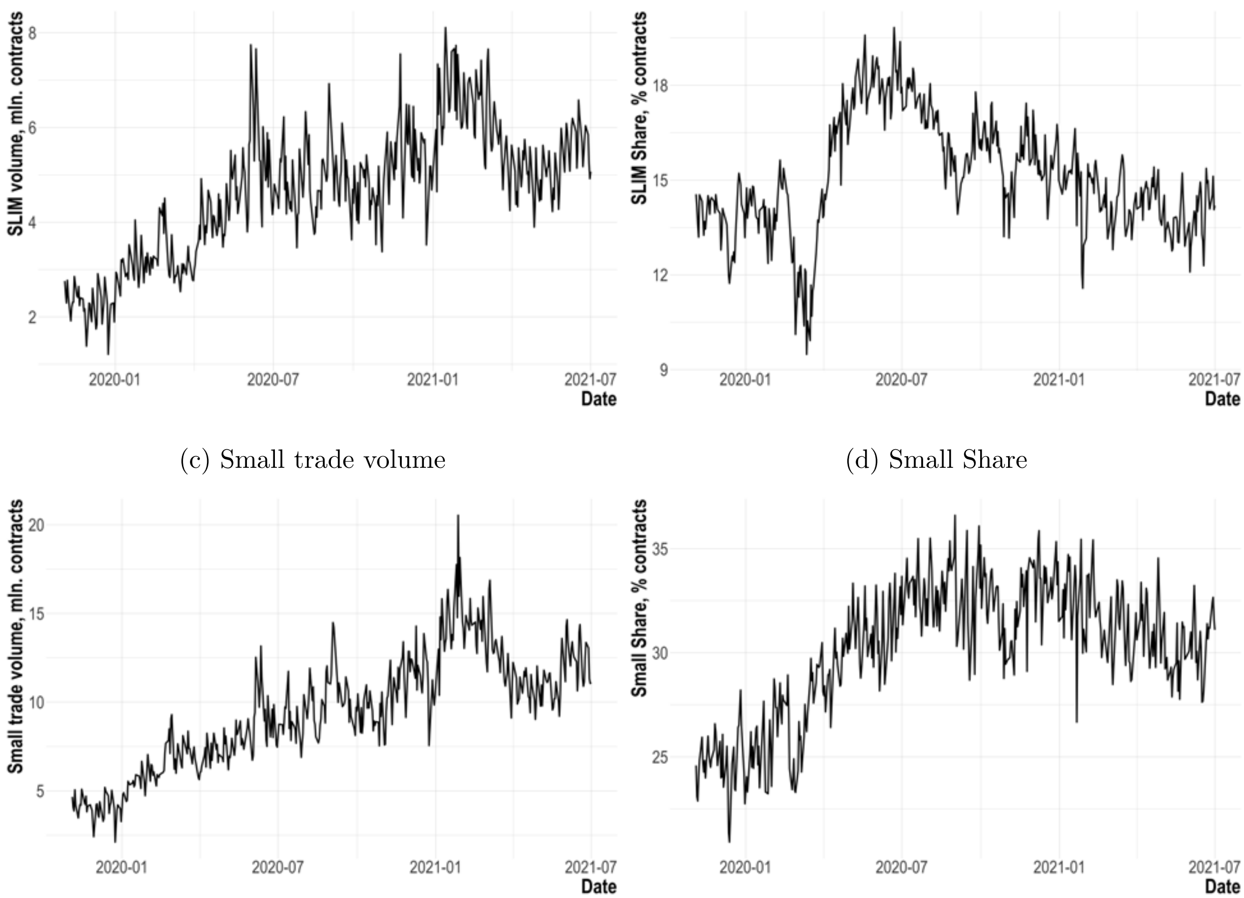
[31]We discuss other measures constructed using SLIM trades, e.g., SLIM Imbalances, later in this section.

[32]Default options often have a significant impact on financial decision-making; see Madrian and Shea (2001), Choi, Laibson, Madrian, and Metrick (2004), Beshears, Choi, Laibson, and Madrian (2009), and Beshears, Choi, Laibson, Madrian, and Skimmyhorn (2022), among others.

Figure 3: Retail investor trading in options

(a) SLIM trade volume (b) SLIM Share

(c) Small trade volume (d) Small Share



This figure characterizes retail investor trading in the U.S. options market between November 2019 and June 2021. Panels (a) and (c) plot total daily trading volumes in SLIM and small trades, respectively. Panels (b) and (d) plot daily SLIM and Small Shares, respectively, averaged across all stocks and ETFs in our sample.

in Table 1 of Liu, Peng, Xiong, and Xiong (2022).[33] Finally, retail investors may simply be cash-constrained. Indeed, weekly options have the lowest prices relative to otherwise identical contracts with longer maturities, so retail investors could opt for the cheapest alternative. At a 12.6% quoted bid-ask spread, the cheapest alternative, however, is by no means cheap to trade. Lured by recent low- or zero-commission offers, retail investors possibly underestimate the indirect trading costs in the options market.[34]

[33] Weekly at-the-money options, favored by retail investors, often have an implied leverage of 58-73. Table A4 in the Appendix reports implied leverage for various option groups. For evidence that retail traders in options are cash-constrained, see Appendix A.10.

[34] The PFOF model and its implications for execution quality and cost transparency have been under scrutiny of regulators for years. See, e.g., the 2021 U.S. Congressional hearing on Robinhood named “Game Stopped? Who Wins and Loses When Short Sellers, Social Media, and Retail Investors Collide.”: `https://`

Table 1: Composition of option trades

Characteristic	Category	SLIM trades Frequency share, %	SLIM trades Volume share, %	SLIM trades Quoted spread, %	SLIM trades Effective spread, %	All trades Frequency share, %	All trades Volume share, %	All trades Quoted spread, %	All trades Effective spread, %
Type	Call	71.5	69.3	13.5	6.6	65.4	62.6	11.4	8.1
	Put	28.5	30.7	14.0	6.9	34.6	37.4	12.8	8.5
Trade size (contracts)	1	45.6	6.2	13.9	6.4	43.4	5.5	11.7	7.9
	2–5	31.0	13.2	12.7	6.2	31.6	12.4	11.5	7.9
	6–10	11.8	14.2	14.1	7.2	11.6	12.6	12.5	8.9
	11–100	11.0	52.6	15.0	8.4	12.3	48.4	12.7	9.6
	Above 100	0.6	13.8	15.0	12.0	1.0	21.2	13.0	10.7
Trade size (dollars)	Below 250	41.2	14.2	23.6	11.7	37.6	12.7	21.0	14.5
	250–500	15.5	8.9	8.7	3.9	15.0	7.7	8.4	5.3
	500–1,000	13.7	11.3	7.4	3.2	14.2	10.1	7.1	4.4
	1,000–2,500	13.8	17.3	6.2	2.6	15.1	16.2	6.0	3.7
	2,500–5,000	7.0	13.5	5.2	2.1	7.9	13.3	5.0	3.1
	5,000–10,000	4.5	13.1	4.5	1.9	4.9	12.4	4.4	2.8
	10,000–20,000	2.5	10.1	3.9	3.2	2.8	10.2	3.8	5.1
	20,000–50,000	1.3	7.7	3.5	6.6	1.7	9.2	3.3	10.2
	Above 50,000	0.5	4.0	3.2	11.9	0.8	8.1	3.0	17.0
Trade direction	Sell	50.1	49.8	14.0	7.1	49.5	48.8	11.0	7.8
	Buy	47.0	47.8	13.0	6.6	47.8	48.9	12.5	9.2
	Midpoint	2.9	2.4	20.2	0.0	2.7	2.3	16.4	0.0
Time to expiration	Less than a week	48.3	50.3	12.6	6.6	43.5	44.0	13.1	9.5
	1–2 weeks	13.9	13.0	12.4	6.0	14.5	13.4	10.3	7.0
	2–4 weeks	15.9	15.1	15.2	7.1	16.9	16.4	11.7	7.5
	1–3 months	13.3	13.4	14.0	6.2	15.0	15.6	10.3	6.5
	3–12 months	7.3	7.1	18.4	7.8	8.2	8.9	11.3	7.9
	Over a year	1.3	1.2	17.7	9.4	1.8	1.8	13.2	11.6
Moneyness	Below -2	0.3	0.2	54.1	28.4	0.3	0.3	48.4	31.6
	-2 to -1	0.3	0.4	50.8	25.6	0.4	0.4	44.9	27.1
	-1 to -0.1	23.4	23.9	28.7	13.9	23.9	24.8	22.3	14.8
	At the money	71.7	71.8	8.7	4.2	70.3	69.9	8.4	5.9
	0.1 to 1	4.0	3.5	8.6	4.8	4.8	4.3	6.2	6.7
	1 to 2	0.2	0.1	9.0	7.7	0.2	0.2	6.9	14.1
	Above 2	0.1	0.1	16.8	11.6	0.1	0.1	12.6	25.2
Trade direction and type	Sell – Call	35.5	34.3	13.7	7.0	32.3	30.4	10.5	7.7
	Sell – Put	14.5	15.5	14.6	7.5	17.2	18.4	11.9	8.1
	Buy – Call	33.9	33.4	12.9	6.6	31.5	30.8	12.1	9.0
	Buy – Put	13.1	14.4	13.0	6.6	16.4	18.1	13.3	9.5
	Midpoint – Call	2.1	1.6	20.8	0.0	1.7	1.4	15.8	0.0
	Midpoint – Put	0.9	0.8	18.6	0.0	1.0	0.9	17.2	0.0
ETF	No	81.3	72.4	14.9	7.2	81.5	70.9	12.6	8.7
	Yes	18.7	27.6	8.4	4.4	18.5	29.1	8.8	6.4
Total		100	100	13.7	6.7	100	100	11.9	8.3

This table reports characteristics of trades by category. Our sample is from November 2019 till June 2021. (Implied) Trade direction is based on whether the trade price is above (buy), below (sell), or at the midpoint. Quoted spread is the spread between the best bid and best ask on the contract (across all exchanges) relative to the midpoint price at the time of the trade. Effective spread is an absolute percentage deviation of the trade price from the midpoint price at the time of the trade, multiplied by 2. For both spreads, we report frequency-weighted averages. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts. The last raw reports frequency-weighted average for the full sample. The overwhelming majority of the reported values for SLIM trades are different from those for non-SLIM trades with the p-value below 1%.

Our analysis reveals that the average quoted bid-ask spread of retail trades across all the maturity buckets is 13.7%, compared to 11.9% for the overall market trades. The former is higher because of the composition of retail trades, which are skewed toward tickers and contracts with higher bid-ask spreads.

Table 1 also reveals that retail investors strongly prefer calls to puts: The volume share in calls is 69%. We see that retail investors trade mostly at-the-money (72% of trades) or slightly-out-of-the-money (24%) options. The latter involves higher trading costs, with

www.nytimes.com/2021/02/19/business/dealbook/robinhood-hearing-congress.html.

average quoted bid-ask spread of 28.7%. Furthermore, 14.2% of retail trades have a "micro" size of up to $250, compared to 12.7% in the whole market, and their average quoted bid-ask spread is 23.6%. These observations suggest that retail investors are entering the options market with an intent to speculate rather than hedge.[35] Furthermore, there is almost perfectly balanced initiation of buy and sell trades in either call or put options. This is consistent with the idea that retail order flow is symmetric and therefore potentially attractive to wholesaler-affiliated market makers who earn profits from executing these trades.[36]

We note that 11.7% of SLIM trades in Table 1 are above $20,000. The literature on retail trading in equities typically considers such large trades to be institutional (starting from Lee and Radhakrishna (2000)). Throughout the Appendix, we show the robustness of all of our subsequent results to using SLIM trades below $20,000 as our proxy for retail trades.[37] We further discuss potential limitations of our measure of retail trading in options in Section 3.3.1.

A natural question to ask is how our measures of retail trading in options behave during retail investor frenzies. To illustrate, Figure 4 plots SLIM and small trade volumes alongside counts of `WallStreetBets` mentions for four "meme" stocks: GameStop, Bed Bath & Beyond, Rocket Companies, and AMC. We should note that our measure of `WallStreetBets` mentions has some missing dates due to the retrieval limitations on `reddit.com`, which appear as gaps in the figure.[38] It is apparent from Figure 4 that both measures adequately capture peaks of `WallStreetBets` mentions of these tickers. In Table 2 below, we establish the cross-sectional relationship between our measure and stock-level retail activity measures formally in a regression framework, for the entire sample.

Having defined our measure of retail activity in the options market, we explore its relationship with the characteristics of both options contracts and their underlying. To do that, we first run the following panel regression, separately for call and put options:[39]

$$SLIM\ Trading_{i,t} = \boldsymbol{\gamma}' X_{i,t} + \boldsymbol{\delta}' C_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}. \tag{1}$$

[35] These observations are consistent with Lakonishok, Lee, Pearson, and Poteshman (2006) and Bauer, Cosemans, and Eichholtz (2009).

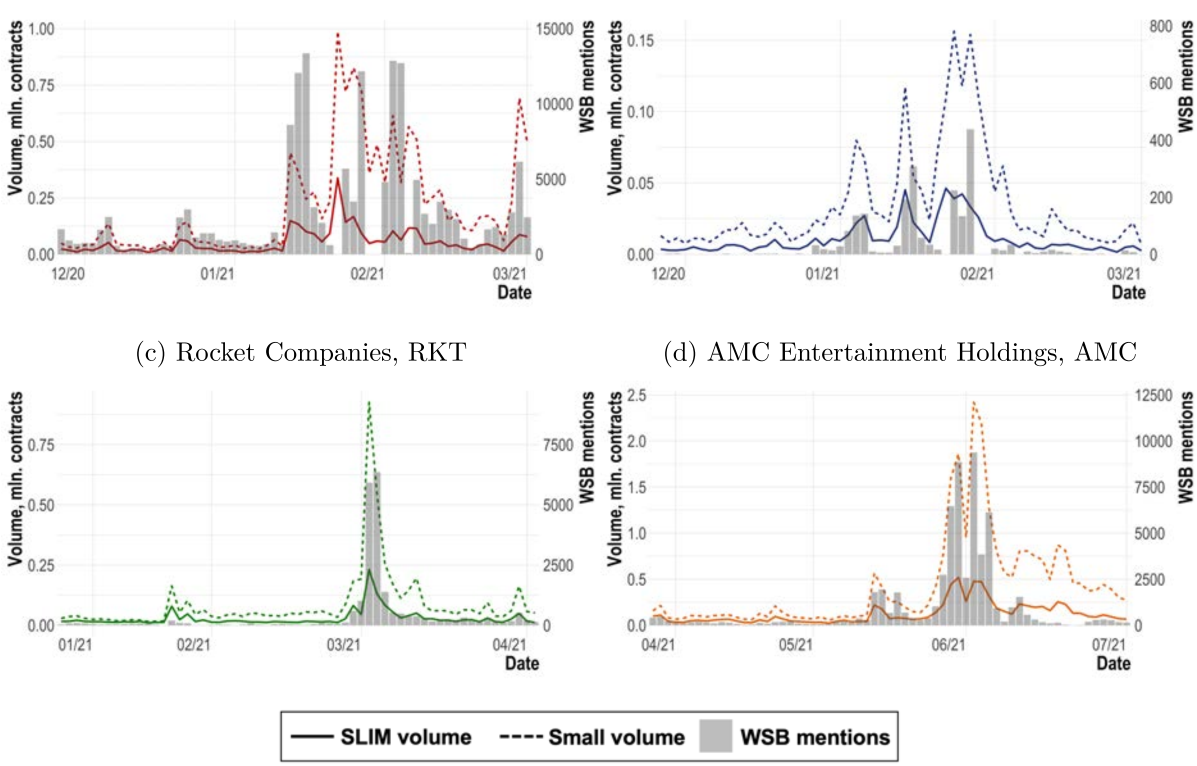
[36] Table A9 in the Appendix shows that our conclusions do not change if we use the fraction of dollar volume in each category instead of frequency or contract volume. All these results are very similar if we use the quote rule to classify trades and exclude open and close trades, as shown in Table A10 in the Appendix.

[37] Table A13 in the Appendix shows the descriptive statistics of trades below $20,000, which are very similar to those without the size filter. Table A14 reports the correlations of SLIM trades below $20,000 with equity-based measures of retail popularity.

[38] These limitations can be circumvented only with real-time scraping of `reddit.com` data.

[39] Splitting the contracts allows us to document differential relationship with the past return on the underlying stock or ETF. All the other results remain similar if we pool both types of contracts together.

Figure 4: "Meme" stocks retail trading and `WallStreetBets` mentions in 2021



(a) GameStop, GME

(b) Bed Bath & Beyond, BBBY

(c) Rocket Companies, RKT

(d) AMC Entertainment Holdings, AMC

This figure plots daily `WallStreetBets` (WSB) mentions (gray bars) and daily volume of SLIM and Small trades.

For call or put contracts of each ticker i on date t separately, we consider two measures for $SLIM\ Trading_{i,t}$. The first one is $SLIM\ Share_{i,t}$, the volume share of SLIM trades among all the options transactions in ticker i on date t, which reflects the general presence of retail investors. The second measure is $SLIM\ Imbalance_{i,t}$, in both calls and puts, which is the volume difference in buy and sell SLIM trades scaled by the total volume of SLIM trades, corresponding to a buy or sell tilt in retail investor trades.

Our vector of characteristics $X_{i,t}$ includes the following ticker-level variables: log dollar trading volume in options on $t-1$, log price on $t-1$, log total trading volume (lit, ATS, and non-ATS OTC) in the underlying stock or ETF over the previous week, relative spread in the underlying averaged over the previous week, volatility of the underlying returns over the previous week, and log market capitalization value as of $t-1$. Our vector of contract characteristics $C_{i,t}$, equal-weighted at ticker i level, includes quoted spread, options moneyness, their time to expiration in months, and leverage.[40] We include ticker and date

[40] Results are not sensitive to whether we use equal-weighting or volume-weighting for contract characteristics at a ticker level. Furthermore, our results are robust to including implied volatility, trade size, delta, and

fixed effects, α_i and μ_t. Finally, we report descriptive statistics for all these variables in Table A11 in the Appendix.

Table 2 presents the results of estimating equation (1). A notable feature of SLIM trades is that retail investor share and order imbalance are higher in the options on the underlying with a larger market capitalization and a higher trading volume in the previous week. The latter is consistent with higher retail participation in attention-grabbing securities. Furthermore, retail investors tend to prefer tickers with lower underlying price (and hence, cheaper options as well). In addition, retail trading is more prevalent in the options on more liquid stocks and ETFs. Earlier studies have documented similar relationships for the stock-level imbalances (see Boehmer, Jones, Zhang, and Zhang (2021) and Welch (2022)).[41]

Notably, we see that SLIM Imbalance in calls is likely to be higher in smaller stocks. However, we also see that our chosen set of characteristics has smaller overall explanatory power for imbalances. It suggests that most of the potential price pressure originated from retail investors in the options market seems to be unrelated to fundamentals. This is consistent with the retail flow being fairly balanced and, hence, attractive to market makers.

How are SLIM Share and SLIM Imbalance related to other measures of retail activity? To answer this question, we run a panel regression similar to that in equation (1) but in addition, consider other measures of retail activity:

$$SLIM\ Trading_{i,t} = \beta Retail_{i,t} + \boldsymbol{\gamma}' X_{i,t} + \boldsymbol{\delta}' C_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{2}$$

where $Retail_{i,t}$ is one of the following measures of retail activity at a ticker level, $share^{small}$ is the volume share of trades up to 10 contracts for ticker i on date t (within call and put options), $Internalized\ volume\ in\ underlying_{i,t}$ is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume of ticker i in the week of date t, *Robinhood ownership breadth,* $log_{i,t}$, is the logarithm of the number of Robinhood users holding the ticker i at the end of date t, and *WSB mentions,* $log_{i,t}$, is the logarithm of the number of times ticker i was mentioned on `WallStreetBets` forum on date t. We use the same set of controls for options contracts ($C_{i,t}$) and their underlying ($X_{i,t}$) as in equation (1).

Table 3 presents the results of estimating equation (2). Our first observation is that the measures of retail trading are positively correlated with both SLIM Share and SLIM Imbalance in the cross-section. This provides further validation of our measure of retail trading in options, with additional supporting evidence presented in Section 3.3.2. However,

other option Greeks, such as theta, vega, and gamma, into the list of contract-level controls.

[41] Both SLIM Share and Imbalance are also correlated with a quasi-Robinhood portfolio, designed to reflect retail-popular tickers. Portfolio weights are based on the previous total trading volume, following the general procedure of Welch (2022). See Table A12 in the Appendix.

Table 2: Retail trading in options and underlying characteristics

	SLIM Share		**SLIM Imbalance**	
	Call	Put	Call	Put
	(1)	(2)	(3)	(4)
Option volume, lagged log	-0.020***	-0.043***	0.040***	0.029***
	(-6.68)	(-17.31)	(13.28)	(9.28)
Underlying price, log	-0.257***	-0.207***	-0.036***	-0.057***
	(-15.42)	(-14.02)	(-3.18)	(-5.55)
Underlying return, past week	-0.005***	0.013***	-0.004***	0.005***
	(-3.87)	(9.84)	(-2.71)	(3.21)
Total volume in underlying, past week log	0.050***	0.042***	0.014***	0.035***
	(8.73)	(8.35)	(2.94)	(6.58)
Underlying spread	-0.028***	-0.012***	-0.017***	-0.013***
	(-7.14)	(-3.19)	(-4.85)	(-3.46)
Underlying volatility, past week	0.000	-0.000	-0.005**	-0.004*
	(0.16)	(-0.02)	(-2.21)	(-1.70)
Market cap, lagged log	0.062**	0.039*	-0.075***	-0.001
	(2.57)	(1.94)	(-4.71)	(-0.08)
Option time to expiration	-0.008***	-0.012***	0.002	-0.001
	(-5.59)	(-9.66)	(1.52)	(-0.86)
Option moneyness	-0.016***	-0.014***	-0.002	0.001
	(-8.70)	(-7.78)	(-1.07)	(0.81)
Option spread	-0.023***	-0.026***	-0.009***	-0.010***
	(-11.76)	(-13.33)	(-3.45)	(-3.68)
Option leverage	0.004**	0.002	0.001	0.001
	(2.04)	(0.88)	(0.30)	(0.46)
Observations	1,436,457	1,248,002	1,106,430	838,604
Adjusted R-squared	0.102	0.077	0.021	0.023

This table reports the results of estimating (1) on daily data from November 2019 till June 2021. SLIM Share is the ticker-level volume shares of SLIM trades. SLIM Imbalance is the ticker-level volume imbalance for SLIM trades. Underlying price (log) is as of the day before. Underlying return is the total return over the last week. Underlying spread is averaged over the previous week. Underlying volatility is return volatility over the previous week. Option spread is the contract quoted relative spread. Option time to expiration (in months), moneyness, spread, and leverage are equal-weighted across trades at a ticker level. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

along with the ticker-level X and C characteristics and fixed effects, they explain only 7%–11% of the total variation in SLIM Share, showing very limited improvement over the explanatory power documented in Table 2.

We note that only `WallStreetBets` mentions seem to exhibit no correlation with SLIM Share, albeit they have a very strong relationship with SLIM Imbalance suggesting that ticker popularity on the investor forum is indeed related to the overall buying pressure in both

Table 3: Retail trading in options and other measures of retail activity

	Retail trading in calls				Retail trading in puts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: SLIM Share								
Small Share	0.057***				0.053***			
	(23.92)				(25.16)			
Internalized volume in underlying		0.025***				0.019***		
		(8.84)				(6.98)		
Robinhood ownership breadth, log			0.032***				0.061***	
			(3.23)				(6.05)	
WSB mentions, log				-0.002				0.002
				(-0.88)				(1.61)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.104	0.102	0.096	0.113	0.079	0.077	0.071	0.084
Panel B: SLIM Imbalance								
Small Imbalance	0.517***				0.516***			
	(258.12)				(226.56)			
Internalized volume in underlying		0.015***				0.004		
		(5.09)				(1.33)		
Robinhood ownership breadth, log			0.042***				0.031***	
			(4.20)				(3.40)	
WSB mentions, log				0.012***				0.009***
				(9.78)				(6.53)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,102,700	1,106,430	436,475	953,691	834,037	838,604	340,258	751,965
Adjusted R-squared	0.184	0.021	0.026	0.020	0.179	0.023	0.025	0.022

This table reports the results of estimating (1) on daily data from November 2019 till June 2021. SLIM and Small Share are the ticker-level volume shares of SLIM and small trades, respectively. SLIM and Small Imbalance are the ticker-level volume imbalance for SLIM and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log, is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

calls and puts, even after conditioning on all the contract and underlying characteristics. The relationship between both SLIM Share and SLIM Imbalance with `WallStreetBets` mentions becomes particularly evident and highly statistically significant if we restrict the sample to micro-trades (of $250 or less), as we show in Appendix A.17. This suggests that micro-trades in options are particularly good in representing the universe of `WallStreetBets` users.

Given that the trading volume in the U.S. options market is highly skewed, one might be concerned that our results hold only for very thinly traded contracts. In Table A16 in the Appendix, we estimate equation (2) for the 341 tickers that constitute the top decile by the total dollar trading volume in our sample. The estimation results are similar to what we document in this section.

3.3.1 Limitations of the measure

Our measure of retail trading has a few limitations. First, 11.7% of the SLIM volume is concentrated in transactions with over $20,000 in value (see Table 1), which is considered a cutoff for retail trades in the related literature on equities (see, e.g., Lee and Radhakrishna (2000)). We therefore exclude trades above this size in our robustness checks. Table A14 in the Appendix confirms that the results are virtually the same. We also do not see a small fraction of retail trades that do not go through a wholesaler and instead are sent directly to exchanges, e.g., those originated by semi-professional traders on Interactive Brokers.

Second, and more importantly, our measure omits some retail trades executed through trade types other than SLIM. These other trades may include non-marketable limit orders. Furthermore, depending on the order attractiveness, some trades may be routed by a wholesaler to an affiliated or other specialist on an exchange, instead of running it through a price improvement mechanism.

Third, our measure omits complex strategies, such as bull spreads, straddles, or butterfly spreads. Complex strategies typically require multi-leg transactions, and, therefore, wholesalers looking for price improvement would usually execute them via multi-leg price improvement auctions, as opposed to single-leg ones. In the OPRA data, these transactions appear as a trade type "MLAN" (multi-leg non-ISO price improvement mechanism) and we refer to them as "MLIM" for consistency. These MLIM trades correspond to about 4% of the total market volume, and they are composed primarily of trades of 'protail' investors – small professional investors and hedge funds – albeit some may be those of retail investors. We have also computed mentions of multi-leg strategies on `WallStreetBets` in our sample period and found that those constitute a very small number relative to the mentions of individual tickers and comments overall. In addition, in Appendix A.19, we report descriptive statistics and cross-sectional correlations of MLIM with the equity-based measures of retail activity. It further demonstrates that these trades are clearly quite different in nature to those going through single-leg actions. Since we want to capture trading of the new generation of retail investors, we are hesitant to include MLIM trades in our analysis.[42]

Despite all these limitations, our measure provides the first comprehensive classification of retail trades in the options market. We are not aware of a reliable approach to estimate type 1/type 2 errors of our method. Nevertheless, we believe that the majority of the trades we capture originated from retail investors, and they can be used as a representative sample to study investor preferences and behavior in the options market. In the next section, we provide more tests, confirming that our approach indeed captures retail trades.

[42] Tables A19 and A20 in the Appendix demonstrate that all multi-leg trades taken together and trades above $50,000 are also clearly different from SLIM trades.

3.3.2 Additional validation

In this section, we offer additional suggestive evidence that our measure captures *retail* trading in the U.S. options market.

First, we exploit the fact that some U.S. retail brokerages handle expiring options on their clients' accounts in a rule-based manner. For example, Robinhood attempts to exercise in-the-money options (if the account has enough buying power) or sells the contract approximately one hour before the market close (if it does not).[43] This gives us a testable prediction for our measure of retail trading in contracts on their expiration day: We expect to see an imbalance in the direction of sell trades in the last one or two trading hours of the day. To test this prediction, we study volume share of buy and sell trades in each trading hour on option expiration day.

On expiration days, as Table A21 in the Appendix reports, there is a significantly larger sell volume share in SLIM trades in the last two hours of the trading day. Notably, this pattern does not emerge on non-expiration days. These features of SLIM trades are consistent with retail brokerages taking an automated action to close retail positions prior to the option's expiration. At the same time, there is no pattern like this for MLIM trades and other multi-leg trades, which are more likely to be institutional. We test these differences more formally in Table A22 and find them to be highly statistically significant.

Second, we study directional order imbalances across trade types during the Robinhood herding events (frenzies) uncovered in Barber, Huang, Odean, and Schwartz (2022).[44] In particular, we estimate equation (2) using a dummy for the Robinhood herding event in ticker i on date t instead of $Retail_{i,t}$. This analysis is performed on a subsample of our data (Novermber 4, 2019 to August 10, 2020) due to availability of Robintrack data with which the investor frenzies are identified.

Table A23 in the Appendix documents higher SLIM Imbalance during Robinhood herding events. We also find that the correlation is the highest for SLIM trades sized below $5,000. Importantly, imbalances in MLIM, all multi-leg, and large trades are not positively related to frenzies. Our results even show negative correlations, suggesting that other types of investors, most likely professional traders or institutions, trade against the retail investors during such events. Overall, we document that during the well-publicized investor frenzies there were directional order imbalances in retail trading in options as well.

The new generation of retail options investors is also more likely to be cash constrained. We look for empirical evidence in support of this around stock splits for micro

[43]See Robinhood's rules here: https://robinhood.com/us/en/support/articles/expiration-exercise-and-assignment/, accessed on March 21, 2022.

[44]We thank Brad Barber for kindly sharing data for this exercise.

SLIM trades, which are more likely to reflect the activity of cash-constrained investors.[45] Note that these events should have no effect on investor positions in the underlying, because trading fractional shares is permitted on most popular investment platforms. However, they may still affect retail options investors because trading fractional options contracts is not permitted during our sample period. We perform a simple event study, reported in Appendix A.10, where we focus on two companies popular with retail investors, Apple (AAPL) and Tesla (TSLA), that executed stock splits on the same day, August 28, 2020. We find that SLIM trading volume on these two names went up significantly relative to a control group of companies popular with retail investors that did not go through a stock split. This formally confirms that SLIM trades, especially of micro sizes, are likely to be originated by cash-constrained investors.

Furthermore, we show that SLIM investors are less likely to exercise their options optimally. The decision to exercise an American option on a dividend-paying asset before maturity involves evaluating the profits from exercise relative to the remaining value of the contract. The latter requires a valuation model, for example, the Black-Scholes model, which novice retail investors are less likely to use. We analyze in-the-money call options, which are optimal to exercise on the last cum-dividend date, and find that a higher SLIM Share is associated with a higher fraction of open interest left unexercised by the ex-dividend date in such options. We also see that there is no such association for other trade types such as MLIM, all multi-leg, and large trades. Table A24 in the Appendix summarizes these results, while the details of the test are reported in Section 3 (see specifically Table 7).

Last but not least, it is reassuring that in an independent contemporaneous work, Ernst and Spatt (2022) rely on the same empirical strategy to classify retail trades in the options market. Their findings are complementary to ours, as they focus on the order execution quality and market microstructure.

3.4 Aggregate performance of retail investors in the U.S. options market

We compute the aggregate retail investor dollar performance over the horizon of h days in the spirit of Barber, Lee, Liu, and Odean (2008):

$$\$Raw\,Perf_h^{SLIM} = \sum_{it} V_{i,t}^{SLIM} \times r_{i,t,t+h},$$

[45] We thank Yang Liu for suggesting this test.

where $V_{i,t}^{SLIM}$ are the net dollar purchases of option contract i corresponding to SLIM trades on day t and $r_{i,t,t+h}$ are the h-day horizon returns on each contract computed as

$$r_{i,t,t+h} = \frac{Close\ midquote_{i,t+h}}{Average\ trade\ price_{i,t}^{SLIM}} - 1.$$

We consider horizons h of one, two, five, and 10 days, as well as until the contract expiration.[46] $Close\,midquote_{i,t+h}$ is the close midquote of contract i on day $t+h$ as reported by OptionMetrics. $Average\,trade\,price_{i,t}^{SLIM}$ is the average trade price of SLIM trades on day $t-1$, which is the average buy price of SLIM trades if $V_{i,t}^{SLIM} > 0$ (retail investors were net buyers of contract i on day t) or the average sell price if $V_{i,t}^{SLIM} < 0$ (retail investors were net sellers of contract i). In the main text of the paper, we report results for equally weighted prices.[47] Furthermore, $r_{i,t,t+h}$ are winsorized at 0.25^{th} and 99.75^{th} percentiles each day.

Table 4 summarizes the performance of retail investor options trades. Under the assumption of a 10-day holding period, retail investors lost \$1.22 billion on their options trades between November 2019 and June 2021. Curiously, retail investor losses were concentrated in at-the-money or slightly in-the-money calls with a very short time to expiration (less than a week).

In Table A26 in the Appendix, we report the overall trade performance by month and day of the week. Retail investor losses are not concentrated in any particular month, while, at the same time, January and February 2021 are the worst months in our sample, corresponding to losses of \$780 and \$337 million, respectively. The same table reveals that, on average, investor performance seems to be lower when the holding period includes weekends. This is especially evident for short-term trades originated on Friday. This could be consistent with the findings of Jones and Shemesh (2018), who demonstrate that options returns are on average lower during the nontrading periods (i.e., primarily weekends).

Table A27 in the Appendix reveals the top and bottom 10 tickers, based on performance of trades originated by retail customers and those of the whole market. Similar to the latter, retail investors on average, realized a gain on such large-cap names as Amazon (AMZN) and Apple (AAPL). Interestingly, however, in contrast to the market, they lost on trading in "meme" stocks, such as GameStop (GME) and AMC Entertainment (AMC).

Our analysis thus far has not taken transaction costs into account. Some of the bro-

[46] Note that at the time of writing, the available OptionMetrics data covered the time period only up to December 31, 2021. Therefore, we are missing performance of the contracts expiring after that date when considering the horizon until expiration.

[47] Results for value-weighted transaction prices are very similar. We report them in Table A25 in the Appendix. Equally weighted prices may be sensitive to outliers, while value-weighted prices might be affected by price impact of large trades. We winsorize trade prices, sizes, and spreads as in our earlier analysis at 99.50^{th} percentile each day.

Table 4: SLIM trade performance, aggregate and by contract characteristics

	SLIM Raw performance, $ billion				
Horizon h	1 day	2 days	5 days	10 days	Expiration
Panel A: All contracts					
	-0.583	-1.083	-1.204	-1.215	-1.339
Panel B: By contract type					
Call	-0.237	-0.751	-0.957	-1.093	-0.975
Put	-0.346	-0.332	-0.247	-0.122	-0.364
Panel C: By moneyness					
Below -2	-0.003	-0.004	-0.003	-0.001	-0.001
-2 to -1	-0.005	-0.004	0.000	0.006	-0.003
-1 to -0.1	0.025	0.069	0.269	0.426	0.394
At the money	0.168	-0.193	-0.490	-0.599	-0.576
0.1 to 1	-0.464	-0.641	-0.666	-0.712	-0.801
1 to 2	-0.156	-0.158	-0.172	-0.181	-0.183
Above 2	-0.147	-0.151	-0.141	-0.152	-0.163
Panel D: By time to expiration					
Less than a week	-0.480	-0.893	-1.264	-1.266	-1.257
1–2 weeks	-0.065	-0.132	-0.182	-0.296	-0.331
2–4 weeks	-0.021	-0.074	-0.071	-0.189	-0.183
1–3 months	0.000	-0.009	0.124	0.214	0.329
3–12 months	0.083	0.111	0.240	0.357	-0.127
Over a year	-0.100	-0.086	-0.052	-0.034	0.230

This table reports the performance of SLIM trades from November 2019 to June 2021. Raw performance at each horizon is computed as explained in Section 3.4. Performance is reported from the perspective of the originating counterparty.

kerages in our sample, such as Robinhood, offer commission-free options trading. However, the majority of brokerages still charge around $0.65 per contract.[48] Using the fraction of PFOF in options paid to Robinhood as the upper bound of their share in the retail options trading, we can therefore estimate the aggregate *direct* transaction costs paid by retail investors. Using 1.93 million contracts as the aggregate SLIM volume and 25% as Robinhood's average share in PFOF for options, the direct transaction costs of retail trades in our sample period amount to $\$0.65 \times 1.93 \times 10^6 \times 0.75 \approx \$941$ million.

Importantly, we also evaluate *indirect* transaction costs at a contract level, aggregated

[48] As of March 2022, TD Ameritrade, Charles Schwab, E*TRADE, and Fidelity all charge $0.65 per contract, according to their websites. Some of the brokers provide commission discounts for frequent traders or for large transactions. However, given the stylized features of retail trading highlighted in Table 1, these discounts are unlikely to have a material impact on our estimates.

across all the contracts. They are computed by summing up the products of effective half-spread and trade size across all SLIM trades in our sample, resulting in $5.2 billion.[49] These costs are not as transparent as brokerage fees and are likely to be overlooked by retail investors. Furthermore, they become revenue for market makers and exchanges executing retail orders (rather than for retail brokerages). These costs are economically large, being five times the direct costs of retail trading, and more than three times larger than the actual trading loss estimate in Table 4. Our calculation approach captures the actual gains and losses of retail trading and does not require any assumptions regarding their opportunity costs. Finally, the overall magnitude of trading costs (relative to the raw trading performance of retail investors) is also consistent with the findings in Barber, Lee, Liu, and Odean (2008) on retail trading in stocks.

One limitation of our data is that some trades might come from multi-leg strategies involving options as well as underlying equities (e.g., a covered call), and we do not observe equity legs of these transactions. However, since the retail investor boom in our sample is largely driven by novice investors, we believe that only a small fraction of them use such sophisticated strategies. Therefore, it should have little impact on our aggregate retail performance estimates.

Finally, results presented in Table A26 in the Appendix also allow us to study whether retail investor performance in the later parts of the sample is better than in the earlier ones, consistent with the research of Seru, Shumway, and Stoffman (2009) and Linnainmaa (2011) on investor learning. We find that, on the contrary, retail investors lost *more* money in the later subsample, especially in January and February of 2021, around the GameStop frenzy. This could happen if retail investors do not learn from their trading experience.[50] A more likely explanation, however, is the changing composition of the investor base. While some of the poor-performing early investors could have exited the sample, it seems that their attrition was more than compensated by the entry of new retail investors in the later months. After all, just in 2021 alone, the account base of Robinhood almost doubled, increasing from 13 to 22.5 million.

[49] To put this number into perspective, the total PFOF in options in our sample is around $2.8 billion.

[50] Prior studies also suggest that investors learn worse after experiencing financial losses, in active trading (relative to observing other people decisions) and when they are emotionally involved in the decision-making. See Kuhnen (2015) and references therein. It would be interesting to extend our data and test these potential mechanisms for the performance of the new generation of retail investors.

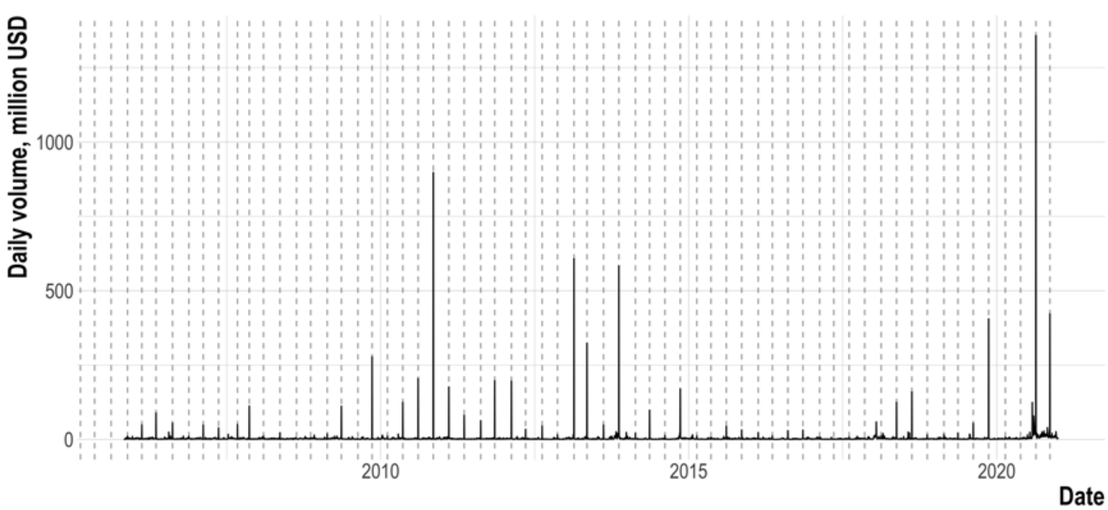
4 Retail investors and option exercise mistakes

Our aim is to study how the inflow of retail investors has affected the behavior of arbitrageurs in the options market. To this end, we focus on a particular arbitrage strategy, a dividend play, in which we can accurately identify trades of arbitrageurs. We present our measure of arbitrageur activity in the dividend play trade and discuss channels through which the inflow of retail investors has made this strategy more profitable for the arbitrageurs.

4.1 Resurgence of dividend play

Daily trading volume in options on high-dividend stocks in the United States exhibits an intriguing seasonality, illustrated in Figure 5 for the case of UPS. The spikes in trading volume, apparent from the figure, occur every quarter, on the last cum-dividend date. The average daily volume of trade in options for UPS is \$125.3 million on cum-dividend dates and \$2.5 million on any other day. This pattern is particularly common for options on high dividend paying stocks; Appendix A.28 presents more examples.

Figure 5: Abnormal trading volume on cum-dividend dates for UPS



This figure plots daily trading volume for all call option contracts on UPS, in millions of U.S. dollars, as reported in OptionMetrics. The dashed lines indicate cum-dividend dates.

On cum-dividend dates, market participants engage in an arbitrage trade known as the dividend play.[51] This strategy is practical only for transactions originating from the floor of the exchange,[52] or, in other words, only to the market participants who must be

[51]See Pool, Stoll, and Whaley (2008) and Hao, Kalay, and Mayhew (2010).

[52]In fact, dividend play could be organized off the exchange floor, but it would then not qualify for transaction

physically located on the trading floor. The strategy involves establishing long and short options positions that are so large that an operational error may potentially destabilize the market. Concerned about the impact of dividend play trades on the orderly functioning of the market, in 2014 the SEC issued a new rule designed to make the strategy impractical (see footnote 14), which curbed this abnormal trading. However, the recent dramatic increase in options trading by inexperienced retail investors appears to have led to a resurgence of the strategy, despite the barriers created by the SEC rule.

Dividend play takes advantage of the specific investor mistake: failing to exercise a call option when it is optimal to do so. It is optimal to exercise a call option, if the value of exercising it on a cum-dividend date and collecting a dividend exceeds the value of the call the next day, when the stock goes ex-dividend. Computing option values, however, involves an application of the Black-Scholes-Merton formula or a more sophisticated option-pricing model, which is typically difficult for a novice retail investor. Alternatively, some retail investors may be unaware of the possibility of early exercise or are simply inattentive.[53] Since a fraction of the in-the-money call options remains suboptimally unexercised, the writers of these options would not be asked to deliver the stock and would earn a windfall gain. It is a zero-sum game. The goal of the dividend play is to divert these gains away from the option writers (or existing holders of the short positions) to an arbitrageur.

When some options contracts are left unexercised, the U.S. Options Clearing Corporation (OCC) randomly assigns short positions that must deliver the stock. The unassigned holders simply hold on to their options and profit from a capital gain. Arbitrageurs can divert this capital gain to themselves by simultaneously buying and selling a large number of in-the-money call options on the same ticker.[54] They exercise all the long positions and deliver on all the assigned short positions. Since the arbitrageurs now account for almost all of the short positions in the call, they receive the windfall gains from the call options not exercised. That is, since some fraction remains (suboptimally) not assigned, they capture dividends on their net long stock positions while staying fully hedged. Usually, two arbitrageurs agree on a dividend play trade in advance and serve as counterparty to each other on their arbitrage positions.

fee caps. In our data, most abnormal volume on cum-dividend dates goes through floor trades on two exchanges, PHLX and BOX, as we discuss below.

[53] There might be other reasons why investors do not exercise, such as the costs of unwinding more complex strategies. Hao, Kalay, and Mayhew (2010) show that dividend play profits outweigh such costs in most cases.

[54] The current SEC rule, presented in footnote 14, makes simultaneous buying and selling of the same contract impractical, due to the new order of position clearing. Yet it does not eliminate a similar trade in the nearby options contracts on the underlying, which would achieve exactly the same goal – artificially inflating the open interest so as to receive most of the assignment.

Table 5: Dividend play: An example

	OI_{t-1} (A)	New positions(t) (B)	Available for ex. (C)	No. exercised (D)	Prob. non-assign. orig. option writer (E)	Prob. non-assign. arbitrageur (F)	Gain per share (G)	Expected gain orig. option writer (E*G*100)	Expected gain arbitrageur (F*G*100)
Case 1. Optimal exercise									
Customer	1	0	1	1	0		5	0	
Case 2. Suboptimal exercise									
Case 2.1. Without dividend play									
Customer	1	0	1	0	1		5	**500**	
Case 2.2. With dividend play									
Customer	1	0	1	0	1/101		5	**5**	
Arbitrageur	0	100	100	100		100/101	5		**495**
Total	1	100	101	100					

This table illustrates the dividend play strategy. Date t refers to the cum-dividend date, and OI_t stands for the open interest on date t.

Table 5 illustrates the mechanics of the dividend play strategy with an example. Suppose there is one call option contract outstanding and it is optimal to exercise it.[55] Case 1 corresponds to the case when the option is exercised. The holder of the short position get assigned to deliver the underlying, so there is no profit for a dividend play strategy to harvest. Case 2 describes what happens if the contract is left unexercised. Without arbitrageur involvement, the short position in the contract does not get assigned, and the option writer receives a windfall gain of $500. Now consider the entry of an arbitrageur. The arbitrageur attempts to pocket most of the potentially harvestable profit of $500. To do so, the arbitrageur buys and simultaneously sells 100 contracts and exercises all their long positions. The probability of assignment increases, but, because of the OCC's random assignment, with probability 100/101, the arbitrageur holds the short position that does not get assigned and hence yields a gain. For the original option writer, this probability is now only 1/101. Hence, the expected gain of the arbitrageur is $495 out of the total gain of $500 and that of the original option writer drops to $5. A dividend play strategy, therefore, dilutes the share of the gain that accrues to the original option writer.

In the next section, we detect dividend play activity at a contract level in the full sample and characterize its importance relative to the overall trading volume on cum-dividend dates.

[55] Appendix A.29 provides another example, in which there are multiple contracts outstanding, some of which are exercised optimally and some not.

Table 6: Characteristics of activity on cum-dividend dates

	Average ticker dollar volume ($ million) on		**Total market dollar volume share (%) on**	
	cum-dividend date	any other date	cum-dividend date	any other date
	(1)	(2)	(3)	(4)
Panel A. Option type				
Call	27.1	1.8	92.1	54.6
Put	2.5	1.7	7.9	45.4
Panel B. Moneyness				
In-the-money	28.2	0.9	78.4	17.8
At-the-money	5.9	2.4	19.6	71.0
Out-of-the-money	0.7	0.4	2.0	11.1
Panel C. Trade size				
Small	1.7	0.8	5.8	26.6
Large	33.6	3.3	94.2	73.4
Panel D. Floor trade				
Yes	51.1	1.0	75.7	5.9
No	7.0	3.0	24.3	94.1
Panel E. Exchange				
PHLX or BOX	26.3	0.6	78.5	14.2
All other	6.3	2.8	21.5	85.8

This table compares option trading activity for dividend-paying tickers (2,153 stocks and ETFs) on cum-dividend date with any other date. The average volume in Columns (1) and (2) is computed at ticker-day level, and the volume share in Columns (3) and (4) is for the entire market. In Panel B, we define "in-the-money" as $(Midpoint\ Price - Strike)/Strike > 0.1$ for call options and $(Midpoint\ Price - Strike)/Strike < -0.1$ for put options. "At-the-money" represents contracts for which this value is between -0.1 and 0.1, and "out-of-the-money" represents all other contracts. In Panel C, we define trade as "small" if the trade size is at or below 10 contracts. In Panel D, we define "floor trades" as trades with SLFT and MLFT OPRA trade types.

4.2 Arbitrageur activity in the dividend play strategy

We first present our measure of arbitrageur activity in the dividend play strategy. Through fee caps, exchanges incentivize cum-dividend day arbitrage strategies to originate from the physical floor. We therefore again exploit OPRA trade types to isolate option transactions that are executed on the floor. The trade types that cover most of the dividend play transactions are SLFT and MLFT, which are single-leg and multi-leg floor trades, respectively (see Appendix A.2 for a more detailed description). Other floor trade types, used infrequently in our sample, are MLCT, MSFL, SLCN, TLFT, and TLFT. To our knowledge, this is the most precise measure of the arbitrageur activity in the dividend play strategy in the literature, which typically uses trading volume on cum-dividend date in excess of the past average volume.

In an effort to reduce operationally risky dividend play trades, in 2014 regulators changed the order of transaction clearing, which made it impractical to buy and sell the same contract for dividend play. As a result, market participants have adjusted their trading strategies and they now implement dividend play through neighboring contracts, which ultimately achieves the same objective. In the data, we see bursts of simultaneous buy and sell activity in neighboring-strike call option contracts, executed normally within several seconds, all coming from the floor. We see no similar bursts of simultaneous buy and sell activity in call option contracts in any other OPRA trade types, which assures us that our measure very accurately captures arbitrageur activity in the dividend play strategy.

Table 6 presents some descriptive statistics of trading activity on cum-dividend vs. any other date for dividend-paying stocks and ETFs. We see an enormous difference in the floor trading volume and volume of large trades on cum-dividend dates relative to other dates. Moreover, on cum-dividend dates we see a colossal spike in volume on two exchanges that cap fees for the dividend play strategy: PHLX and BOX. Splitting the trades by moneyness, we see that the primary increase in volume comes from trading deep-in-the-money calls (which are more likely to be optimal to exercise). This pattern is a signature of the dividend play strategy. The sheer size of the dividend play positions is astonishing, especially after the 2014 OCC/SEC rule intended to clamp down on this strategy.

4.3 Failure to exercise and dividend play profits

In this section, we compute exploitable profits from a dividend play strategy. Some of them come from an increase in the open interest, some from investors' failure to exercise, and some from the value of early exercise of each contract. With an inflow of inexperienced investors in the options market, we expect the first two components to increase. We therefore find it useful to decompose the exploitable profit from a contract into three parts: the (i) open interest, (ii) fraction unexercised, and (iii) early exercise value.

The exploitable dividend play profit on all the interest for each contract is defined as

$$\pi_t = OI_{t-1} \times f_t \times EEV_t, \tag{3}$$

where $t-1$ is the day before the cum-dividend date, OI_{t-1} denotes open interest on that date (measured after all trades, exercises, and assignments on that date), $f_t \equiv OI_t/OI_{t-1}$ is the fraction unexercised, and EEV_t the early exercise value, computed below. Note that the fraction unexercised reflects the fraction of open interest in an option contract that remains outstanding after the cum-dividend date (after all trades, exercises, and assignments on that date). Both EEV_t and f_t are estimated quantities. Open interest as of the day before the cum-dividend day (OI_{t-1}) and fraction not exercised (f_t) are available from OptionMetrics. In rational and frictionless markets, we expect $f_t = 0$, if $EEV > 0$.

The early exercise value is model-based, and we rely on the Black-Scholes-Merton option pricing formula to compute it.[56] Denote the expected ex-dividend price of an option by c_{ex}, its strike by K, and the current (cum-dividend) underlying stock price by S. The expected option ex-dividend price represents the expected time value of the option. *Early exercise value (EEV)* is therefore the difference between the current stock price, strike, and this expected time value of the option: $S - K - c_{ex}$.[57] The details of the computation of c_{ex} are in Appendix A.30.

In the following analyses, we restrict our sample to call options contracts that are optimal to exercise on cum-dates and refer to it as the *dividend play sample*. Further details related to its construction are provided in Appendix A.31, and Table A29 in the Appendix presents the descriptive statistics for our dividend play sample.

How do retail trading trends relate to cum-dividend date exercise rates? To answer this question, we run the following regression:

$$Y_{c,t} = \beta\, share^{SLIM}_{c,t} + \boldsymbol{\gamma}' X_{c,t} + \alpha_{i,t} + \varepsilon_{c,t}, \tag{4}$$

where, for each contract c on cum-date t, we consider two dependent variables, $Y_{c,t}$: fraction of open interest not exercised by ex-dividend date and potential profits from dividend play strategy as defined in equation (3). $share^{SLIM}_{c,t}$ is the average dollar volume share in OPRA trade type SLAN over one trading week before the cum-dividend date t to capture recent interest of retail investors. In some specifications we also use Small Share ($share^{small}_{c,t}$) and ticker-level measures of retail investor popularity such as *Internalized volume in underlying* and *WSB mentions, log*, all computed over one trading week before the cum-dividend date t.[58] These measures are defined underneath equation (2). Our vector of controls $X_{c,t}$ include the following contract-level variables: log OI, EEV, log dollar trading volume, relative spread, implied volatility, moneyness, and days to expiration.[59] Finally, our specification

[56] To make sure our results are robust to the choice of the underlying pricing model, we considered the sample of broad-index ETFs and computed their corresponding option prices with the Merton and Bates models, following Bakshi, Cao, and Chen (1997) and Cosma, Galluccio, Pederzoli, and Scaillet (2020). Options on these ETFs represent over 10% of contracts in our dividend play sample and 55% of potential dividend play profits. All our results go through in that sample and are available upon request.

[57] We follow the definition of Pool, Stoll, and Whaley (2008), which is equivalent to the definition in Hao, Kalay, and Mayhew (2010). The latter uses dividend instead: $Dividend - c_{ex} + S_{ex} - K$.

[58] We have also explored an alternative specification, in which we measure retail trading over two weeks preceding a cum-dividend date. Our results are quantitatively similar.

[59] Since log OI and EEV are components of potential dividend play profits, we do not include them in the specification in Panel B below.

Table 7: Suboptimal exercise and retail investor popularity

	Dividend play profitability feature					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Fraction of OI not exercised, %						
SLIM Share	4.561***	4.515***	5.155***	5.085***	4.718***	4.918***
	(5.40)	(5.36)	(3.84)	(5.46)	(5.58)	(5.52)
Small Share		2.867***				
		(2.87)				
SLIM Share				-3.609**		
× D(expiring within 2 days)				(-2.09)		
D(expiring within 2 days)				-4.552***		
				(-4.21)		
Internalized volume						27.631***
in underlying						(3.11)
WSB mentions, log						0.399**
						(2.33)
Observations	41,737	41,737	13,759	41,737	41,737	40,252
Adjusted R-squared	0.206	0.206	0.286	0.208	0.183	0.183
Panel B. Potential profits, log U.S. dollar						
SLIM Share	1.525***	1.524***	1.581***	1.611***	1.591***	1.640***
	(11.67)	(11.72)	(7.99)	(11.63)	(12.01)	(11.80)
Small Share		0.087				
		(0.43)				
SLIM Share				-0.689**		
× D(expiring within 2 days)				(-2.49)		
D(expiring within 2 days)				1.153***		
				(9.17)		
Internalized volume						2.904**
in underlying						(2.52)
WSB mentions, log						0.035
						(1.52)
Observations	41,737	41,737	13,759	41,737	41,737	40,252
Adjusted R-squared	0.248	0.248	0.254	0.255	0.231	0.230
Sample	All	All	Top EEV tercile	All	All	All
FE	Ticker*Date	Ticker*Date	Ticker*Date	Ticker*Date	Ticker and Date	Ticker and Date
Contract controls	Y	Y	Y	Y	Y	Y
Ticker controls	N	N	N	N	Y	Y

This table reports estimates of equation (4) in our dividend play sample. SLIM Share and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the cum-dividend date. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF, averaged over one trading week before the cum-dividend date. WSB mentions, log, is the logarithm of total mentions of the ticker on `WallStreetBets` forum. In Panel B, contract controls include log dollar trading volume, relative spread, IV, moneyness, days to expiration. In Panel A, they additionally include log OI and EEV. Ticker controls include: underlying price, underlying volatility, underlying relative bid-ask spread, and underlying market cap. S.E. are clustered by ticker and date. Robust t-statistics are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

also includes the ticker by date fixed effects $\alpha_{i,t}$ as we aim to compare contracts within the same ticker yet different SLIM Share.

Panel A of Table 7 reports the results of the regression in (4), with the fraction of open interest unexercised as the outcome variable. We find that there is a strong positive relationship between retail investor trading, as measured by SLIM Share, and the fraction of options that were suboptimally not exercised on the cum-dividend day. This effect is highly significant regardless of whether we also include Small Share into the model or not. A one standard deviation increase in the share of SLIM trades in the contract in the week preceding the cum-date raises the fraction unexercised by about one percentage point, depending on the specification. This result is robust, and the magnitudes of the coefficients of interest do not significantly change as we relax the specification of fixed effects and switch on ticker-level controls instead (see Columns (5)–(6)). Interestingly, the fraction of unexercised options is lower when retail investors bought them only a day or two before the expiration date (see Column (4)). This could be driven by a high level of leverage embedded in these contracts, which makes retail investors particularly attentive to their exercise value.

In Section 3.3 we introduced another measure of retail trading: Share of the internalized volume in the total trade volume in the underlying stock or ETF. Although this is a measure of retail investor trading in the underlying, it is also likely to be correlated with retail trading in the options on that security. Indeed, we find that a one standard deviation increase in the internalized volume increases the fraction of options left unexercised by 4.5 percentage points (Column (6)). Finally, WSB mentions also reflect a positive and significant relationship between retail investor interest and the fraction of options left unexercised.

A possible alternative explanation for our findings is that the failures to exercise the options may be driven by transactions costs that make exercise impractical. To rule out this explanation, we restrict our sample to the most profitable contracts to exercise, the top EEV tercile (Column (3)). We find that the size of the effect goes up significantly relative to our base case, implying that investor mistakes are a more likely driver of our findings.

Could investors hold the call options in our sample as part of a sophisticated strategy, with exercise breaking one of its legs? First, the shares of multi-leg orders and complex trades are not positively related to exercise mistakes (see Table A24 in the Appendix). Furthermore, to engage in such strategies, investors must qualify for a certain level of investment proficiency, required by investing platforms. Although we do see mentions of a number of multi-leg options strategies on `WallStreetBets`, we believe that the new generation of retail investors that drive our results are financial novices and relatively few of them engage in complex options strategies.

Panel B of Table 7 reveals that our main measure of retail activity, SLIM Share, is also positively related to the profitability of the dividend play in retail-popular contracts. A one standard deviation increase in SLIM Share corresponds to around $1,500 higher profit *per contract*. In other words, the higher the retail activity in a contract in a week preceding the cum-dividend date, the more profitable it is for arbitrageurs to engage in dividend play in the contract. Higher profits come from both (i) higher fraction unexercised (documented in Panel (a)) and (ii) higher open interest in the contracts popular with retail investors.

5 Money left on the table: A puzzle

In this section, we show that arbitrageurs engaging in dividend play leave money on the table by failing to capture arbitrage profits in some call option contracts. We then discuss potential reasons for this puzzling behavior.

5.1 Case study

November 11, 2020, was a cum-dividend date for UPS, a high-dividend paying stock. A number of call options on UPS were deeply in-the-money and optimal to exercise. Table 8 focuses on a particular pair of such contracts, both expiring on November 20, 2020. Notice that the November 11 trading volume in Contract 2 exceeds that in Contract 1 by two orders of magnitude. Furthermore, Contract 2 has a very high share of orders from the trading floor on that day, while Contract 1 has zero. This means that arbitrageurs engaging in a dividend play trade entered Contract 2 but not Contract 1.

Here is the core of our puzzle: Why did the arbitrageurs leave money on the table in Contract 1? The contract had a high EEV and a large fraction of options left unexercised. Using equation (3) to compute the arbitrageur's forgone profits from not entering Contract 1, we arrive at $1,945 \times 0.76 \times 0.29 \times 100 \approx \$42,900$, a significant sum of money.[60]

Table 8: Case study of the arbitrageur activity: Two UPS call options on the cum-dividend date

	Strike	EEV	OI (t-1)	Moneyness	Spread	Fraction not exercised	Cum-date volume	Floor share
Contract 1	160	0.29	1,945	3.15	0.045	0.76	45	0.000
Contract 2	155	0.43	2,487	4.62	0.039	0.47	3,255	0.998

Trading costs do not explain market participants' reluctance to trade Contract 1. Exchanges offer daily fee caps for the dividend play strategy at the ticker level, so if arbitrageurs entered Contract 2, they should have also entered Contract 1. Contract bid-ask spreads in Table 4 are also very similar. In the following section, we show that this pattern is a widespread phenomenon of our sample and arbitrageurs leave about 50% of potential profits on the table.

Who is the recipient of these windfall gains? The unexploited profit in Contract 1 accrued to the writer of this contract, which could be a market maker or perhaps a retail investor. The latter is less likely because retail brokerages take an automated action to

[60]Each options contract in our sample is for 100 shares of the underlying stock or ETF.

close short positions that have dividend risk on behalf of their clients.[61] Appendix A.32 presents an excerpt from Robinhood's Terms and Conditions to provide an example of such automated action. It is therefore more likely that the writer of the contract who received the windfall gain was a market maker. The market maker who is a writer of the contract, of course, has no incentive to engage in a dividend play strategy in this contract because this would mean sacrificing his or her own profit. Yet, it is puzzling why other arbitrageurs would not wish to enter Contract 1 and reap associated profits.

Table A5 in the Appendix generalizes this case study and reports forgone profits by ticker for the top 40 underlying stocks and ETFs sorted by the total size of forgone profits in our sample.[62] We also report the number of profitable individual contracts per ticker. The total amount of harvested profits in the top 40 tickers in our sample is around $64 million, whereas the total amount of forgone profit stands at $80 million. In the full sample, these numbers stand at $96 million and $97 million, correspondingly. For a virtually riskless arbitrage strategy, the amount of money left on the table is striking.

Furthermore, Table A5 does not reveal any particular pattern in harvested versus forgone profits: There is a large variation in arbitrageur participation across and within tickers. In what follows, we examine possible explanations for the puzzling reluctance of market participants to harvest arbitrage profits in some contracts.

5.2 Lack of arbitrageur entry in selected contracts

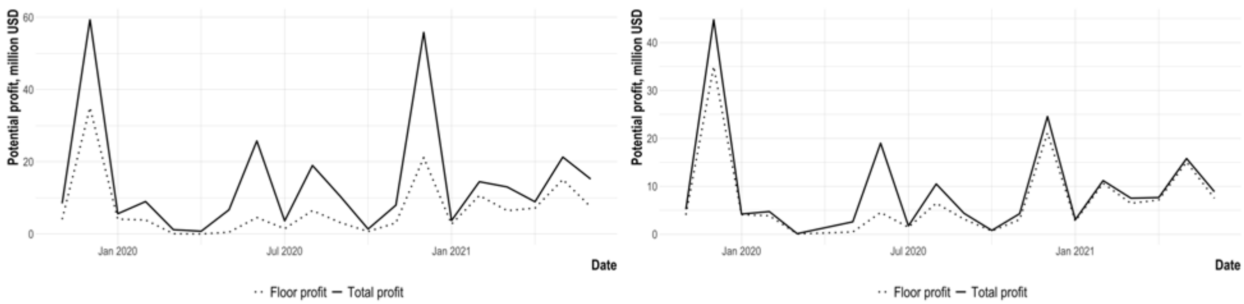
We start by documenting a puzzling feature of the arbitrageur activity, whereby they avoid many profitable contracts altogether. Figure 6 presents potential profit of the dividend play strategy in all outstanding contracts, computed using equation (3), and profits harvested by floor traders. It emerges from Panel (a), that a large fraction of potential profit, over 50%, remains unharvested. If we restrict the sample, however, to the contracts with non-zero floor trading volume—that is, contracts in which we detect dividend play activity—most of the potential profit resulting from the failure of investors to exercise their options on cum-dividend dates is harvested. In other words, arbitrageurs selectively enter profitable contracts, capturing almost 100% of exploitable gains there, but forgo arbitrage opportunities in other profitable contracts.

[61] Since each options contract is for the delivery of 100 shares of the underlying, for small retail investors the cash outlay needed for purchases of the underlying stock and delivering it could be quite significant. A brokerage would therefore close a short position if there are not enough funds in the account to buy and deliver the underlying.

[62] In the absence of direct measure of the open interest held by arbitrageurs, in our calculations of forgone profits we assume that arbitrageurs capture as much as their trading share on the last cum-dividend date in the contract.

Figure 6: Total and floor trader profit from dividend play strategy

(a) All profitable contracts (b) Contracts with non-zero floor share



This figure illustrates the implied share of potential dividend play profits captured by arbitrageurs on the trading floor. The solid plot is for the potential profit from the dividend play strategy, and the dashed plot is for the profit harvested by floor traders (arbitrageurs).

5.3 Drivers of the arbitrageur activity

We now explore the features of the profitable contracts into which arbitrageurs are likely to enter and show how this relates to the retail interest in those contracts.

The total exploitable profit is a calculated quantity, not known for sure on cum-dividend dates. The determinants of potential profit from a contract, however, are well-understood (see equation (3)), and the projected fraction of the suboptimally unexercised call options is one of them. As we know from Table 7, this fraction is increasing in retail investor popularity. We therefore examine whether arbitrageurs exploit increased investor inattention in contracts popular with retail investors. We estimate the following regression in the sample of contracts that should optimally be exercised on cum-date:

$$share^{floor}_{c,t} = \beta\, share^{SLIM}_{c,t} + \boldsymbol{\gamma}' X_{c,t} + \alpha_{i,t} + \varepsilon_{c,t}, \tag{5}$$

where the regressors are as in our previous specification (4) and the outcome variable is now the share of floor trades, which are predominantly dividend play trades in contract c on date t.

Table 9, Panel A, reveals that arbitrageur activity, as measured by floor trading share, is positively related to the retail investor trading over the preceding week. This suggests that arbitrageurs exploit suboptimal options exercise by retail investors, and implement more dividend play trades in the contracts popular among retail investors. A one-standard-deviation increase in SLIM Share increases the share of floor trading by about 0.034*100*0.18≈1 percentage point. These results are statistically significant and consistent for different measures of the floor trading (Columns (4)–(5)).

Table 9: Arbitrageur activity and retail investor popularity

	Floor trading share on cum-date			D(floor share > 0)	Floor trading volume, log
	(1)	(2)	(3)	(4)	(5)
Panel A: Contracts expiring after two days					
SLIM Share	0.034**	0.032**	0.028	0.033**	0.455***
	(2.30)	(2.23)	(1.33)	(2.19)	(3.40)
Small Share		0.076**			
		(2.23)			
Observations	33,564	33,564	10,509	33,564	33,564
Adjusted R-squared	0.416	0.417	0.522	0.412	0.501
Panel B: Contracts expiring within two days					
SLIM Share	-0.050*	-0.049*	0.021	-0.049	-0.280
	(-1.90)	(-1.88)	(0.53)	(-1.63)	(-1.23)
Small Share		0.011			
		(0.36)			
Observations	7,972	7,972	3,178	7,972	7,972
Adjusted R-squared	0.464	0.464	0.482	0.462	0.533
Sample	All	All	Top EEV tercile	All	All
FE	Ticker*Date	Ticker*Date	Ticker*Date	Ticker*Date	Ticker*Date
Contract controls	Y	Y	Y	Y	Y

This tables reports estimates of equation (5) in our dividend play sample described in Section A.31 of the Appendix. Panel A includes contracts expiring after two days from the cum-dividend date. Panel B includes contracts expiring within two days from the cum-dividend date. Floor trading share on cum-date is the contract-level volume share of trades executed on the traded floor in the total traded volume on the cum-dividend date. SLIM Share and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the cum-dividend date. Contract controls include: log OI, EEV, log dollar trading volume, relative spread, IV, moneyness, days to expiration. All regressions include ticker by date fixed effects. t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

If we restrict our sample to the most profitable contracts (Column (3)), the relationship between retail trading and arbitrageur activity weakens and becomes insignificant. This is quite natural to expect: If the early exercise value in a given contract is already high (and hence, other things being equal, so is profitability of dividend play) arbitrageurs are likely to engage in their strategy regardless of the level of retail presence.

To ascertain the robustness of our key result in Panel A, we also pursue an alternative empirical strategy, one based on propensity score matching. Matching is a natural choice in our setup because the set of characteristics on which one should match options to keep the expected profitability constant, is well understood. We report the results of this estimation in Appendix A.33 and they confirm our findings.[63]

[63]In fact, we can also use matching on the same profitability characteristics to illustrate selective entry of

Panel B of Table 9 examines the drivers of arbitrageur activity in the contracts that expire within two days from the cum-dividend date. Intriguingly, in this case arbitrageurs engage *less* with the contracts that are popular among retail investors. This behavior is also puzzling to us. We revisit this feature of the dividend play in the sext section, where we provide some suggestive evidence for an economic channel, consistent with these findings.

5.4 Potential explanations

We now turn to discussing potential explanations for the dividend play puzzle. We start by ruling out the impact of transactions costs and capital/margin constraints.

Transactions costs do not seem to explain the puzzling amount of profits left on the table by arbitrageurs. First, there exist dividend-play-specific fee caps on Philadelphia and, more recently, Boston options exchanges (PHLX and BOX).[64] Those fee caps limit the total costs paid by the market maker on a particular day at the options class level: Harvesting the profit from an additional contract would not increase payments to the exchanges once the limit is reached. Second, given that dividend play trades are usually pre-arranged, it is highly likely that participating parties agree on the transaction price that allows for mutually beneficial profit sharing. There is no clear reason why they would omit any particular contract from their agreement due to its otherwise lower liquidity. Finally, in the above analysis, we always control for contract liquidity or match on contract relative spread. It is therefore unlikely that the contracts in which arbitrageurs do not engage in dividend play are systematically less liquid.

Another potential explanation could be arbitrageurs' capital/margin constraints. However, most regulatory requirements typically involve netted positions, which are relatively low, given the symmetric and fully hedged nature of the strategy. So it is not clear why capital constraints may bind unless they are specific to the arbitrageurs' internal risk management guidelines. Relatedly, such large trades are associated with high operational risks. According to SIFMA, Bank of America Merrill Lynch incurred a $10 million loss due to a human error when executing the dividend play strategy.[65] Still, such explanations cannot produce the variation in floor trader activity within and across tickers that we document.

arbitrageurs even among most similar profitable contracts. In Appendix A.34, we show that across the whole propensity score spectrum, there exist contracts with both zero and positive floor volume in the latter case floor traders represent almost 100% of trading so they seem to exhaust most of the potential profits. This result suggests that profitability characteristics do not predict entry very well, hence emphasizing the puzzle.

[64] See PHLX pricing schedule: `https://listingcenter.nasdaq.com/rulebook/phlx/rules/phlx-options-7` and BOX fee schedule: `https://boxoptions.com/regulatory/fee-schedule/`.

[65] See https://www.reuters.com/article/us-usa-options-apple-idUSKBN0IQ2FA20141106.

Even sophisticated market players exhibit limits to attention (Kacperczyk, van Nieuwerburgh, and Veldkamp (2016)). Indeed, there may be hundreds of potentially profitable contracts available to dividend play on each cum-dividend day (thousands in the case of SPY). Perhaps traders simply cannot evaluate all relevant pricing parameters, enter into an agreement with each other, and process the necessary number of trades? First, it is not clear why other exchange members do not enter to reap arbitrage profits if such limits exist. In undocumented tests, we used the number of stock-level EPS (Hirshleifer, Lim, and Teoh (2009)) and macroeconomic announcements (Savor and Wilson (2014)) as proxies for limits to attention and did not find that those affected floor trader activity.

Some profits could be left unexploited because of the stigma and reputational costs associated with the dividend play strategy. The SEC has clearly signaled its disapproval of the strategy in its 2014 rule aimed at making the strategy impractical (see footnote 14). Reputational costs could also explain the lack of entry of new arbitrageurs. However, they cannot explain why arbitrageurs who regularly engage in this strategy, and hence are willing to incur reputation costs, still leave money on the table.

The only explanation that has some potential in explaining our puzzling empirical results is related to the market concentration among arbitrageurs. Why could market concentration matter for the surprising reluctance of market participants to engage in dividend play in certain contracts? One potential hypothesis is that a large market maker, who receives buy retail call options orders, also serves as a counterparty to these transactions and *writes these options*. Therefore, it is this market maker who would then receive windfall gains from investors' failures to exercise these options. Hence, such market makers have no incentive to engage in dividend play in these contracts. It is puzzling, however, why *other* arbitrageurs do not wish to trade in these contracts and divert windfall gains from the market maker holding a short position in them.

We start by attempting to quantify the number of arbitrageurs simultaneously engaging in a dividend play strategy in a particular contract and show that this number is typically quite small. A signature of the dividend play strategy is the bursts (several trades within milliseconds of one another) of simultaneous buy and sell activity in neighboring-strike call option contracts, originating from the floor. Sizes of trades within each burst are always the same in our sample, but they differ across bursts. We use the number of unique trade sizes executed on the floor as a proxy for the number of arbitrageurs engaging in a dividend play trade in a call option contract. Figure A11 in the Appendix plots a percentage split of dividend play trades by unique trade sizes. The figure reveals that the most common trade size by far is one, which means that there is often only a single pair of arbitrageurs, with one of them entering the long side of the contract. We also observe two or three unique trade sizes, but the occurrence of trade sizes higher than three is fairly rare. This provides suggestive evidence that the number of arbitrageurs entering a given contract is very low.

In Section 3.2, we showed that the market for PFOF in options is highly concentrated: The share of the Big Three—Citadel, Susquehanna, and Wolverine—stands at nearly 90% in the second quarter of 2021. Unfortunately, we do not have actual trader identities in our options dataset, so there is no direct mapping into their share of their overall trading volume in the options market. However, FINRA provides a breakdown by firm of the internalized volume in the underlying *equities and ETFs*, which we summarize in Appendix A.7. Two of the Big Three wholesalers in options, Citadel and Susquehanna, also belong to the top three providers of PFOF in equities in our sample (together with Virtu).[66] Since retail investor frenzies in equities and options tend to occur at the same time, we use the trading volume share in equities internalized by the Big Three as our proxy for their trading volume share in options. Importantly, this data is available at a weekly frequency.

Figure A12 in the Appendix displays the weekly share of internalized trade volume in stocks for a cross-section of tickers. First, over the last few years, internalized volume in equities has become so high that we see stocks and ETFs in which the Big Three had a share in the *total* trading volume of around 15% (see Figure A12, Panel (a)). Those stocks include many usual suspects from the highly publicized retail investor frenzies (ticker and the mean Big Three share in brackets): Rocket Companies (RKT, 16%) and Nvidia Corporation (NVDA, 14%), Pfizer (PFE, 12%) and AstroZeneca (AZN, 11%), Metro-Goldwyn-Mayer Studios (MGM, 16%), AMC Entertainment Holdings (AMC, 11%), and even Apple (AAPL, 13%) and Microsoft (MSFT, 11%).

Table 10 attempts to further shed light on the economic mechanism behind the puzzling decision of the arbitrageurs to leave money on the table. We capture arbitrageur concentration by using the internalized share of volume in the underlying stocks and ETFs by the Big Three (Citadel, Susquehanna, and Wolverine) in the previous week, and create a dummy indicating when the Big Three share is particularly high, over 10% of the overall trading volume in equities. This dummy variable also captures the top decile of tickers with the highest internalization by the Big Three. We report estimation results for a continuous measure of the market concentration in Table A31 in the Appendix.

Table 10, Columns (1) and (2), demonstrate an important role of SLIM order imbalance—that is, more buy relative to sell orders from retail investors and vice versa—in the week preceding the cum-dividend day. SLIM order imbalance has a negative effect on the arbitrageurs' decision to engage in a dividend play strategy. In particular, consistent with our hypothesis that recent retail purchases of a contract indicate that wholesalers are likely to

[66]See Appendix A.3 for a detailed description.

Table 10: Arbitrageur activity and market concentration

	Floor trading share on cum-date			
	(1)	(2)	(3)	(4)
D(SLIM imbalance)	-0.033***		-0.032***	
	(-3.12)		(-2.91)	
D(SLIM buy imbalance)		-0.017***		-0.016***
		(-3.45)		(-2.82)
D(Big Three share > 10%)			0.025	0.011
			(1.16)	(0.54)
D(Big Three share > 10%) × D(SLIM trade imbalance)			-0.030**	
			(-2.35)	
D(Big Three share > 10%) × D(SLIM buy trade imbalance)				-0.018*
				(-1.87)
SLIM Share	0.069***	0.031**	0.065***	0.024*
	(4.54)	(2.31)	(4.09)	(1.71)
Observations	41,737	41,737	41,737	41,737
Adjusted R-squared	0.415	0.415	0.376	0.376
FE	Ticker*Date	Ticker*Date	Ticker and Date	Ticker and Date
Contract controls	Y	Y	Y	Y
Ticker controls	N	N	Y	Y

This table further explains floor trader activity in our dividend play sample. Big Three share is the total share of the Big Three wholesalers' non-ATS OTC volume in the total stock trading volume over the past trading week. $D(SLIM\,imbalance) = 1$ if there was an order imbalance in SLIM trades over the past trading week. $D(SLIM\,buy\,imbalance) = 1$ if there was a positive order imbalance in SLIM trades over the past trading week. Contract controls include: SLIM Share, log OI, EEV, log dollar trading volume, relative spread, IV, moneyness, days to expiration. Ticker controls include: underlying price, underlying volatility, underlying relative bid-ask spread, underlying market cap. t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1

have a short position in it, and hence, would be reluctant to engage in dividend play against themselves. Furthermore, Columns (3) and (4) reveal that the role of SLIM order imbalance is particularly negative, when the Big Three share in the underlying is especially high. That is, arbitrageurs seem more reluctant to enter a contract that has experienced exceptionally high trading volume in the underlying that was internalized by the Big Three.[67]

A possible interpretation of these results follows. Due to a large imbalance in options, the Big Three may have written a large number of call options on that stock/ETF in the preceding week and are holding a large short position in these calls in their inventory on

[67]We also use matching as an alternative approach for studying the effects of arbitrageur concentration in the dividend play strategy. Our estimation results, reported in Appendix A.38, also support the negative effect of the Big Three wholesaler internalization in the underlying stocks/ETFs on the dividend play participation. The set of characteristics on which we match option contracts is the same as in Table A30.

a cum-dividend date. They are then set to benefit from retail investor suboptimal option exercise behavior, documented in Section 4.1, and to collect a windfall gain. Therefore, they have no incentive to engage in dividend play as that would not bring any additional profits. Consistent with this explanation, arbitrageurs are less likely to enter retail-popular contracts, other things being equal, with only up to two days until expiration (see Table 9, Panel B). These extremely short-term contracts are even more likely to stay in the inventory of the original options writers, which lowers their incentive to engage in dividend play. That said, it is still unclear why other arbitrageurs do not enter such contracts.

6 Discussion and policy implications

Our paper calls for more transparency in reporting wholesaler activities in the options market, consistent with the current requirement by FINRA in equities. In particular, it would be useful to know how often market makers affiliated with wholesalers get order allocations through price improvement auctions. The current highly concentrated market is likely to favor leading wholesalers and calls into question the extent of price improvement of retail orders. One particularly fruitful avenue for future research is uncovering the barriers to entry in this market and characterizing the optimal market structure.

We would not be the first ones calling for more transparency in trading costs in zero-commission offers of retail brokerages.[68] However, most prior calls were related to equities. Trading costs in options are orders of magnitude higher, so a regulatory requirement to disclose these costs to investors would be a welcome first step.

Frequent trading produces large order flow and revenue from PFOF for retail investing platforms. Trading less liquid assets, such as options, enhances these profits further. This may create an incentive for retail brokerages to encourage more trading in less liquid asset classes or securities. It would be interesting to explore this issue in follow-up work.

The new generation of investors, while tech-savvy and active on investing forums, could still be lacking in financial education that is required to successfully trade options. For example, retail investors trade options frequently, opting for contracts with very short maturities. This behavior is associated with significant trading costs, which are masked by zero-commission option trading offers by investing platforms (e.g., Robinhood). Another example of retail investor mistakes is that they fail to exercise their options when it is

[68] Regulators have long been interested in various aspects of the system of payment for order flow and, in particular, whether internalization of orders really provides price improvement for the clients. In 2017 SEC found that some of the algorithms used by Citadel Securities to route retail orders, did not seek to obtain the best price on the marketplace, leading to a settlement fee of $22.6 mln (see https://www.sec.gov/news/pressrelease/2017-11.html).

optimal to do so. The question of optimal options exercise requires knowledge of option pricing models, which retail investors are likely to be lacking. One possibility would be to require retail brokerages to report options' early exercise values to investors. The early exercise value could be computed from the Black-Scholes model. Another possibility is to make *automatic* early exercise of calls on cum-dividend dates (when it is optimal to do so) a default option for investors, from which they can opt out if they wish.

Naturally, to better understand retail investor strategies, their potential pitfalls, or discuss investor protection policies, it would be ideal to couple our analysis with account-level data from retail brokerages.

7 Conclusion

This paper focuses on the recent boom in retail investor trading in options. The new generation of retail investors are young and tech-savvy, yet amateur investors. Exploiting a new OPRA reporting requirement, we develop a novel measure of retail investor trading in options and document a rapid rise in retail investor trading in our sample. We argue that retail investors enter the options market for speculative reasons. Lured by recent low- or zero-commission offers, they prefer options with very short maturities, primarily calls. These contracts have high relative bid-ask spreads, making the options business a very lucrative one for wholesalers that execute retail order flow. This is further supported by the ballooning PFOF for options received by retail brokerages.

Retail investors are more likely to make early exercise mistakes. Arbitrageurs exploit these mistakes by engaging in a dividend play strategy, and their profits from it have been boosted by the retail investor boom. This trade is potentially disruptive, as it involves very high trading volumes and carries operational risks. It improves neither market efficiency nor liquidity; it simply redistributes profits from option writers to arbitrageurs. Regulators could act to reduce this trade – for example, by suggesting to abolish fee caps for the dividend play strategy on exchanges. It remains a puzzle to us, however, that while we see a clear signature of this strategy in our data, arbitrageurs leave around 50% of potential profits from it unexploited. This pattern may be a symptom of a bigger problem, which extends beyond the dividend play strategy. Future research may be able to identify it.

References

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *The Journal of Finance* 52, 2003–2049.

Barber, Brad M., Xing Huang, Terrance Odean, and Chris Schwartz, 2022, Attention-Induced Trading and Returns: Evidence from Robinhood Users, *The Journal of Finance,* forthcoming.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2008, Just How Much Do Individual Investors Lose by Trading?, *The Review of Financial Studies* 22, 609–632.

Barber, Brad M., and Terrance Odean, 2001, Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment, *Quarterly Journal of Economics* 116, 261–292.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, *American Economic Review* 98, 2066–2100.

Barbu, Alexandru, 2022, Ex-Post Loss Sharing In Consumer Financial Markets, working paper, London Business School.

Barraclough, Kathryn, and Robert E. Whaley, 2012, Early Exercise of Put Options on Stocks, *The Journal of Finance* 67, 1423–1456.

Battalio, Robert, Stephen Figlewski, and Robert Neal, 2020, Option Investor Rationality Revisited: The Role of Exercise Boundary Violations, *Financial Analyst Journal* 76, 82–99.

Battalio, Robert, Todd Griffith, and Robert Van Ness, 2021, Do (Should) Brokers Route Limit Orders to Options Exchanges That Purchase Order Flow?, *Journal of Financial and Quantitative Analysis* 56, 183–211.

Battalio, Robert, Andriy Shkilko, and Robert Van Ness, 2016, To Pay or Be Paid? The Impact of Taker Fees and Order Flow Inducements on Trading Costs in U.S. Options Markets, *Journal of Financial and Quantitative Analysis* 51, 1637–1662.

Bauer, Rob, Mathijs Cosemans, and Piet Eichholtz, 2009, Option Trading and Individual Investor Performance, *Journal of Banking & Finance* 33, 731–746.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian, 2009, The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States, pp. 167–195 In Brown, J. R., J.B. Liebman, and D.A. Wise (Eds.). Social Security Policy in a Changing Environment.

——— , and William L. Skimmyhorn, 2022, Borrowing to Save? The Impact of Automatic Enrollment on Debt, *Journal of Finance,* forthcoming.

Bhattacharya, Vivek, Gaston Illanes, and Manisha Padi, 2019, Fiduciary Duty and the Market for Financial Advice, working paper, available at SSRN: https://ssrn.com/

abstract=3281345.

Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang, 2021, Tracking Retail Investor Activity, *The Journal of Finance* 76, 2249–2305.

Boyer, Brian H., and Keith Vorkink, 2014, Stock Options as Lotteries, *The Journal of Finance* 69, 1485–1527.

Campbell, John Y., Howell E. Jackson, Brigitte C. Madrian, and Peter Tufano, 2011, Consumer Financial Protection, *Journal of Economic Perspectives* 25, 91–114.

Chapkovski, Philipp, Mariana Khapko, and Marius Zoican, 2021, Does Gamified Trading Stimulate Risk Taking?, working paper, available at SSRN: https://ssrn.com/abstract_id=3971868.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, 2004, *For Better or for Worse: Default Effects and 401(k) Savings Behavior* (Chicago: University of Chicago Press) in David A. Wise, ed., Perspectives on the Economics of Aging.

Christie, William G., and Paul H. Schultz, 1994, Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes?, *The Journal of Finance* 49, 1813–1840.

Christoffersen, Peter, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui, 2018, Illiquidity Premia in the Equity Options Market, *The Review of Financial Studies* 31, 811–851.

Cosma, Antonio, Stefano Galluccio, Paola Pederzoli, and Olivier Scaillet, 2020, Early Exercise Decision in American Options with Dividends, Stochastic Volatility, and Jumps, *Journal of Financial and Quantitative Analysis* 55, 331–356.

Eaton, Gregory W., T. Clifton Green, Brian Roseman, and Yanbin Wu, 2021, Zero-Commission Individual Investors, High Frequency Traders, and Stock Market Quality, working paper, available at SSRN: https://ssrn.com/abstract_id=3776874.

Egan, Mark, 2019, Brokers versus Retail Investors: Conflicting Interests and Dominated Products, *The Journal of Finance* 74, 1217–1260.

Ernst, Thomas, and Chester S. Spatt, 2022, Payment for Order Flow and Asset Choice, *NBER Working Paper 29883*.

Fedyk, Valeria, 2021, This Time is Different: Investing in the Age of Robinhood, working paper, London Business School.

Hagströmer, Björn, 2021, Bias in the Effective Bid-Ask Spread, *Journal of Financial Economics* 142, 314–337.

Hao, Jia, Avner Kalay, and Stewart Mayhew, 2010, Ex-dividend Arbitrage in Option Markets, *The Review of Financial Studies* 23, 271–303.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, 2009, Driven to Distraction: Extraneous Events and Underreaction to Earnings News, *The Journal of Finance* 64, 2289–2325.

Jensen, Mads V., and Lasse H. Pedersen, 2016, Early Option Exercise: Never Say Never, *Journal of Financial Economics* 121, 278–299.

Jones, Christopher S., and Joshua Shemesh, 2018, Option Mispricing around Nontrading Periods, *The Journal of Finance* 73, 861–900.

Kacperczyk, Marcin, Stijn van Nieuwerburgh, and Laura Veldkamp, 2016, A Rational Theory of Mutual Funds' Attention Allocation, *Econometrica* 84, 571–626.

Kalda, Ankit, Benjamin Loos, Alessandro Previtero, and Andreas Hackethal, 2021, Smart(Phone) Investing? A within Investor-time Analysis of New Technologies and Trading Behavior, *NBER Working Paper 28363*.

Kuhnen, Camelia M., 2015, Asymmetric Learning from Financial Information, *The Journal of Finance* 70, 2029–2062.

Lakonishok, Josef, Inmoo Lee, Neil D. Pearson, and Allen M. Poteshman, 2006, Option Market Activity, *The Review of Financial Studies* 20, 813–857.

Lee, Charles M.C., and Balkrishna Radhakrishna, 2000, Inferring Investor Behavior: Evidence From TORQ Data, *Journal of Financial Markets* 3, 83–111.

Lee, Charles M. C., and Mark J. Ready, 1991, Inferring Trade Direction from Intraday Data, *The Journal of Finance* 46, 733–746.

Linnainmaa, Juhani T., 2011, Why Do (Some) Households Trade So Much?, *The Review of Financial Studies* 24, 1630–1666.

Liu, Hongqi, Cameron Peng, Wei A. Xiong, and Wei Xiong, 2022, Taming the Bias Zoo, *Journal of Financial Economics* 143, 716–741.

Madrian, Brigitte C., and Dennis F. Shea, 2001, The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior, *The Quarterly Journal of Economics* 116, 1149–1187.

Mayhew, Stewart, 2002, Competition, Market Structure, and Bid-Ask Spreads in Stock Option Markets, *The Journal of Finance* 57, 931–958.

Muravyev, Dmitriy, 2016, Order Flow and Expected Option Returns, *The Journal of Finance* 71, 673–708.

——— , and Neil D. Pearson, 2020, Options Trading Costs Are Lower Than You Think, *The Review of Financial Studies* 33, 4973–5014.

Pool, Veronika Krepely, Hans R. Stoll, and Robert E. Whaley, 2008, Failure to Exercise Call Options: An Anomaly and a Trading Game, *Journal of Financial Markets* 11, 1–35.

Poteshman, Allen M., and Vitaly Serbin, 2003, Clearly Irrational Financial Market Behavior: Evidence from the Early Exercise of Exchange Traded Stock Options, *The Journal of Finance* 58, 37–70.

Ramachandran, Lakshmi S., and Jitendra Tayal, 2021, Mispricing, Short-Sale Constraints, and the Cross-Section of Option Returns, *Journal of Financial Economics* 141, 297–321.

Savickas, Robert, and Arthur J. Wilson, 2003, On Inferring the Direction of Option Trades, *The Journal of Financial and Quantitative Analysis* 38, 881.

Savor, Pavel, and Mungo Wilson, 2014, Asset Pricing: A Tale of Two Days, *Journal of Financial Economics* 113, 171–201.

Seru, Amit, Tyler Shumway, and Noah Stoffman, 2009, Learning by Trading, *The Review of Financial Studies* 23, 705–739.

Welch, Ivo, 2022, The Wisdom of the Robinhood Crowd, *The Journal of Finance,* forthcoming.

A Internet Appendix

A.1 Share of non-directed orders, by broker

Table A1: Share of non-directed orders by broker

Broker	**Options**	**Stocks**	
		SP500	**Other**
Ally	100.0	100.0	100.0
Apex	97.7	80.7	77.4
Charles	100.0	99.7	99.4
E*TRADE	99.9	99.5	99.1
FIDELITY	88.5	8.1	7.3
Robinhood	99.9	100.0	100.0
TD	99.5	100.0	99.9
Tradestation	99.4	98.2	98.8
Vanguard	100.0	.	.
Virtu	.	95.6	96.4
Webull	100.0	100.0	100.0
tastyworks	100.0	100.0	100.0
Average	**98.6**	**89.2**	**88.9**

This table reports the share of non-directed orders in all orders for each broker in Q1/2020–Q4/2021. Non-directed orders are orders routed to wholesalers and/or exchanges listed in Table A3. All data is from SEC Rule 606 reports.

A.2 OPRA trade types

The table below presents OPRA trade types, together with their descriptions, implemented on November 4, 2019. We also include the corresponding Trade Condition IDs from LiveVol, our data provider.

Table A2: OPRA trade types for transactions in U.S. options exchanges

OPRA Type Description	**OPRA Message Type**	**LiveVol Trade Condition ID**	**OPRA Condition Description**
AUTO		18	Transaction was executed electronically. Prefix appears solely for information; process as a regular transaction.
CANC		40	Transaction previously reported (other than as the last or opening report for the particular option contract) is now to be cancelled.
CBMO	Multi Leg Floor Trade of Proprietary Products	133	Transaction represents execution of a proprietary product non-electronic multi leg order with at least 3 legs. The trade price may be outside the current NBBO.
CNCL		41	Transaction is the last reported for the particular option contract and is now cancelled.
CNCO		42	Transaction was the first one (opening) reported this day for the particular option contract. Although later transactions have been reported, this transaction is now to be cancelled.
CNOL		43	Transaction was the only one reported this day for the particular option contract and is now to be cancelled.
ISOI		95	Transaction was the execution of an order identified as an Intermarket Sweep Order. Process like normal transaction.
LATE		13	Transaction is being reported late, but is in the correct sequence; i.e., no later transactions have been reported for the particular option contract.
MASL	Multi Leg Auction against single leg(s)	125	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period and trades against single leg orders/ quotes. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
MESL	Multi Leg auto-electronic trade against single leg(s)	123	Transaction represents an electronic execution of a multi Leg order traded against single leg orders/ quotes.
MLAT	Multi Leg Auction	120	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period in a complex order book. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
MLET	Multi Leg auto-electronic trade	119	Transaction represents an electronic execution of a multi leg order traded in a complex order book.

continuation on the next page

Table A2: OPRA trade types for transactions in U.S. options exchanges (*cont.*)

MLCT	Multi Leg Cross	121	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross and QCC with two or more options legs.
MLFT	Multi Leg floor trade	122	Transaction represents a non-electronic multi leg order trade executed against other multi-leg order(s) on a trading floor. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.
MSFL	Multi Leg floor trade against single leg(s)	126	Transaction represents a non-electronic multi leg order trade executed on a trading floor against single leg orders/ quotes. Execution of Paired and Non-Paired Auctions on an exchange floor are also included in this category.
OPEN		6	Transaction is a late report of the opening trade and is out of sequence; i.e., other transactions have been reported for the particular option contract.
OPNL		7	Transaction is a late report of the opening trade, but is in the correct sequence; i.e., no other transactions have been reported for the particular option contract.
OSEQ		2	Transaction is being reported late and is out of sequence; i.e., later transactions have been reported for the particular option contract.
REOP		21	Transaction is a reopening of an option contract in which trading has been previously halted. Prefix appears solely for information; process as a regular transaction.
SCLI	Single Leg Cross ISO	117	Transaction was the execution of an Intermarket Sweep electronic order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross.
SLAI	Single Leg Auction ISO	115	Transaction was the execution of an Intermarket Sweep electronic order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism marked as ISO.
SLAN	Single Leg Auction Non ISO	114	Transaction was the execution of an electronic order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Soliciation Mechanism.
SLCN	Single Leg Cross Non ISO	116	Transaction was the execution of an electronic order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross and QCC with a single option leg.

continuation on the next page

Table A2: OPRA trade types for transactions in U.S. options exchanges (*cont.*)

MLCT	Multi Leg Cross	121	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross and QCC with two or more options legs.
SLFT	Single Leg Floor Trade	118	Transaction represents a non-electronic trade executed on a trading floor. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.
TASL	Stock Options Auction against single leg(s)	131	Transaction was the execution of an electronic multi leg stock/options order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period and trades against single leg orders/ quotes. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
TESL	Stock Options auto-electronic trade against single leg(s)	130	Transaction represents an electronic execution of a multi Leg stock/options order traded against single leg orders/ quotes.
TFSL	Stock Options floor trade against single leg(s)	132	Transaction represents a non-electronic multi leg stock/options order trade executed on a trading floor against single leg orders/ quotes. Execution of Paired and Non-Paired Auctions on an exchange floor are also included in this category.
TLAT	Stock Options Auction	124	Transaction was the execution of an electronic multi leg stock/options order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period in a complex order book. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
TLCT	Stock Options Cross	128	Transaction was the execution of an electronic multi leg stock/options order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross.
TLET	Stock Options auto-electronic trade	127	Transaction represents an electronic execution of a multi leg stock/options order traded in a complex order book.
TLFT	Stock Options floor trade	129	Transaction represents a non-electronic multi leg order stock/options trade executed on a trading floor in a Complex order book. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.

This table reports OPRA trade types and their descriptions. The type of each transaction in U.S. options exchanges has to be classified using a type description from the table and reported to OPRA. This reporting requirement was implemented on November 4, 2019.

A.3 Payment for order flow, by broker and firm

Table A3: Payment for order flow: Data description

Firms	Broker											Total paid, $ mln.	Total paid, %
	TD Ameritrade	Robinhood	E*TRADE	Charles Schwab	Webull	Fidelity	tastyworks	Tradestation	Apex	Ally	Vanguard		
Panel A: Stocks													
CITADEL	388.1	215.3	115.2	71.4	56.8	0	1	9.6	10.5	4.8		**872.7**	36.4
SUSQUEHANNA	121.9	81.9	67.5	42.7		0	0.5	0	3.6	3.2		**321.3**	13.4
VIRTU	299.5	140.4	94.9	58.6	22.5	-0.4		22	9.8	3		**650.3**	27.1
WOLVERINE		29.3	0				0			0.1		**29.4**	1.2
DASH					0				0			**0.0**	0.0
MORGAN STANLEY						-0.5						**-0.5**	0.0
TWO SIGMA	94.8	65.5	16.2	8.2	7.1	0		6.8	1			**199.6**	8.3
NASDAQ		0	6.3	0.9	0.1	43.1		0		0		**50.4**	2.1
UBS	80.6		15.7	32.6		-0.1		6.2				**135.0**	5.6
CBOE			11.7	0.4	0	48.2		1.1				**61.4**	2.6
OTHER	8.6	0	6.1	2.4	31.3	-0.7		12.8	15.5	0		**76.0**	3.2
Total received, $ mln.	**993.5**	**532.4**	**333.6**	**217.2**	**117.8**	**89.6**	**1.5**	**58.5**	**40.4**	**11.1**	**0.0**	**2395.6**	
Total received, %	41.5	22.2	13.9	9.1	4.9	3.7	0.1	2.4	1.7	0.5	0.0		
Panel B: Options													
CITADEL	713.4	507.6	185.9	101.6	64.6	93	45	10.3	1.7	6.7	2.6	**1732.4**	42.1
SUSQUEHANNA	516.8	298.9	134.4	100.3	45.2	36.2	21.5	0.6	4.5	6.7	0.6	**1165.7**	28.3
VIRTU												**0.0**	0.0
WOLVERINE	142.6	238.7	69.4	73.4	6.6	44.3	0	9.6	3.7	4.6	0.3	**593.2**	14.4
DASH	125.3		89.2	36.6	37.8	15.4	30	11.7	5.5			**351.5**	8.5
MORGAN STANLEY	76.1	83.7	36.9	26.8		9		8.4				**240.9**	5.9
TWO SIGMA			5.1		0.3				0			**5.4**	0.1
NASDAQ						0						**0.0**	0.0
UBS												**0.0**	0.0
CBOE						0						**0.0**	0.0
OTHER	2.1		0.9	3.4	0	6.6	7.4	0.9	2.4	0	0.6	**24.3**	0.6
Total received, $ mln.	**1576.3**	**1128.9**	**521.8**	**342.1**	**154.5**	**204.5**	**103.9**	**41.5**	**17.8**	**18.0**	**4.1**	**4113.4**	
Total received, %	38.3	27.4	12.7	8.3	3.8	5.0	2.5	1.0	0.4	0.4	0.1		

This table reports the total payment for order flow in stocks (Panel A) and options (Panel B) for each broker-firm pair in Q1/2020–Q4/2021. Relationships with missing values do not exist. PFOFs with zero values are rounded to zero. Negative values indicate fees paid. All data is from SEC Rule 606 reports. NASDAQ and CBOE represent exchanges within NASDAQ and CBOE groups, respectively.

A.4 Embedded leverage in options trades

Table A4: Embedded leverage by moneyness and maturity bin

	Frequency-weighted within a bin					Volume-weighted within a bin				
	Time to expiration					**Time to expiration**				
Moneyness	Less than a week (1)	1-2 weeks (2)	2-4 weeks (3)	1-3 months (4)	Above 3 months (5)	Less than a week (6)	1-2 weeks (7)	2-4 weeks (8)	1-3 months (9)	Above 3 months (10)
Panel A: SLIM trades										
Below -2	12.0	4.8	3.3	3.3	2.9	36.3	8.5	3.7	3.8	2.9
-2 to -1	6.4	5.6	4.9	5.1	3.8	7.0	6.3	6.2	6.2	4.2
-1 to -0.1	19.8	13.4	10.2	8.6	6.2	23.1	15.3	11.6	9.6	6.3
At the money	57.9	22.4	15.6	11.0	5.8	83.4	25.5	17.8	12.7	6.7
0.1 to 1	5.5	4.5	3.9	3.6	2.5	5.5	4.3	3.7	3.4	2.7
1 to 2	2.0	2.1	2.8	2.1	2.4	2.4	5.1	9.8	2.8	5.5
Above 2	25.8	20.2	13.5	14.0	13.6	60.5	58.7	32.4	31.6	34.0
Panel B: SLIM trades below $250										
Below -2	12.9	5.0	3.5	3.6	3.1	41.0	9.4	3.8	4.1	3.0
-2 to -1	7.0	6.1	5.3	5.5	4.3	7.5	6.8	6.7	6.6	4.7
-1 to -0.1	21.5	14.2	10.9	9.2	7.6	24.8	16.0	12.2	10.1	7.3
At the money	71.5	25.1	17.7	13.3	7.6	93.2	28.3	20.4	15.6	10.3
0.1 to 1	5.3	3.9	3.3	2.6	1.5	5.4	3.9	3.4	2.8	3.9
1 to 2	2.3	2.9	5.4	3.3	6.7	3.4	12.5	30.5	6.2	29.4
Above 2	69.4	41.9	36.9	46.2	60.1	153.1	110.2	81.4	95.7	134.4
Panel C: All trades										
Below -2	11.3	4.5	3.3	2.8	2.3	15.8	5.8	3.6	3.3	2.4
-2 to -1	6.3	5.3	4.6	4.5	3.3	7.0	6.0	5.7	5.5	3.8
-1 to -0.1	21.0	14.7	11.7	9.3	5.4	24.0	16.7	13.3	10.4	5.9
At the money	62.8	23.3	16.8	11.8	6.0	93.2	27.5	20.2	14.1	7.7
0.1 to 1	5.6	4.7	4.1	3.7	2.7	5.6	4.6	4.0	3.7	2.8
1 to 2	2.7	2.8	2.7	2.5	3.8	3.7	10.8	5.2	3.8	10.6
Above 2	27.8	24.1	14.6	17.6	14.5	53.5	61.3	29.2	45.9	38.1

This table reports the average embedded leverage of options by their moneyness and maturity bin at the time of the trade. Panel A reports averages for SLIM trades only, Panel B – for SLIM trades below $250 in value, and Panel C – for all options trades. Leverage is computed as $|\Delta \times S/p|$, where Δ is the option's delta at the time of the trade, S is the underlying midpoint price at the time of the trade, and p is the option's trade price. All the values are first wighted either by frequency (Columns (1)–(5)) or trading volume (Columns (6)–(10)) within a corresponding bin and then frequency-weighted across time. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts.

A.5 Dividend play profits by ticker

Table A5: Dividend play profits by ticker

Ticker	Profit, USD Harvested	Forgone	No. contracts Fully harvested	Partly harvested	Forgone	Traded volume (contracts)
Ticker	(1)	(2)	(3)	(4)	(5)	(6)
SPY	7,554,679.0	31,400,000.0	478	95	1737	5,811,339
AAPL	4,484,501.0	12,400,000.0	434	303	465	1,842,771
EEM	11,000,000.0	4,938,097.0	246	9	51	6,056,682
IWM	1,484,352.0	3,551,356.0	54	5	218	521,360
XLE	1,827,032.0	2,841,853.0	233	12	112	577,320
VALE	1,944,431.0	2,755,483.0	93	10	29	2,022,030
QQQ	107,597.0	2,054,353.0	24	2	274	29,950
EFA	4,525,486.0	1,973,938.0	202	8	39	2,573,722
EWZ	3,446,402.0	1,332,894.0	157	2	52	1,596,842
KO	421,466.6	1,141,860.0	92	37	110	351,700
DIA	370,758.0	1,050,798.0	225	67	460	145,725
HYG	958,013.9	961,677.6	22	5	92	372,660
XOM	8,332,642.0	825,996.9	270	179	195	2,168,205
SAN	-	753,507.4	0	0	12	-
HD	983,639.4	714,949.4	143	45	190	305,970
COST	10,337.7	681,858.1	47	19	86	15,878
IBM	1,009,330.0	674,709.7	213	228	176	637,022
QCOM	850,742.1	639,122.6	131	52	144	559,989
XLF	494,604.8	621,714.4	71	8	65	569,965
BHP	175,916.1	553,665.3	45	6	14	60,570
ET	688,937.0	503,247.8	61	47	60	736,770
AVGO	1,038,727.0	486,157.6	238	112	121	418,627
GOLD	330,311.0	480,721.0	36	7	52	73,789
T	3,071,579.0	477,051.0	194	76	103	2,790,898
JPM	997,002.3	472,951.6	120	60	134	1,188,446
VIAC	1,780,227.0	437,255.1	114	7	82	471,970
DIS	856,860.6	423,720.4	67	19	20	529,346
WMT	326,273.2	402,605.6	97	40	100	258,917
XLI	10,878.1	401,909.4	13	1	28	17,423
INTC	176,308.0	397,908.8	145	66	147	394,345
GILD	369,587.1	379,696.8	95	46	116	308,665
RIO	27,355.7	378,665.9	18	9	14	73,662
BP	456,491.3	375,388.2	148	87	130	425,279
XLP	137,822.7	371,263.9	27	2	52	58,136
NVDA	-	339,488.3	0	0	107	-
CVX	635,446.6	335,764.3	313	131	183	632,813
FXI	882,075.1	315,215.0	103	6	45	1,371,416
MRO	-	307,658.0	0	0	26	-
PFE	512,518.9	290,129.6	137	97	97	701,674
UPS	1,247,195.0	272,550.6	288	60	149	776,732
Total	**63,527,526.2**	**79,717,182.3**	**5,394**	**1,965**	**6,287**	**37,448,608**

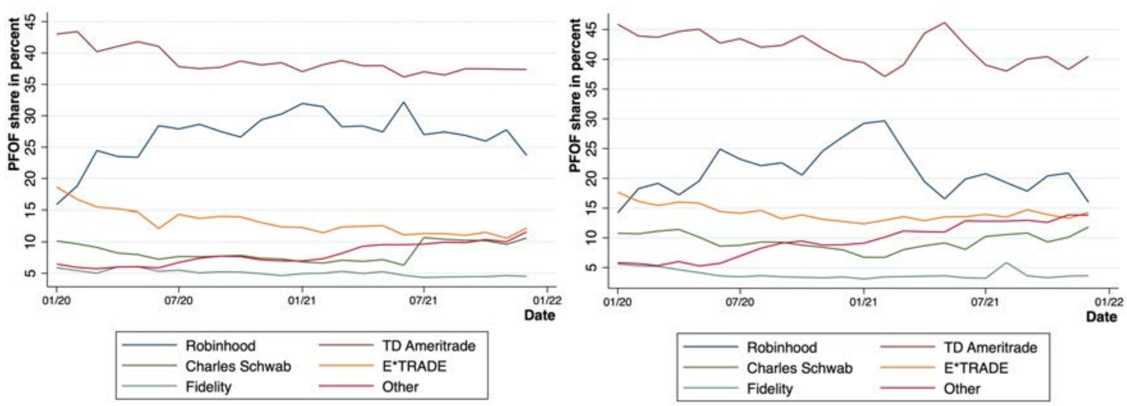
This table reports the top 40 tickers in terms of dividend play profits forgone by floor traders in our sample. Values are aggregated across all contracts within a ticker from November 2019 till June 2021. Total dividend play potential profits are computed as in equation (3) (we do not winsorize components or the resulting profits). To compute harvested profits, we multiply the total profits by the floor volume share on cum-dividend date, and attribute the residual to forgone profits. No. of fully harvested contracts in Column (3) is the number of contracts with floor share above 90%, and in Column (5) – with zero floor share.[a] Traded volume in Column (6) is the total floor trading volume in all contracts.

[a]The average floor share is 99% in fully harvested contracts and 64% in partly harvested contracts.

A.6 PFOF trends, by broker

Figure A1: Share in the payment for order flow: Options versus stocks, by brokerage

(a) Options (b) Stocks



This figure plots the share in monthly payments for order flow of the largest U.S. retail brokerages.

A.7 OTC trading volume, by venue

Table A6: Top 15 internalizers in the United States

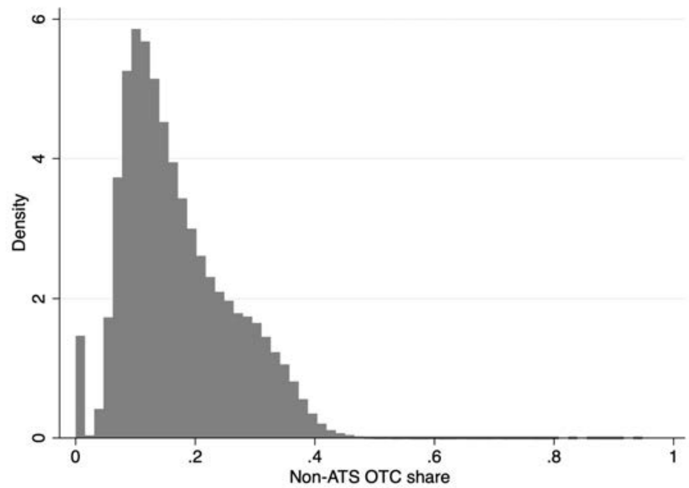
Firm	OTC volume, billion shares	Venue share in total volume, %	Cumulative share, %
CITADEL SECURITIES	477.82	44.31	44.31
VIRTU	357.61	33.16	77.47
SUSQUEHANNA	119.10	11.04	88.52
TWO SIGMA	48.50	4.50	93.01
JANE STREET CAPITAL	28.49	2.64	95.66
UBS	25.35	2.35	98.01
WOLVERINE	7.29	0.68	98.68
COMHAR CAPITAL MARKETS	3.84	0.36	99.04
HRT EXECUTION SERVICES	3.46	0.32	99.36
LEK SECURITIES CORPORATION	2.27	0.21	99.57
GOLDMAN	2.20	0.20	99.77
ACS EXECUTION SERVICES	0.44	0.04	99.81
IMC	0.32	0.03	99.84
MORGAN STANLEY	0.29	0.03	99.87
COWEN	0.28	0.03	99.90

This table reports the top 15 firms in terms of their total OTC non-ATS (i.e., internalized) stock volume between November 2019 and June 2021. It us based on FINRA OTC Transparency data.

A.8 A measure of internalized volume in equities

Figure A2 plots a histogram of weekly non-ATS OTC trading volume (internalized volume) as a share of the total weekly stock trading volume. The average share of internalized volume in the total one is 17% in our sample, and it is trending upwards.

Figure A2: Histogram of non-ATS OTC share



This figure plots the share of non-ATS OTC volume in the total trading volume for all equities and ETFs with options traded in the U.S. between November 2019 and June 2021.

A.9 Price improvement mechanism fees, by exchange

In this section, we summarize fees related to price improvement mechanisms across the U.S. exchanges. In particular, we consider two scenarios. The first one is when a customer order is paired in an auction and the wholesaler-affiliated market maker trades gets the full allocation (i.e., the auction not broken). In the second scenario, a customer order is paired in an auction and an unaffiliated market maker trades in full (i.e., the auction is broken as an unaffiliated market maker provides a better price).

Table A7: PIM-related exchange fees across the U.S. exchanges

				1. Customer order is paired in an auction and wholesaler trades in full					**2. Customer order is paired in an auction and unaffiliated market maker trades in full**				
Exchange Code	Full Name	PIM Name	SLAN trade volume share, %	Customer Exchange Fee/Rebate	Break-up credit	Affiliated market maker	Non-affiliated market maker (responder fee)	Exchange	Customer Exchange Fee/Rebate	Break-up credit	Affiliated market maker	Non-affiliated market maker (responder fee)	Exchange
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Penny program securities													
PHLX	Philadelphia Stock Exch.	PIXL	31.00	(0.17)	NA	0.05	NA	0.12	(0.17)	(0.25)†	NA	0.25	(0.08)
CBOE	Chicago Board Options Exch.	AIM / C-AIM	21.70	(0.14)	NA	0.07	NA	0.07	(0.14)	(0.25)	NA	0.50	(0.11)
AMEX	American Stock Exch.	CUBE	15.50	(0.12)	NA	0.05	NA	0.07	(0.12)	(0.30)	NA	0.50	(0.08)
MIAX	MIAX Options Exch.	MIAX PRIME	12.00	(0.11)	NA	0.05	NA	0.06	(0.11)	(0.25)	NA	0.50	(0.14)
BOX	Boston Stock Exch.	PIP	6.80	(0.11)	NA	0.05	NA	0.06	(0.11)	(0.34)	NA	0.50	(0.05)
EDGX	Direct Edge X	AIM	5.10	(0.06)	NA	0.05	NA	0.01	(0.06)	(0.25)	NA	0.50	(0.19)
MRX	ISE Mercury	PIM	4.10	-	NA	0.02	NA	(0.02)	-	(0.25)	NA	0.50	(0.25)
ISE	International Securities Exch.	PIM	2.60	(0.02)	NA	0.10	NA	(0.08)	(0.02)	(0.15)	NA	0.50	(0.33)
GEMX	ISE Gemini	PIM	1.10	-	NA	0.05	NA	(0.05)	-	-	NA	0.05	(0.05)
NASDBX	NASDAQ OMX BX Options	PRISM	0.06	-	NA	0.05	NA	(0.05)	-	(0.35)	NA	0.49	(0.14)
EMLD	MIAX Emerald Options Exch.	Emerald PRIME	0.05	(0.10)	NA	0.05	NA	0.05	(0.10)	(0.53)	NA	0.55	0.08
Panel B: Non-penny program securities													
PHLX	Philadelphia Stock Exch.	PIXL	31.00	(0.17)	NA	0.05	NA	0.12	(0.17)	(0.70)	NA	0.40	(0.23)
CBOE	Chicago Board Options Exch.	AIM / C-AIM	21.70	(0.14)	NA	0.07	NA	0.07	(0.14)	(0.60)	NA	1.05	(0.31)
AMEX	American Stock Exch.	CUBE	15.50	(0.12)	NA	0.05	NA	0.07	(0.12)	(0.70)	NA	1.05	(0.23)
MIAX	MIAX Options Exch.	MIAX PRIME	12.00	(0.11)	NA	0.05	NA	0.06	(0.11)	(0.60)	NA	1.10	(0.39)
BOX	Boston Stock Exch.	PIP	6.80	(0.11)	NA	0.05	NA	0.06	(0.11)	(0.81)	NA	1.15	(0.23)
EDGX	Direct Edge X	AIM	5.10	(0.06)	NA	0.05	NA	0.01	(0.06)	(0.60)	NA	1.05	(0.39)
MRX	ISE Mercury	PIM	4.10	-	NA	0.02	NA	(0.02)	-	(0.60)	NA	1.10	(0.50)
ISE	International Securities Exch.	PIM	2.60	(0.02)	NA	0.10	NA	(0.08)	(0.02)	(0.15)	NA	1.10	(0.93)
GEMX	ISE Gemini	PIM	1.10	-	NA	0.05	NA	(0.05)	-	-	NA	0.05	(0.05)
NASDBX	NASDAQ OMX BX Options	PRISM	0.06	-	NA	0.05	NA	(0.05)	-	(0.70)	NA	0.94	(0.24)
EMLD	MIAX Emerald Options Exch.	Emerald PRIME	0.05	(0.10)	NA	0.05	NA	0.05	(0.10)	(1.05)	NA	1.10	0.05

This table reports the exchange fees related to price improvement mechanisms (PIM) on all U.S. options exchanges where this mechanisms are used, as of May 10, 2022. Panel A reports fees for securities in the penny program, and Panel B for those not in the penny program. Columns (5)–(9) report fees in a scenario when customer order is paired in an auction and the wholesaler trades in full. Columns (10)–(14) report fees in a scenario when customer order is paired in an auction and an unaffiliated market maker trades in full. Negative values indicate rebates. Rebates typically vary by volume tier, and we report the highest rebate. These fees and rebates are for the majority of underlying securities (they do not include securities with special fees such as SPY). All values are in $ per contract. † signifies break-up credit fees that we could not locate within the corresponding exchange fee schedule, yet its value has been reported by an active market maker participating in PIM.

A.10 Are retail investors in the US options market cash-constrained?

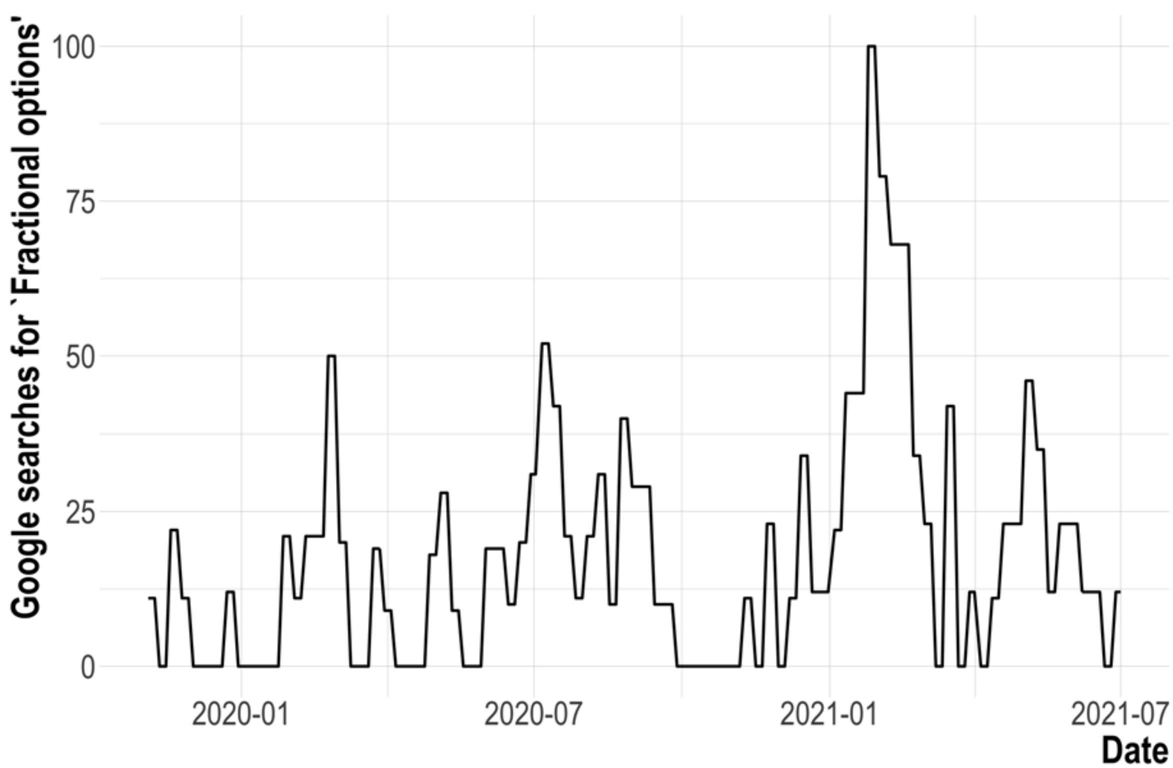
In this section, we present suggestive evidence for binding cash constraints for retail investors in the US options market.

First of all, we see that, during retail frenzies, Google users are more likely to search for "fractional options." Trading fractional options is not permitted in the U.S. in our sample, yet it could allow constrained investors to trade in contracts on an underlying with a high price. Figure A3 plots Google searches for fractional options in our sample. It demonstrates that people are more actively searching for this phrase during the periods of retail frenzies, that is, in June-July 2020 and January 2021.

Second, we see that stock splits on retail-popular yet expensive underlying stocks are associated with an increase in the retail trading volume in options. Figure A4 shows that the average daily volume in SLIM trades below \$250[69] in Apple (AAPL) and Tesla (TSLA)

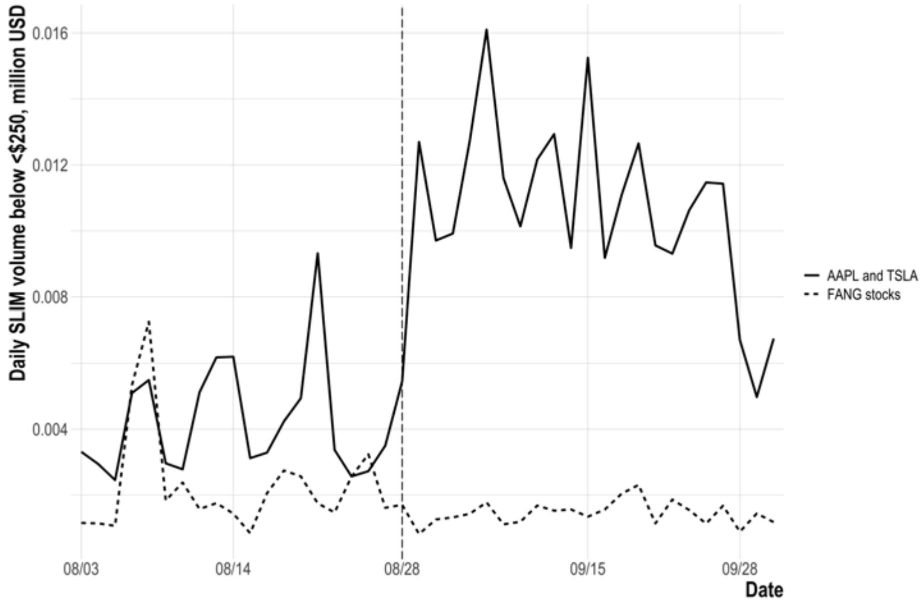
[69] We focus on SLIM trades below \$250 as this measure most likely reflects retail investors who are cash-constrained.

Figure A3: Google searches for fractional options



This figure plots weekly Google searches for fractional options between November 2019 and June 2021. Data source is Google Trends (see https://trends.google.com/trends/), accessed on May 8, 2022.

Figure A4: SLIM trading volume around stock splits



This figure plots the dollar volume in SLIM trades below $250 in size two months around August 28, 2020, when AAPL and TSLA had stock splits (4:1 and 5:1, respectively). The solid line is the average daily SLIM volume of AAPL and TSLA, while the dashed line is the average of FANG companies (Facebook, Amazon, Netflix, and Alphabet). The vertical dashed line indicates the day of the split.

has risen sharply right after their stock splits (both on August 28, 2020) while SLIM trading in FANG stocks, equally popular among retail investors, remained roughly the same.

Finally, we investigate whether a change in SLIM trading is related to stock splits for all underlying securities that had a split in our sample period. Specifically, we estimate the following cross-sectional regression:

$$Y_i = \beta\, Split\, ratio_i + \boldsymbol{\gamma}' X_i + \varepsilon_i. \tag{6}$$

Y_i is one of the following measures of trading activity change around the split of shares in company i: Δ *SLIM volume (contracts)* is a log difference between the daily average number of contracts in SLIM trades below $250 one month after the split and one month before the split, Δ *SLIM volume (USD)* is the same for the average daily dollar volume, Δ *SLIM freq. share* is the difference between the average daily frequency share of SLIM trades below $250 in the total options trading volume one month after the split and one month before the split, and *Internalized volume in equities* is the difference between the average share of non-ATS OTC volume in the total underlying volume one month after the split and one month before the split. X_i are controls related to the underlying stock or ETF, all averaged over one month before the split: price, volatility, return, volume (log), and market capitalization (log).

If retail investors are cash-constrained, we expect their activity to increase more when the constraint is becoming less binding. Consistent with this hypothesis, Table A8 reveals that retail trading in options tends to increase more when the split ratio is higher. This is true for all the measures we consider: contract volume change, dollar volume change, or the change in the share in the total option trading volume for that underlying. Furthermore, this effect is large both statistically and economically, as the size of the split ratio explains 35%–40% of the variation in SLIM volume around the event date. Notably, the change in internalized volume in equities is not sensitive to the split ratio size, which is consistent with the availability of trading fractional shares the United States.

Table A8: SLIM trading activity and split ratio

	Δ SLIM volume (contracts)	Δ SLIM volume (USD)	Δ SLIM freq. share	Internalized volume in equities
	(1)	(2)	(3)	(4)
Panel A: Without controls				
Split ratio	0.214***	0.181***	0.008***	0.002
	(6.84)	(5.43)	(2.66)	(1.04)
Observations	75	75	75	75
Adjusted R-squared	0.390	0.351	0.130	0.005
Panel B: With controls				
Split ratio	0.198***	0.136***	0.008*	0.002
	(6.63)	(3.43)	(1.86)	(0.99)
Observations	75	75	75	75
Adjusted R-squared	0.399	0.428	0.144	-0.027

This table reports estimates of equation (6) on daily data from November 2019 till June 2021. Controls in Panel B include average underlying price, average underlying volatility, average underlying return, average underlying volume (log), and average underlying market capitalization (log), all computed over one month before the split. Heteroscedasticity-robust t-statistics are in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

A.11 Additional descriptive statistics on SLIM trades

Table A9: Composition of SLIM trades, additional statistics

Characteristic	**Category**	Dollar volume share, %	Dollar spread, $	Implied volatility	Trade price, $
Type	Call	70.25	0.22	0.85	4.87
	Put	29.75	0.21	0.80	4.48
Trade size (contracts)	1	13.32	0.25	0.85	5.62
	2-5	22.31	0.21	0.83	4.70
	6-10	18.37	0.17	0.83	3.66
	11-100	39.71	0.13	0.81	2.75
	Above 100	6.28	0.06	0.68	1.31
Trade size (dollars)	Below 250	2.03	0.08	0.94	0.74
	250-500	2.76	0.15	0.81	2.29
	500-1,000	4.84	0.21	0.79	3.71
	1,000-2,500	10.78	0.30	0.76	6.47
	2,500-5,000	12.16	0.44	0.73	11.47
	5,000-10,000	15.67	0.54	0.71	16.68
	10,000-20,000	16.75	0.66	0.69	22.78
	20,000-50,000	19.53	0.82	0.68	29.18
	Above 50,000	15.47	1.13	0.68	42.10
Trade direction	Sell	51.11	0.23	0.83	4.96
	Buy	46.97	0.20	0.85	4.67
	Midpoint	1.92	0.15	0.69	2.74
Time to expiration	Less than a week	40.39	0.17	0.89	4.06
	1-2 weeks	12.29	0.18	0.84	4.41
	2-4 weeks	14.39	0.21	0.85	4.26
	1-3 months	16.60	0.25	0.73	5.63
	3-12 months	12.43	0.40	0.69	7.84
	Over a year	3.89	0.85	0.60	14.16
Moneyness	Below -2	0.06	0.21	2.48	0.93
	-2 to -1	0.11	0.27	1.92	1.40
	-1 to -0.1	10.90	0.18	1.25	2.08
	At the money	77.19	0.20	0.67	5.16
	0.1 to 1	10.90	0.55	1.17	12.96
	1 to 2	0.57	0.80	1.51	18.25
	Above 2	0.28	0.87	1.61	17.99
Trade direction and type	Sell - Call	35.95	0.23	0.84	5.11
	Sell - Put	15.16	0.22	0.80	4.60
	Buy - Call	32.95	0.21	0.87	4.74
	Buy - Put	14.02	0.19	0.81	4.50
	Midpoint - Call	1.35	0.16	0.70	2.85
	Midpoint - Put	0.57	0.13	0.68	2.46
ETF	No	79.02	0.24	0.92	5.18
	Yes	20.98	0.09	0.46	2.91

This table reports characteristics of trades by category. Our sample is from November 2019 to June 2021. (Implied) Trade direction is based on whether the trade price is above (buy), below (sell), or at the midpoint. Dollar spread, $ is the spread between the best bid and best ask on the contract (across all exchanges) in U.S. dollars at the time of the trade. Implied volatility is trade-implied volatility reported by OPRA. For all measures, we report frequency-weighted averages. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts.

A.12 Descriptive statistics on SLIM trades, without open and close trades

Table A10: Composition of option trades

Characteristic	Category	SLIM trades: Frequency share, %	SLIM trades: Volume share, %	SLIM trades: Quoted spread, %	SLIM trades: Effective spread, %	All trades: Frequency share, %	All trades: Volume share, %	All trades: Quoted spread, %	All trades: Effective spread, %
Type	Call	71.2	69.0	13.3	6.5	65.0	62.3	10.8	7.7
	Put	28.8	31.0	13.7	6.7	35.0	37.7	11.9	8.0
Trade size (contracts)	1	45.6	6.2	13.7	6.3	44.4	6.0	10.9	7.4
	2-5	30.9	13.3	12.4	6.1	31.6	13.3	10.8	7.5
	6-10	11.8	14.5	13.8	7.1	11.4	13.3	12.0	8.4
	11-100	11.0	53.1	14.6	8.2	11.6	48.7	12.3	9.1
	Above 100	0.6	12.9	14.6	11.6	0.9	18.8	12.6	10.2
Trade size (dollars)	Below 250	41.4	14.3	23.4	11.6	37.9	13.3	19.9	13.7
	250-500	15.4	8.9	8.3	3.7	15.0	8.0	7.7	4.8
	500-1,000	13.6	11.3	7.0	3.0	14.2	10.3	6.5	4.0
	1,000-2,500	13.8	17.3	5.9	2.5	15.0	16.4	5.5	3.3
	2,500-5,000	7.0	13.5	4.9	2.0	7.8	13.3	4.6	2.8
	5,000-10,000	4.6	13.1	4.2	1.8	4.9	12.2	4.0	2.5
	10,000-20,000	2.5	10.1	3.7	2.9	2.7	10.0	3.5	4.8
	20,000-50,000	1.3	7.6	3.3	6.2	1.6	8.9	3.1	10.4
	Above 50,000	0.5	3.9	3.0	11.4	0.8	7.7	2.9	17.6
Trade direction	Sell	46.2	46.4	13.6	7.6	47.5	47.2	10.1	7.6
	Buy	43.4	44.8	13.0	7.1	45.8	46.9	12.0	9.1
	Midpoint	10.4	8.8	14.2	0.0	6.7	6.0	13.3	0.0
Time to expiration	Less than a week	47.2	49.6	12.3	6.5	42.4	42.8	12.5	9.1
	1-2 weeks	13.8	12.8	12.2	6.0	14.4	13.2	9.7	6.6
	2-4 weeks	16.0	15.2	14.9	7.0	17.1	16.5	10.8	7.0
	1-3 months	13.6	13.6	13.7	6.1	15.4	15.9	9.5	6.0
	3-12 months	7.6	7.3	18.3	7.7	8.6	9.4	10.5	7.3
	Over a year	1.4	1.3	17.4	9.1	1.9	1.9	12.2	10.9
Moneyness	Below -2	0.3	0.3	53.3	27.9	0.3	0.3	46.2	30.1
	-2 to -1	0.4	0.4	50.2	25.3	0.4	0.4	42.8	25.7
	-1 to -0.1	23.8	24.1	28.4	13.7	24.2	25.2	21.1	14.0
	At the money	71.2	71.5	8.4	4.1	70.0	69.4	7.8	5.5
	0.1 to 1	4.1	3.6	8.2	4.6	4.8	4.4	5.7	6.4
	1 to 2	0.2	0.1	8.6	7.4	0.2	0.2	6.3	13.8
	Above 2	0.1	0.1	16.3	11.2	0.1	0.1	11.6	24.8
Trade direction and type	Sell - Call	32.7	31.8	13.4	7.4	30.8	29.3	9.6	7.5
	Sell - Put	13.5	14.5	14.3	8.0	16.7	17.9	10.9	7.8
	Buy - Call	31.2	31.2	13.0	7.1	29.9	29.4	11.6	8.9
	Buy - Put	12.2	13.6	13.1	7.1	15.9	17.5	12.7	9.3
	Midpoint - Call	7.3	6.0	14.5	0.0	4.2	3.6	12.9	0.0
	Midpoint - Put	3.1	2.9	13.4	0.0	2.5	2.3	14.0	0.0
ETF	No	81.4	72.4	14.6	7.1	81.6	71.3	11.8	8.2
	Yes	18.6	27.6	8.3	4.4	18.4	28.7	8.4	6.0

This table reports characteristics of trades by category. Our sample is from November 2019 to June 202. It is limited to trades after 9:45 a.m. and before 3:50 p.m., and trades are classified using the quote method. (Implied) Trade direction is based on whether the trade price is above (buy), below (sell), or at the midpoint. Quoted spread is the spread between the best bid and best ask on the contract (across all exchanges) relative to the midpoint price at the time of the trade. Effective spread is an absolute percentage deviation of the trade price from the midpoint price at the time of the trade, multiplied by 2. For both spreads, we report frequency-weighted averages. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts. The overwhelming majority of the reported values for SLIM trades are different from those for non-SLIM trades with the p-value below 1%.

A.13 Descriptive statistics for the ticker-level sample

Table A11: Descriptive statistics for the ticker-level variables

	Call options					**Put options**				
	Mean	Median	St. Dev.	p1	p99	Mean	Median	St. Dev.	p1	p99
SLIM Share	0.20	0.13	0.23	0.00	1.00	0.16	0.06	0.24	0.00	1.00
SLIM < \$250 Share	0.06	0.01	0.15	0.00	1.00	0.06	0.01	0.16	0.00	1.00
SLIM < \$5k Share	0.17	0.09	0.22	0.00	1.00	0.14	0.04	0.23	0.00	1.00
SLIM < \$20k Share	0.19	0.12	0.23	0.00	1.00	0.15	0.06	0.24	0.00	1.00
SLIM > \$20k Share	0.01	0.00	0.04	0.00	0.16	0.00	0.00	0.04	0.00	0.12
Small Share	0.59	0.53	0.34	0.00	1.00	0.63	0.63	0.36	0.00	1.00
MLIM Share	0.02	0.00	0.08	0.00	0.44	0.04	0.00	0.11	0.00	0.67
Complex Share	0.11	0.02	0.20	0.00	0.97	0.16	0.03	0.25	0.00	1.00
Large Share	0.06	0.00	0.15	0.00	0.72	0.06	0.00	0.15	0.00	0.79
> \$50k Share	0.02	0.00	0.08	0.00	0.43	0.02	0.00	0.08	0.00	0.44
SLIM Imbalance	-0.11	-0.11	0.65	-1.00	1.00	-0.17	-0.23	0.70	-1.00	1.00
SLIM < \$250 Imbalance	-0.17	-0.19	0.65	-1.00	1.00	-0.22	-0.31	0.70	-1.00	1.00
SLIM < \$5k Imbalance	-0.12	-0.12	0.64	-1.00	1.00	-0.18	-0.23	0.70	-1.00	1.00
SLIM < \$20k Imbalance	-0.11	-0.11	0.64	-1.00	1.00	-0.18	-0.23	0.70	-1.00	1.00
SLIM > \$20k Imbalance	-0.04	-0.04	0.80	-1.00	1.00	-0.08	-0.13	0.83	-1.00	1.00
Small Imbalance	-0.05	-0.04	0.52	-1.00	1.00	-0.03	-0.02	0.58	-1.00	1.00
MLIM Imbalance	-0.08	0.00	0.51	-1.00	1.00	-0.11	-0.03	0.54	-1.00	1.00
Complex Imbalance	-0.04	0.00	0.47	-1.00	1.00	-0.06	0.00	0.51	-1.00	1.00
Large Imbalance	-0.03	0.00	0.73	-1.00	1.00	-0.05	-0.01	0.75	-1.00	1.00
> \$50k Imbalance	-0.01	0.00	0.74	-1.00	1.00	-0.05	-0.04	0.77	-1.00	1.00
Internalized volume in underlying	0.17	0.15	0.09	0.00	0.39	0.17	0.15	0.08	0.00	0.38
Robinhood ownership breadth, log	6.90	6.80	1.76	3.30	11.78	7.02	6.91	1.78	3.33	11.93
WSB mentions, log	0.18	0.00	0.56	0.00	3.22	0.19	0.00	0.58	0.00	3.26
Option trading volume, lagged log	5.41	5.30	2.89	0.19	12.38	4.90	4.70	2.85	0.18	11.84
Underlying price, log	3.30	3.37	1.29	0.33	6.03	3.39	3.45	1.26	0.44	6.10
Underlying return, past week	0.01	0.00	0.09	-0.24	0.32	0.01	0.00	0.09	-0.25	0.33
Total volume in underlying, log	15.43	15.39	1.50	11.89	19.18	15.60	15.56	1.46	12.15	19.27
Underlying spread	0.04	0.04	0.03	0.00	0.18	0.05	0.04	0.03	0.00	0.18
Underlying volatility	0.48	0.36	0.42	0.04	2.39	0.49	0.36	0.43	0.04	2.45
Market cap, log	7.57	7.57	1.94	3.24	12.13	7.76	7.76	1.90	3.46	12.21
D(is ETF)	0.15	0.00	0.35	0.00	1.00	0.14	0.00	0.35	0.00	1.00
Option spread	0.49	0.36	0.41	0.05	2.00	0.48	0.34	0.42	0.05	2.00
Option moneyness	-0.05	-0.04	0.13	-0.47	0.43	-0.10	-0.07	0.18	-0.94	0.37
Option time to expiration	0.08	0.06	0.06	0.00	0.27	0.07	0.05	0.06	0.00	0.30
Option leverage	14.51	10.64	12.65	2.38	75.02	13.61	10.05	12.35	0.97	71.17
Option delta	0.42	0.42	0.14	0.11	0.83	-0.35	-0.33	0.15	-0.82	-0.06
Option gamma	0.12	0.08	0.15	0.01	0.66	0.12	0.07	0.16	0.00	0.71
Option vega	6.64	3.37	10.50	0.19	47.13	6.31	3.16	9.99	0.15	44.84
Option theta	-18.88	-7.67	52.77	-172.61	-0.45	-20.49	-8.68	54.33	-180.76	-0.47

This table reports the descriptive statistics for the daily ticker-level sample from November 2019 till June 20211, separately for call and put options. The sample includes all stock and ETF tickers with lagged price above \$1. SLIM and Small Share are the ticker-level volume shares of SLIM and small trades, respectively. SLIM and Small Imbalance are the ticker-level volume imbalance for SLIM and small trades, respectively. Share and imbalance are constructed similarly for SLIM < \$250, SLIM < \$5,000, SLIM < \$20,000, SLIM \$5,000 − 20,000, MLIM, complex (all multi-leg), large (above 100 contracts) trades and trades above \$50,000. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log, is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying price (log) is as of the day before. Underlying return is the total return over the last week. Total volume in underlying, log, is the logarithm of the total trading volume (lit, ATS and non-ATS OTC) in underlying ticker over the previous week. Underlying spread is averaged over the previous week. Underlying volatility is return volatility over the previous week. Option spread is the contract quoted relative spread. Option time to expiration (in months), moneyness, spread, and leverage are equal-weighted across trades at a ticker level. Option Greeks are from OptionMetrics (not winsorized, equal-weighted across trades at a ticker level). WSB mentions, Robinhood ownership breadth, underlying volatility, and spread as well as option spread, time to expiration, and lambda are winsorized at 99^{th} percentile. Underlying return and option moneyness are winsorized at 0.5^{th} and 99.5^{th} percentiles.

A.14 SLIM volume and quasi-Robinhood portfolio

Table A12: SLIM trading and quasi-Robinhood portfolio

	SLIM Share		**SLIM Imbalance**	
	Call	Put	Call	Put
	(1)	(2)	(3)	(4)
QRH weight	0.017***	0.023***	-0.003	0.034***
	(3.04)	(4.25)	(-0.59)	(6.81)
Observations	1,430,765	1,242,849	1,101,529	834,658
Adjusted R-squared	0.102	0.077	0.021	0.023

This table reports the results of estimating (1) on daily data from November 4, 2019 till August 10, 2020, separately for call and put options. The sample includes all stock and ETF tickers with lagged price above $1. As a dependent variable, we use SLIM Share or SLIM Imbalance. SLIM is a single-leg price improvement auction, through which we measure retail activity. QRH weight is a log weight of the ticker in a quasi-Robinhood portfolio suggested in Welch (2022), using a three-month lag instead of a 12-month lag. All regressions include X and C controls, as described in Section 3.3, as well as date and ticker fixed effects. t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.15 Descriptive statistics on SLIM trades below \$20,000 by category

Table A13: Composition of SLIM trades below \$20,000 in size

Characteristic	**Category**	Frequency share, %	Volume share, %	Quoted spread, %	Effective spread, %
Type	Call	71.5	70.0	13.7	6.5
	Put	28.5	30.0	14.2	6.9
Trade size (contracts)	1	46.4	7.0	13.9	6.3
	2-5	31.3	14.8	12.7	6.1
	6-10	11.7	15.7	14.4	7.3
	11-100	10.2	52.4	16.1	9.1
	Above 100	0.4	10.2	20.2	16.6
Trade size (dollars)	Below 250	42.0	16.0	23.6	11.7
	250-500	15.8	10.0	8.7	3.9
	500-1,000	14.0	12.8	7.4	3.2
	1,000-2,500	14.1	19.6	6.2	2.6
	2,500-5,000	7.1	15.3	5.2	2.1
	5,000-10,000	4.6	14.8	4.5	1.9
	10,000-20,000	2.5	11.4	3.9	3.2
	20,000-50,000				
	Above 50,000				
Trade direction	Sell	50.0	49.8	14.2	7.0
	Buy	47.0	47.8	13.1	6.7
	Midpoint	3.0	2.4	20.3	0.0
Time to expiration	Less than a week	48.5	51.7	12.8	6.7
	1-2 weeks	14.0	13.1	12.6	6.1
	2-4 weeks	15.9	15.1	15.4	7.1
	1-3 months	13.2	12.8	14.2	6.1
	3-12 months	7.2	6.5	19.0	7.5
	Over a year	1.3	1.0	18.7	7.9
Moneyness	Below -2	0.3	0.3	54.1	28.4
	-2 to -1	0.4	0.4	51.0	25.7
	-1 to -0.1	23.7	25.7	28.8	13.9
	At the money	71.6	70.7	8.9	4.2
	0.1 to 1	3.8	2.8	8.9	3.8
	1 to 2	0.1	0.1	9.6	4.5
	Above 2	0.1	0.1	18.1	8.0
Trade direction and type	Sell - Call	35.5	34.6	13.9	6.8
	Sell - Put	14.5	15.2	14.9	7.5
	Buy - Call	33.9	33.8	13.1	6.7
	Buy - Put	13.1	14.0	13.2	6.7
	Midpoint - Call	2.1	1.6	21.0	0.0
	Midpoint - Put	0.9	0.8	18.8	0.0
ETF	No	81.4	74.0	15.1	7.1
	Yes	18.6	26.0	8.5	4.5

This table reports characteristics of SLIM trades (single-leg price improvement auctions) that are smaller than \$20,000 in size by category. Our sample is from November 2019 till June 2021. (Implied) Trade direction is based on whether the trade price is above (buy), below (sell), or at the midpoint. Quoted spread is the spread between the best bid and best ask on the contract (across all exchanges) relative to the midpoint price at the time of the trade. Effective spread is an absolute percentage deviation of the trade price from the midpoint price at the time of the trade, multiplied by 2. For both spreads, we report frequency-weighted averages. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts.

A.16 SLIM trades below $20,000 and other measures of retail activity

Table A14: Share of SLIM option trades below $20,000 in size and other measures of retail activity

	SLIM < $20k trades in calls				SLIM < $20k trades in puts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: SLIM < $20k Share								
Small Share	0.080***				0.072***			
	(32.77)				(34.16)			
Internalized volume in underlying		0.023***				0.019***		
		(8.35)				(7.25)		
Robinhood ownership breadth, log			0.027***				0.061***	
			(2.69)				(6.01)	
WSB mentions, log				-0.002				0.003*
				(-1.16)				(1.77)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.115	0.110	0.102	0.124	0.085	0.082	0.077	0.090
Panel B: SLIM < $20k Imbalance								
Small Imbalance	0.522***				0.521***			
	(262.74)				(227.77)			
Internalized volume in underlying		0.015***				0.004		
		(5.10)				(1.42)		
Robinhood ownership breadth, log			0.042***				0.029***	
			(4.20)				(3.24)	
WSB mentions, log				0.012***				0.010***
				(9.91)				(6.91)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,101,380	1,105,025	435,904	952,614	832,607	837,046	339,427	750,667
Adjusted R-squared	0.187	0.021	0.027	0.020	0.182	0.023	0.025	0.022

This table reports the results of estimating (1) on daily data from November 2019 to June 2021. SLIM < $20k and Small Share are the ticker-level volume shares of SLIM (below $20,000) and small trades, respectively. SLIM < $20k and Small Imbalance are the ticker-level volume imbalance for SLIM (below $20,000) and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log, is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.17 SLIM trades below \$250 and other measures of retail activity

Table A15: Share of SLIM option trades below \$250 in size and other measures of retail activity

	SLIM < \$250 **trades in calls**				**SLIM** < \$250 **trades in puts**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: SLIM < \$250 **Share**								
Small Share	0.227***				0.206***			
	(74.54)				(72.49)			
Internalized volume in underlying		0.010***				0.009***		
		(3.44)				(3.29)		
Robinhood ownership breadth, log			0.007				0.027***	
			(0.70)				(3.01)	
WSB mentions, log				0.006***				0.004***
				(4.41)				(3.46)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.149	0.114	0.102	0.127	0.102	0.070	0.068	0.076
Panel B: SLIM < \$250 **Imbalance**								
Small Imbalance	0.481***				0.474***			
	(205.67)				(160.60)			
Internalized volume in underlying		0.017***				0.013***		
		(5.80)				(3.81)		
Robinhood ownership breadth, log			0.042***				0.028***	
			(3.71)				(2.69)	
WSB mentions, log				0.018***				0.013***
				(13.51)				(9.40)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	938,895	939,586	366,203	823,682	686,300	686,847	271,801	623,959
Adjusted R-squared	0.147	0.029	0.038	0.029	0.140	0.030	0.032	0.029

This table reports the results of estimating (1) on daily data from November 2019 till June 2021. SLIM < \$250 and Small Share are the ticker-level volume shares of SLIM (below \$250) and small trades, respectively. SLIM < \$250 and Small Imbalance are the ticker-level volume imbalance for SLIM (below \$250) and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.18 SLIM trades and other measures of retail activity, most traded tickers only

Table A16: Retail trading in options and other measures of retail activity, most traded tickers only

	Retail trading in calls				Retail trading in puts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: SLIM Share								
Small Share	0.311***				0.209***			
	(17.12)				(16.97)			
Internalized volume in underlying		0.092***				0.067***		
		(3.60)				(4.10)		
Robinhood ownership breadth, log			0.049				0.019	
			(0.95)				(0.46)	
WSB mentions, log				-0.014				0.021**
				(-1.31)				(2.54)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	135,444	135,444	57,838	126,756	135,159	135,159	57,617	126,471
Adjusted R-squared	0.413	0.380	0.358	0.375	0.346	0.328	0.286	0.326
Panel B: SLIM Imbalance								
Small Imbalance	0.282***				0.212***			
	(40.24)				(34.12)			
Internalized volume in underlying		0.033***				0.001		
		(3.30)				(0.13)		
Robinhood ownership breadth, log			0.069***				0.061**	
			(2.64)				(2.49)	
WSB mentions, log				0.048***				0.022***
				(8.93)				(4.15)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	134,974	134,980	57,489	126,292	134,182	134,194	56,850	125,506
Adjusted R-squared	0.112	0.045	0.044	0.045	0.095	0.057	0.053	0.056

This table reports the results of estimating (1) on daily data for the underlying securities in the top decile by their total option dollar traded volume from November 2019 to June 2021 (355 tickers). SLIM and Small Share are the ticker-level volume shares of SLIM and small trades, respectively. SLIM and Small Imbalance are the ticker-level volume imbalance for SLIM and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.19 Characteristics of MLIM trades

In this section, we describe trades that are multi-leg and that went through price improvement auctions. These trades are on average larger than SLIM trades, more balanced by option type, and negatively correlated with equity-based measures of retail activity. Furthermore, a larger fraction of these trades is executed at midpoint.

Table A17: MLIM trades in options and other measures of retail activity

	MLIM trades in calls				**MLIM trades in puts**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: MLIM Share								
Small Share	0.048***				0.045***			
	(16.42)				(22.09)			
Internalized volume in underlying		-0.004				0.006**		
		(-1.44)				(2.23)		
Robinhood ownership breadth, log			0.008				0.012	
			(1.12)				(1.36)	
WSB mentions, log				-0.008***				-0.002
				(-5.62)				(-1.35)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.060	0.058	0.043	0.061	0.048	0.047	0.041	0.049
Panel B: MLIM Imbalance								
Small Imbalance	0.281***				0.373***			
	(59.75)				(75.38)			
Internalized volume in underlying		-0.000				0.000		
		(-0.10)				(0.03)		
Robinhood ownership breadth, log			-0.024*				-0.016	
			(-1.88)				(-1.04)	
WSB mentions, log				-0.001				-0.005***
				(-0.81)				(-3.43)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	542,812	543,022	198,014	502,477	457,425	457,960	174,656	427,021
Adjusted R-squared	0.033	0.014	0.016	0.013	0.058	0.021	0.023	0.019

This table reports the results of estimating (1) on daily data from November 2019 till June 2021. MLIM and Small Share are the ticker-level volume shares of MLIM and small trades, respectively. MLIM and Small Imbalance are the ticker-level volume imbalance for MLIM and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log, is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

Table A18: Composition of MLIM trades

Characteristic	**Category**	Frequency share, %	Volume share, %	Quoted spread, %	Effective spread, %
Type	Call	54.1	53.5	15.4	7.0
	Put	45.9	46.5	18.2	8.5
Trade size (contracts)	1	54.4	10.3	17.6	8.6
	2-5	28.6	16.9	15.4	6.5
	6-10	9.8	16.6	16.2	6.8
	11-100	6.7	40.1	16.1	7.3
	Above 100	0.5	16.2	14.6	7.7
Trade size (dollars)	Below 250	40.6	15.4	29.9	14.9
	250-500	14.6	8.1	10.0	2.9
	500-1,000	13.9	10.1	8.3	2.2
	1,000-2,500	14.3	15.4	7.2	1.8
	2,500-5,000	7.2	11.9	6.2	1.5
	5,000-10,000	4.5	10.8	5.4	1.4
	10,000-20,000	2.5	9.1	4.8	5.3
	20,000-50,000	1.5	8.8	4.2	12.8
	Above 50,000	0.8	10.4	3.7	19.3
Trade direction	Sell	54.3	53.4	13.9	6.1
	Buy	39.6	41.0	20.1	11.0
	Midpoint	6.0	5.6	19.7	0.0
Time to expiration	Less than a week	36.1	40.4	23.1	12.7
	1-2 weeks	14.9	14.7	14.8	6.2
	2-4 weeks	21.8	19.0	13.8	4.5
	1-3 months	20.3	17.5	10.3	3.2
	3-12 months	5.7	6.9	15.3	7.4
	Over a year	1.1	1.5	14.8	9.8
Moneyness	Below -2	0.1	0.3	79.8	37.4
	-2 to -1	0.2	0.3	68.5	24.7
	-1 to -0.1	25.0	22.9	32.8	14.8
	At the money	69.8	71.2	11.4	5.2
	0.1 to 1	4.8	5.1	5.9	5.3
	1 to 2	0.1	0.1	6.7	16.2
	Above 2	0.0	0.1	12.2	23.3
Trade direction and type	Sell - Call	29.0	28.3	13.0	5.8
	Sell - Put	25.3	25.1	14.9	6.5
	Buy - Call	22.0	22.3	18.2	9.7
	Buy - Put	17.6	18.7	22.4	12.7
	Midpoint - Call	3.1	2.9	18.2	0.0
	Midpoint - Put	2.9	2.7	21.3	0.0
ETF	No	74.6	70.6	17.7	7.2
	Yes	25.4	29.4	13.9	9.2

This table reports characteristics of MLIM trades (multi-leg price improvement auctions) by category. Our sample is from November 2019 to June 2021. (Implied) Trade direction is based on whether the trade price is above (buy), below (sell), or at the midpoint. Quoted spread is the spread between the best bid and best ask on the contract (across all exchanges) relative to the midpoint price at the time of the trade. Effective spread is an absolute percentage deviation of the trade price from the midpoint price at the time of the trade, multiplied by 2. For both spreads, we report frequency-weighted averages. Moneyness for calls is measured as $(MidpointPrice - Strike)/Strike$, with the opposite sign for puts.

A.20 Complex strategy trades and measures of retail activity

Table A19: Complex strategy trades in options and measures of retail activity

	Trades in calls				Trades in puts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Complex Share								
Small Share	-0.008***				0.020***			
	(-3.01)				(7.62)			
Internalized volume in underlying		-0.006**				0.000		
		(-2.08)				(0.08)		
Robinhood ownership breadth, log			-0.012				0.003	
			(-1.25)				(0.23)	
WSB mentions, log				-0.011***				-0.002
				(-5.89)				(-0.99)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.126	0.126	0.122	0.135	0.108	0.107	0.101	0.114
Panel B: Complex Imbalance								
Small Imbalance	0.403***				0.514***			
	(105.45)				(127.45)			
Internalized volume in underlying		0.000				0.003		
		(0.10)				(1.03)		
Robinhood ownership breadth, log			-0.018*				-0.002	
			(-1.90)				(-0.17)	
WSB mentions, log				-0.003**				-0.003**
				(-2.17)				(-2.56)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	809,093	810,180	307,652	725,746	706,819	709,103	277,897	641,913
Adjusted R-squared	0.067	0.009	0.012	0.008	0.120	0.012	0.015	0.010

This table reports the results of estimating (1) on daily data from November 2019 to June 2021. Complex and Small Share are the ticker-level volume shares of all multi-leg strategy and small trades, respectively. Complex and Small Imbalance are the ticker-level volume imbalance for all multi-leg and small trades, respectively. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log, is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.21 Trades above $50,000 and measures of retail activity

Table A20: Trades in size above $50,000 in options and measures of retail activity

	Trades in calls				Trades in puts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Share of trades sized above $50,000								
Small Share	-0.209***				-0.193***			
	(-33.84)				(-28.99)			
Internalized volume in underlying		0.017***				-0.005*		
		(7.40)				(-1.77)		
Robinhood ownership breadth, log			0.044***				-0.009	
			(4.25)				(-0.67)	
WSB mentions, log				0.011***				0.003
				(4.54)				(1.12)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,436,457	1,436,457	587,030	1,169,587	1,248,002	1,248,002	514,122	1,051,468
Adjusted R-squared	0.155	0.125	0.120	0.126	0.132	0.104	0.107	0.105
Panel B: Imbalance in trades sized above $50,000								
Small Imbalance	0.248***				0.238***			
	(30.38)				(25.80)			
Internalized volume in underlying		-0.001				0.010		
		(-0.19)				(1.10)		
Robinhood ownership breadth, log			-0.053**				-0.021	
			(-2.26)				(-0.84)	
WSB mentions, log				0.005**				0.006***
				(2.37)				(2.60)
Underlying controls X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contract controls C	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	181,716	181,931	65,258	169,372	132,775	133,173	55,581	123,645
Adjusted R-squared	0.024	0.016	0.020	0.014	0.029	0.021	0.026	0.020

This table reports the results of estimating (1) on daily data from November 2019 to June 2021. Small Share is the ticker-level volume share of small trades. Small Imbalance is the ticker-level volume imbalance for small trades. Internalized volume in underlying is the share of non-ATS OTC (i.e., internalized) volume in the total trading volume in the underlying stock or ETF. Robinhood ownership breadth, log, is the logarithm of the total number of Robinhood users holding the ticker at the end of each day. WSB mentions, log is the logarithm of the number of mentions a ticker gets on `WallStreetBets` during the day. Underlying controls X and contract controls C are described in Section 3.3. All regressions include date and ticker fixed effects. All variables are standardized within the contract type (call or put). t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.22 Characteristics of SLIM and other trade types on option expiration day

Table A21: Composition of option trades on expiration day

Hour to expiration	Trade direction	SLIM		MLIM		Complex	
		Frequency share, %	Volume share, %	Frequency share, %	Volume share, %	Frequency share, %	Volume share, %
1	Sell	5.12	5.99	5.54	5.85	5.40	5.71
1	Buy	3.86	4.38	5.88	6.68	6.98	8.02
1	Midpoint	0.17	0.17	0.51	0.53	0.40	0.38
2	Sell	6.34	6.58	7.09	7.06	6.10	6.14
2	Buy	4.67	5.44	6.80	7.25	6.84	7.26
2	Midpoint	0.22	0.21	0.63	0.59	0.45	0.38
3	Sell	4.63	5.42	4.84	4.80	4.47	4.49
3	Buy	4.10	5.06	4.46	4.79	4.88	5.05
3	Midpoint	0.19	0.20	0.48	0.45	0.36	0.33
4	Sell	5.21	5.65	5.18	5.09	4.82	4.74
4	Buy	4.72	5.53	4.72	5.00	5.15	5.18
4	Midpoint	0.22	0.21	0.51	0.48	0.38	0.33
5	Sell	6.47	6.50	6.30	6.13	5.84	5.66
5	Buy	5.90	6.51	5.61	5.77	6.11	6.10
5	Midpoint	0.27	0.25	0.60	0.55	0.45	0.38
6	Sell	9.84	8.91	8.89	8.42	8.27	7.92
6	Buy	9.12	9.00	7.59	7.74	8.36	8.35
6	Midpoint	0.41	0.36	0.78	0.68	0.61	0.51
7	Sell	14.18	11.36	12.55	11.45	12.04	11.29
7	Buy	13.83	11.83	10.12	9.93	11.40	11.19
7	Midpoint	0.53	0.42	0.92	0.78	0.72	0.59

This table reports characteristics of trades by category for options on their expiration day. Our sample is from November 2019 till June 2021. SLIM (MLIM) stand for the trades that went through a single-leg (multi-leg) price improvement auction, while Complex trades correspond to all milti-leg trades in options. Trade direction is based on the classification method of Muravyev (2016), and 'Midpoint' refers to the trades we could not classify (for additional details, see Section 3.1).

To shed light on statistical significance of observations in Table A21 and Section 3.3.2 in general, we regress the daily series of differences between buy and sell shares onto dummies for each trading hour interacted with trade type. In particular, we estimate the following regression:

$$
\begin{aligned}
Volume\,Share^{buy}_{i,h,t}-&Volume\,Share^{sell}_{i,h,t} \\
&= \beta \sum_{j=1}^{7} D(End\,of\,day - j\,hour(s))_{i,h,t} * D(SLIM)_{i,h,t} \\
&+ \delta \sum_{j=1}^{7} D(End\,of\,day - j\,hour(s))_{i,h,t} * D(MLIM)_{i,h,t} \\
&+ \gamma \sum_{j=1}^{7} D(End\,of\,day - j\,hour(s))_{i,h,t} * D(Complex)_{i,h,t} + \varepsilon_{i,h,t}.
\end{aligned}
$$

Table A22 reports the results. SLIM trades exhibit a statistically significant intraday pattern compared to other trade types: On the option expiration days, there is a larger sell volume share in the last two hours of the trading day. This is consistent with retail brokerages taking an automated action to close retail positions prior to the option's expiration. This pattern does not emerge if the estimation is done on non-expiration days.

Table A22: Intra-day buy-sell patterns on option expiration days

	Buy-sell volume share by trade type:					
	SLIM		MLIM		Complex	
Variable	Coef. (1)	t-stat. (2)	Coef. (3)	t-stat. (4)	Coef. (5)	t-stat. (6)
D(EOD -1 hour)	-0.242***	(-16.39)	-0.087*	(-1.91)	0.100**	(2.08)
D(EOD -2 hours)	-0.084***	(-5.65)	-0.034**	(-2.47)	0.025	(1.27)
D(EOD -3 hours)	0.002	(0.16)	-0.017***	(-5.50)	0.012	(0.80)
D(EOD -4 hours)	0.053***	(3.52)	-0.012**	(-2.71)	0.006	(0.37)
D(EOD -5 hours)	0.088***	(5.81)	-0.015*	(-1.87)	0.005	(0.31)
D(EOD -6 hours)	0.144***	(9.01)	-0.028	(-1.58)	0.010	(0.44)
D(EOD -7 hours)	0.278***	(16.76)	-0.037	(-0.93)	-0.003	(-0.07)
Test equality to SLIM -1 hour				10.58***		55.33***
Test equality to SLIM -2 hours				6.67**		52.55***

This table reports estimation results from a pooled regression of hourly volume share difference between buy and sell trades on hourly dummies interacted with trade types on option expiration days from November 2019 till June 2021. Total number of observations is 18,432. D(EOD -X hours) equals 1 for X^{th} hour to the end of the trading day (EOD) for the respective trade type: SLIM trades in Column (1), MLIM trades in Column (3), and all milti-leg trades in Column (5). Constant is excluded. t-statistics are based on standard errors double-clustered by date and trade type. The last two rows report results of a Wald test for the same buy-sell volume share in the last two trading hours of SLIM trades compared to MLIM and Complex trades (i.e., comparing the corresponding coefficients in front of D(EOD -1 hour) and D(EOD -2 hours) across different trade types). *** p<0.01, ** p<0.05, and * p<0.1.

A.23 SLIM volume and Robinhood herding events (frenzies)

Table A23: Options trade imbalances and herding events

	Imbalance in trades of type:							
	SLIM		**SLIM** < $250		**SLIM** < $5**k**		**SLIM** < $20**k**	
	Call	Put	Call	Put	Call	Put	Call	Put
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D(Robinhood frenzy)	0.073***	0.094***	0.128***	0.179***	0.088***	0.133***	0.075***	0.104***
	(3.09)	(3.01)	(5.28)	(5.60)	(3.66)	(4.39)	(3.16)	(3.40)
Observations	450,681	350,957	377,592	280,253	446,646	346,076	450,103	350,102
Adjusted R-squared	0.026	0.024	0.037	0.031	0.028	0.025	0.026	0.024
	MLIM		**All complex**		**All** > $50**k**		**All** > 100 **contracts**	
	Call	Put	Call	Put	Call	Put	Call	Put
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
D(Robinhood frenzy)	-0.115***	-0.019	-0.064*	-0.012	0.128	-0.126	-0.035	-0.006
	(-2.81)	(-0.30)	(-1.96)	(-0.37)	(1.55)	(-1.44)	(-0.92)	(-0.09)
Observations	204,043	179,808	317,816	286,963	67,277	57,106	130,141	98,373
Adjusted R-squared	0.015	0.023	0.012	0.015	0.020	0.026	0.021	0.030

This table reports the results of estimating (1) on daily data from November 4, 2019 August 10, 2020, separately for call and put options. The sample includes all stock and ETF tickers with lagged price above $1. As a dependent variable, we use imbalance of contract volume traded via the indicated trade type, aggregated at the ticker level. SLIM is a single-leg price improvement auction, through which we measure retail activity. SLIM < $250, < $5k, and < $20k correspond to SLIM trades of the respective dollar size. MLIM is a multi-leg price improvement auction. D(Robinhood frenzy) equals 1 if the ticker experienced a Robinhood herding event using the data of Barber, Huang, Odean, and Schwartz (2022). All regressions include X and C controls, as described in Section 3.3, as well as date and ticker fixed effects. t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1.

A.24 Fraction not exercised, and trade types

Table A24: Suboptimal exercise and trading via different trade types

	Fraction of OI not exercised, %			
	(1)	(2)	(3)	(4)
SLIM Share	4.561***			
	(5.40)			
MLIM Share		-0.729		
		(-0.53)		
Complex Share			-2.541***	
			(-3.91)	
Large Share				-3.384
				(-1.48)
Observations	41,737	41,737	41,737	41,737
Adjusted R-squared	0.206	0.205	0.206	0.205
Contract controls	Y	Y	Y	Y

This table reports estimates of equation (4) in our dividend play sample. SLIM Share is the contract-level volume shares of SLIM trades, averaged over one trading week before the cum-dividend date (similar for MLIM, complex, and large trades). MLIM trades are trades that went through multi-leg price improvement auctions. Complex trades are all multi-leg trades. Large trades are trades with lot size above 100. Contract controls include log dollar trading volume, relative spread, IV, moneyness, days to expiration, log OI, and EEV. All regressions include ticker by date fixed effects. S.E. are clustered by ticker and date. Robust t-statistics in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

A.25 Aggregate SLIM performance, value-weighted prices

Table A25: SLIM trade performance, aggregate and by contract characteristics, using value-weighted prices

	SLIM Raw performance, $ billion				
Horizon h	1 day	2 days	5 days	10 days	Expiration
Panel A: All contracts					
	-0.705	-1.205	-1.326	-1.336	-1.461
Panel B: By contract type					
Call	-0.376	-0.890	-1.096	-1.232	-1.114
Put	-0.329	-0.315	-0.230	-0.104	-0.346
Panel C: By moneyness					
Below -2	-0.003	-0.004	-0.003	-0.001	-0.001
-2 to -1	-0.005	-0.005	-0.001	0.005	-0.003
-1 to -0.1	-0.005	0.038	0.239	0.396	0.364
At the money	0.076	-0.286	-0.583	-0.692	-0.669
0.1 to 1	-0.460	-0.637	-0.663	-0.709	-0.798
1 to 2	-0.156	-0.158	-0.172	-0.181	-0.183
Above 2	-0.148	-0.153	-0.142	-0.153	-0.164
Panel D: By time to expiration					
Less than a week	-0.507	-0.920	-1.291	-1.294	-1.285
1-2 weeks	-0.091	-0.158	-0.208	-0.322	-0.357
2-4 weeks	-0.051	-0.105	-0.101	-0.220	-0.213
1-3 months	-0.025	-0.033	0.100	0.189	0.305
3-12 months	0.072	0.100	0.228	0.346	-0.138
Over a year	-0.103	-0.088	-0.054	-0.036	0.228

This table reports the performance of SLIM trades from November 2019 till June 2021. Raw performance at each horizon is computed as explained in Section 3.4. We use value-weighted average transaction prices.

A.26 Aggregate SLIM performance, by month and weekday

Table A26: SLIM trade performance, by month and weekday

	SLIM Raw performance, $ billion				
Horizon h	1 day	2 days	5 days	10 days	Expiration
Panel A: By month					
Nov-19	-0.017	-0.009	-0.011	-0.021	0.099
Dec-19	-0.012	-0.013	-0.013	-0.007	0.151
Jan-20	0.016	0.023	0.069	0.149	0.326
Feb-20	-0.101	-0.149	-0.205	-0.163	-0.355
Mar-20	0.084	0.076	0.139	0.191	-0.391
Apr-20	0.003	-0.007	-0.003	-0.022	0.075
May-20	-0.030	-0.016	-0.012	-0.017	0.014
Jun-20	-0.088	-0.179	-0.048	-0.098	0.043
Jul-20	0.016	0.037	0.125	0.168	0.159
Aug-20	0.044	0.078	0.039	0.026	0.019
Sep-20	0.059	0.010	0.001	0.000	0.026
Oct-20	0.053	0.006	0.022	0.060	0.028
Nov-20	-0.025	-0.032	-0.027	0.039	0.099
Dec-20	0.073	0.077	0.030	-0.031	0.025
Jan-21	-0.156	-0.493	-0.848	-0.780	-0.898
Feb-21	-0.042	-0.135	-0.209	-0.337	-0.343
Mar-21	-0.145	-0.095	-0.079	-0.113	-0.053
Apr-21	-0.166	-0.170	-0.184	-0.216	-0.206
May-21	-0.078	-0.016	0.001	-0.016	-0.033
Jun-21	-0.091	-0.092	-0.002	-0.033	-0.160
Panel B: By weekday					
Mon	0.092	0.118	-0.084	-0.216	-0.386
Tue	-0.005	-0.228	-0.141	-0.111	0.077
Wed	0.050	-0.169	-0.190	-0.146	-0.209
Thu	-0.258	-0.312	-0.266	-0.288	-0.364
Fri	-0.464	-0.494	-0.523	-0.455	-0.458

This table reports the performance of SLIM trades in November 2019 to June 2021. Raw performance at each horizon is computed as explained in Section 3.4.

A.27 Aggregate SLIM performance, best and worst tickers

Table A27: Best and worst performing tickers, based on trades originated by SLIM investors and the whole market

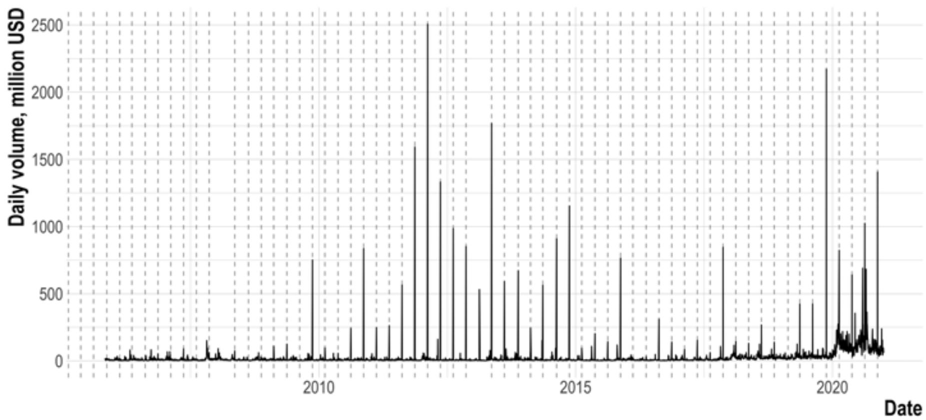
Ticker	Name	SLIM Raw performance, $ billion					Ticker	Name	Market Raw performance, $ billion				
		1 day	2 days	5 days	10 days	Expiration			1 day	2 days	5 days	10 days	Expiration
Top 10 tickers for SLIM trades							**Top 10 tickers for all market trades**						
AMZN	Amazon.com Inc	0.419	0.415	0.341	0.277	0.057	AMZN	Amazon.com Inc	2.184	2.359	2.414	2.412	1.371
NVDA	NVIDIA Corp	0.032	0.061	0.100	0.117	0.064	GME	GameStop Corp	1.094	1.101	1.709	1.543	1.052
AAPL	Apple Inc	-0.012	0.003	0.027	0.107	0.097	SHOP	Shopify Inc	0.689	0.763	0.869	0.721	0.702
SHOP	Shopify Inc	0.076	0.076	0.087	0.079	0.057	AAPL	Apple Inc	0.762	0.874	0.672	0.653	0.295
MRNA	Moderna Inc	0.012	0.021	0.032	0.034	0.053	AMC	AMC Entertainment Holdings Inc	0.024	0.045	0.387	0.487	0.619
GOOGL	Alphabet Inc	0.006	0.020	0.036	0.026	-0.001	NVDA	NVIDIA Corp	0.155	0.282	0.224	0.354	-0.192
BABA	Alibaba Group Holding Ltd	0.023	0.031	0.040	0.025	0.009	BA	Boeing Co	0.345	0.406	0.256	0.338	0.284
DIS	Walt Disney Co	-0.001	0.004	0.007	0.022	-0.041	ZM	Zoom Video Communication Inc	0.360	0.370	0.348	0.328	0.469
LQD	iShares iBoxx $ Investment Grade Corporate Bond ETF	0.000	0.000	0.006	0.018	0.008	NIO	Nio Inc - ADR	0.277	0.300	0.289	0.327	0.650
PLTR	Palantir Technologies Inc	-0.003	0.002	0.007	0.017	0.008	ROKU	Roku Inc	0.177	0.272	0.334	0.306	0.134
Bottom 10 tickers for SLIM trades							**Bottom 10 tickers for market trades**						
TSLA	Tesla Inc	-0.471	-0.612	-0.901	-0.903	-0.416	TSLA	Tesla Inc	-1.803	-1.643	-1.322	-1.389	-1.595
SPY	SPDR S&P 500 ETF	-0.289	-0.476	-0.495	-0.433	-0.381	SPY	SPDR S&P 500 ETF	-0.258	-0.786	-1.323	-1.291	-0.467
QQQ	Invesco Nasdaq-100 ETF	-0.052	-0.115	-0.103	-0.158	-0.201	TLRY	Tilray Brands Inc	-0.390	-0.397	-0.335	-0.393	-0.556
GME	GameStop Corp	0.035	-0.121	-0.106	-0.112	-0.134	IWM	iShares Russell 2000 ETF	-0.038	-0.071	-0.120	-0.287	-0.116
AMC	AMC Entertainment Holdings Inc	-0.062	-0.094	-0.106	-0.096	-0.103	NFLX	Netflix Inc	-0.131	-0.131	-0.167	-0.143	-0.013
MSTR	MicroStrategy Inc	-0.032	-0.030	-0.031	-0.039	-0.025	LQD	iShares iBoxx $ Investment Grade Corporate Bond ETF	0.010	0.002	-0.046	-0.138	0.034
IWM	iShares Russell 2000 ETF	-0.010	-0.006	-0.008	-0.039	-0.074	EEM	iShares MSCI Emerging Markets ETF	-0.088	-0.041	-0.052	-0.114	0.043
MARA	Marathon Digital Holdings Inc	-0.002	-0.003	-0.008	-0.035	-0.044	XOP	SPDR S&P Oil & Gas Exploration & Production ETF	0.008	0.004	-0.039	-0.073	-0.274
RIOT	Riot Blockchain Inc	-0.002	-0.003	-0.018	-0.035	-0.072	NKE	NIKE Inc	-0.017	-0.028	-0.048	-0.071	-0.074
PLUG	Plug Power Inc	-0.006	-0.014	-0.025	-0.026	-0.074	MARA	Marathon Digital Holdings Inc	0.014	0.003	0.004	-0.071	-0.017

This table reports the performance of top 10 and bottom 10 tickers, based on trades originated by SLIM investors and the market overall from November 2019 to June 2021. We assume a 10-day holding period. Raw performance at each horizon is computed as explained in Section 3.4. Performance is reported from the perspective of the originating counterparty (and hence, total profits/losses within a ticker do not necessarily need to sum up to zero).

A.28 Abnormal trading volume on cum-dividend dates: Further examples

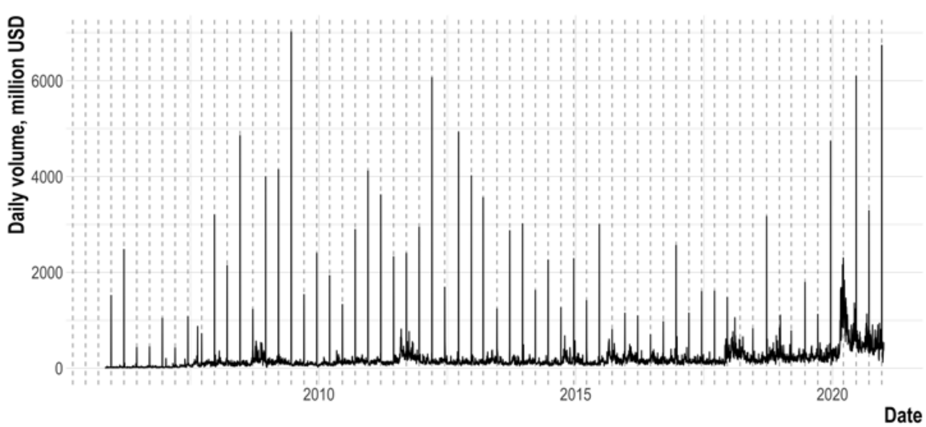
This appendix contains two further examples of abnormal trading volume on cum-dividend dates. The figures below plot daily trading volume of options on Microsoft, MSFT, and on the largest S&P 500 ETF, SPY.

Figure A5: Abnormal trading volume on cum-dividend dates for Microsoft



This figure plots daily trading volume for all call option contracts on MSFT, in millions of U.S. dollars, as reported in OptionMetrics. The dashed lines indicate cum-dividend dates.

Figure A6: Abnormal trading volume on cum-dividend dates for SPY, S&P 500 ETF



This figure plots daily trading volume for all call option contracts on SPY, in millions of U.S. dollars, as reported in OptionMetrics. The dashed lines indicate cum-dividend dates.

A.29 Dividend play: Another example

Table A28 provides an additional example illustrating the mechanics of the dividend play strategy. Case 1 corresponds to the case when all 1,000 outstanding contracts are exercised: All 1,000 short positions get assigned, so there is no profit for a dividend play strategy to harvest. Case 2 describes what happens if 500 of 1,000 outstanding contracts are left unexercised. Without arbitrageur involvement, half of the short positions in the contract get assigned; the remaining positions deliver a gain of \$0.5 per share and \$25,000 in total for the unassigned short positions, a gain to the original customers with short positions. Now consider the entry of arbitrageurs. The arbitrageurs attempt to recover most of the potentially harvestable profit of \$25,000. To do so, they buy and simultaneously sell 5,000 contracts and exercise all their long positions. The probability of assignment increases, but, because of the OCC's random assignment, some of the short positions of the arbitrageurs remain unassigned and hence yield a gain. In our example, arbitrageurs harvest \$20,850 out of the total gain of \$25,000. To divert a larger fraction of the total gain from the original customers with short positions, arbitrageurs simply increase the number of contracts they buy and sell.

Table A28: Dividend play: Another example

	OI_{t-1}	New posi-tions(t)	Available for ex.	No. ex-ercised	Prob. Assign.	No. assign.	No. not assign.	Gain per share	Total gain on unassign. positions	OI_t	Fraction unex.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Case 1. Optimal exercise											
Customer	1000	0	1000	1000	100%	1000	0.00	0.5	0	0.00	
Case 2. Suboptimal exercise											
Case 2.1. Without dividend play											
Customer	1000	0	1000	500	50%	500	500	0.5	25000	500	0.5
Case 2.2. With dividend play											
Customer	1000	0	1000	500		916.7	83.33	0.5	4166.7		
Arbitrageurs	0	5000	5000	5000		4583.3	416.67	0.5	20833.3		
Total	1000	5000	6000	5500	92%	5500	500		25000	500	0.5

This table illustrates the dividend play strategy. Date t refers to the cum-dividend date, and OI_t stands for the open interest on date t. This table is similar to Table 1 in Pool, Stoll, and Whaley (2008).

A.30 Dividend play: Technical details

We compute the expected call option ex-dividend price using the Black-Scholes-Merton formula:

$$
\begin{aligned}
c_{ex} &= S_{ex}e^{-y(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2), \\
d_1 &= \frac{1}{\sigma\sqrt{T-t}} ln\left(\frac{S_{ex}}{K} + \left[r - y + \frac{\sigma^2}{2}\right](T-t)\right), \\
d_2 &= d_1 - \sigma\sqrt{T-t}, \\
y &= Dividend_{ex}/S_{ex},
\end{aligned}
$$

where S_{ex} is the expected price after the stock goes ex-dividend, that is the price at close on the cum-dividend day minus expected dividend, $T-t$ is time to maturity in years, that is difference in the expiration date and the current date in days divided by 360, K is the contract strike, σ^2 is the annualized implied volatility,[70] r is the interpolated maturity-specific interest rate provided by OptionMetrics (annualized %), and $Dividend_{ex}$ is the expected dividend after the ex-date.[71]

A.31 Dividend play sample: Data filters and calculated variables

We use our dataset described in 3.1 together with the following filters to arrive at the final dividend play sample. We include all call option contracts on dividend-paying stocks with $EEV > 0$. Furthermore, since our valuation might be imperfect, we add a market-based filter of the optimality of exercise: We only keep contracts with a decline in open interest on the cum-dividend date.[72] By implication, we only have contracts with non-missing open interest on the cum-dividend date and the date before that. Following the early papers on dividend play, we remove contracts with missing trading volume on cum-dividend date (either in OPRA or in OptionMetrics).

To measure arbitrageur activity, we use floor trading share, defined as the total volume in transactions of OPRA types SLFT and MLFT, divided by the total volume on the cum-dividend date.[73] For both SLIM and Small Share, we compute a one-week moving average (requiring a minimum of a one-day observation) and use its lagged value on the cum-dividend

[70] We use the daily contract-level implied volatility from OptionMetrics. If it is missing, we interpolate it from the neighboring strikes.

[71] We assume that its size is equal to the current dividend if the stock pays one more dividend after the current dividend until the option expires and 0 otherwise.

[72] This is consitent with Hao, Kalay, and Mayhew (2010).

[73] In unreported tests, we confirm that using dollar-volume-based measures instead yields similar results.

Table A29: Dividend play sample descriptive statistics

	Mean	Median	St. Dev.	p1	p99
Fraction of OI not exercised, %	17.50	1.99	28.17	0.00	98.71
Floor trades volume share on cum-date	0.49	0.58	0.47	0.00	1.00
D(floor share > 0)	0.54	1.00	0.50	0.00	1.00
SLIM Share	0.14	0.06	0.18	0.00	0.81
Small Share	0.84	1.00	0.22	0.00	1.00
Internalized volume in equities	0.17	0.16	0.05	0.07	0.30
WSB mentions, log	-1.05	-0.89	3.20	-4.61	6.99
OI, log	4.28	4.19	2.21	0.00	9.60
Early exercise value (EEV), $	0.52	0.34	0.60	0.00	3.01
Market EEV, $	0.07	0.02	0.37	-0.57	1.07
Potential profit, $	4,466.65	53.66	48,262.21	0.00	70,017.45
Potential profit, log $	3.70	4.00	3.46	0.00	11.16
Dollar volume, log	1.85	1.55	1.39	0.00	6.61
Relative spread	0.09	0.05	0.13	0.00	0.65
Implied volatility, annualized	0.44	0.37	0.45	0.00	1.72
Moneyness	12.09	5.48	20.75	0.51	108.35
Days to expiration	50.14	14.00	108.12	1.00	603.00

This table reports descriptive statistics for all contracts in the dividend play sample (29,111 observations). SLIM and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the cum-dividend date. Internalized volume in equities is the ticker-level share of volume executed in the non-ATS OTC space relative to the total trading volume, averaged over one trading week before the cum-dividend date. WSB mentions is the number of underlying ticker mentions on `WallStreetBets` forum, averaged over one trading week before the cum-dividend date. Relative spread is options contract quoted spread at the time of the trade relative to the midpoint price. Implied volatility is as reported in LiveVol, interpolated using nearest strikes if missing. Moneyness of call options is measured as $(Midpoint\, Price - Strike)/Strike$.

date. We use the same rolling measures for the retail activity variables described in the main text, as well as volume, spread, and implied volatility controls.

We compute relative spread quoted at the time of each option trade as $2(best\, ask - best\, bid)/(best\, ask + best\, bid)$ (relative to the midpoint price). We compute moneyness of a call option as $0.5(underlying\, bid + underlying\, ask)/strike - 1$.[74]

[74] In the absence of TAQ data, we use underlying bid-ask midpoint as a high-frequency price.

A.32 Dividend risk and automatic actions of retail brokerages

This appendix presents an example of an automatic action to close short positions exposed to dividend risk on cum-dividend dates undertaken by retail brokerages. The example is from the Robinhood Terms and Conditions.

Figure A7: Excerpt from Robinhood's Terms and Conditions

Options Dividend Risk

Dividend risk is the risk that you'll get assigned on any short call position (either as part of a covered call or spread) the trading day before the underlying security's ex-dividend date. If this happens, you'll open the ex-date with a short stock position and actually be responsible for paying that dividend yourself. You can potentially avoid this by closing any position that includes a short call option at any time before the end of the regular-hours trading session the day before the ex-date.

Robinhood may take action in your account to close any positions that have dividend risk the day before an ex-dividend date. Generally, we'll only take action if your account wouldn't be able to cover the dividend that would be owed after an assignment. This is done on a best-efforts basis.

A.33 Arbitrageur activity and retail popularity: Matched contracts

We study the relationship between floor trading share on cum-dividend date and retail popularity. To do this, we isolate contracts with high retail popularity (top quintile of SLIM Share and construct the control group of contracts matched on the profitability characteristics from contracts with low retail popularity. In the basic set of control characteristics, we use open interest, early exercise value, and moneyness. We also report results with the characteristics extended to include contract and underlying spreads, as well as the price of the underlying. We report the covariate balance plot in Figure A8 below.

Table A30 confirms the findings of Section 5.3. We again see that contracts that had experienced a larger volume of retail trading in the week preceding the cum-dividend date are more targeted by the arbitrageurs. The magnitudes are very similar to those in Table 9. Varying controls and the number of neighbors do not affect the magnitudes of the effects. The coefficients are also statistically indistinguishable from the OLS estimates from the same specification (Column (4)), which offers further evidence that our empirical findings are robust.

Figure A8: Covariate balance for SLIM dummy variable in Table A30

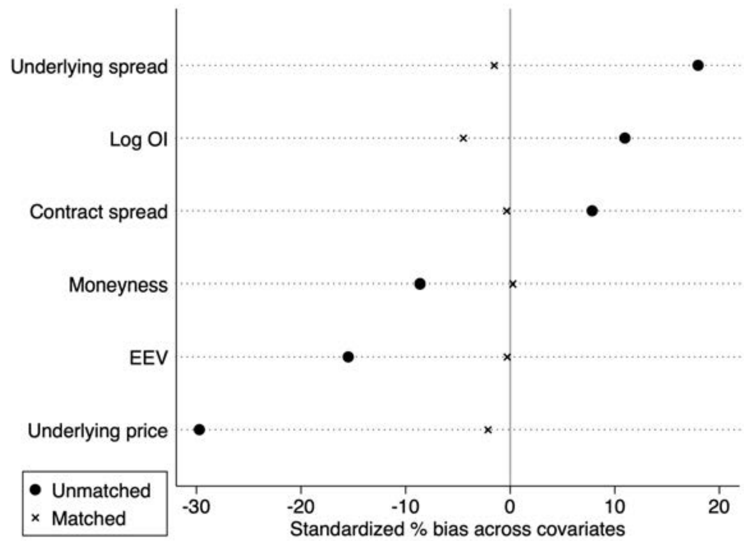


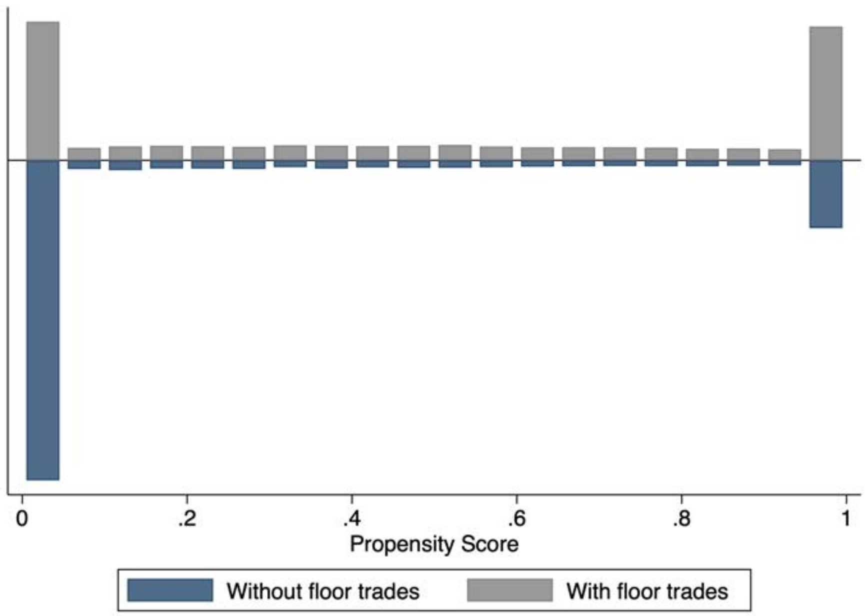
Table A30: Arbitrageur activity and retail popularity: Matched contracts

	Floor trading share on cum-date			
		Matched		OLS
	(1)	(2)	(3)	(4)
D(SLIM Share in top quintile)	0.015**	0.016***	0.013**	0.014**
	(2.08)	(2.73)	(2.17)	(2.31)
Observations	33,684	33,684	33,684	33,684
No. neighbors	1	10	10	
Short controls	Y	Y	Y	Y
Extended controls	N	N	Y	Y

This table reports the results of propensity score matching estimation and OLS estimates for the same set of contracts as in our dividend play sample (contracts expiring in more than two days). Columns (1)–(3) report ATE. SLIM and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the cum-dividend date. Short controls include: log OI, EEV, and moneyness. Extended controls include contract spread, underlying price, and underlying spread. Robust z-statistics are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

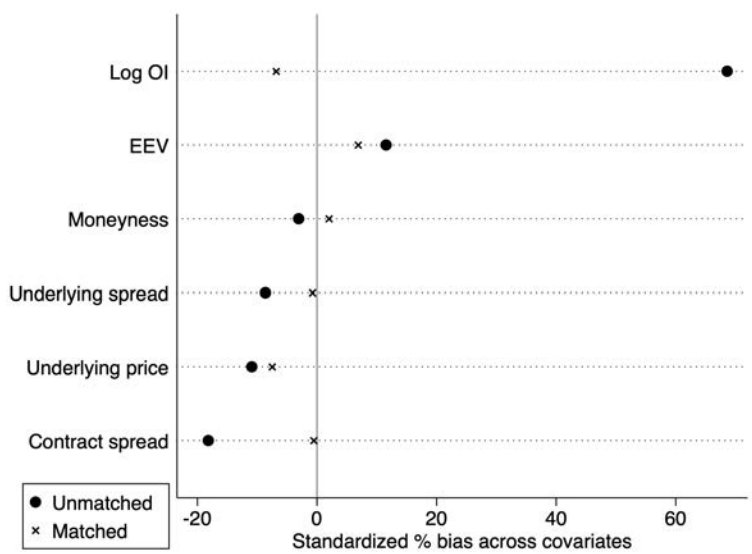
A.34 Dividend play puzzle in matched contracts

Figure A9: Floor traders' entry across propensity score levels



This figure depicts the number of contracts with and without floor trades across the scores of propensity to have floor trades. The propensity scores are based on the same set of controls: log OI, EEV, contract spread, moneyness, underlying spread, and underlying price.

Figure A10: Covariate balance for Floor dummy variable in Figure A9

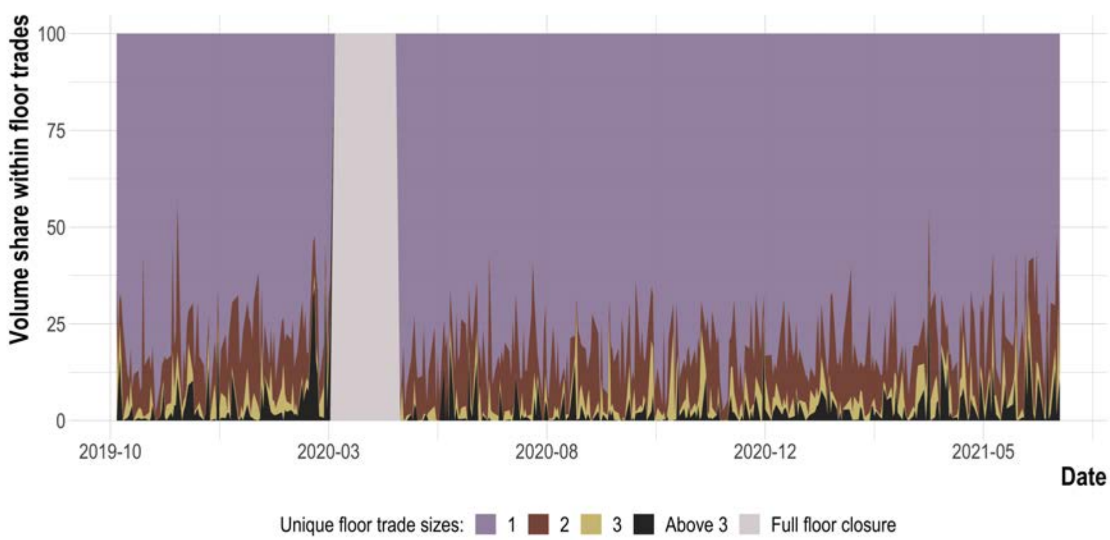


A.35 Few arbitrageurs engaging in dividend play

This appendix provides suggestive evidence for the number of arbitrageurs simultaneously engaging in a dividend play strategy in a particular contract. Figure A11 plots a percentage split of dividend play trades by unique trade sizes, which is our proxy for a number of arbitrageurs engaging in dividend play in each contract.

The gray-shaded area in Figure A11 corresponds to the closure of all exchange floors in the U.S. due to the COVID-19 pandemic. Our measure of floor trading is indeed zero over this period. Furthermore, the total trading volume on cum-dividend dates during the closures is the same as on any other day, which provides additional validation of the measure. Even when the PHLX floor was closed but ARCA and BOX floors were open, the mean trading volume on cum-dates was one order of magnitude lower.

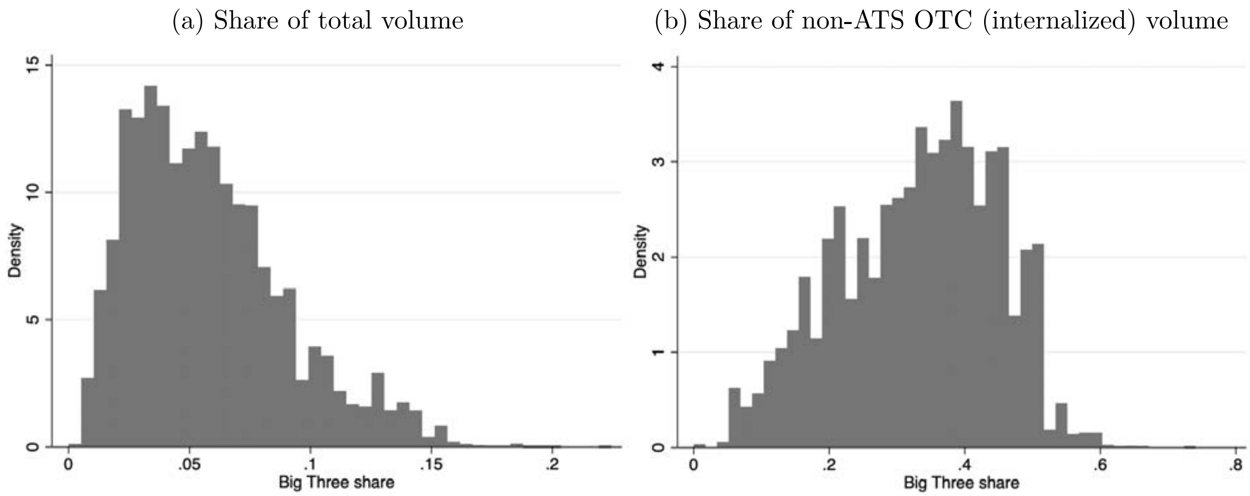
Figure A11: Floor trading by number of floor trade sizes



This figures depicts percentage split of trades executed on exchange floor by the number of unique trade sizes. We only include contracts in our dividend play sample. The gray shaded area corresponds to the period of floor closures on all exchanges.

A.36 The Big Three share of trading volume in underlying equities and ETFs

Figure A12: The Big Three: Share of trading volume in underlying equities and ETFs



This figure plots histograms of weekly non-ATS OTC volume share in underlying equities and ETFs by the Big Three wholesalers in options, Citadel, Susquehanna, and Wolverine. Panel (a) uses share in total trading volume (on lit exchanges, ATS and non-ATS OTC), Panel (b) uses share in total non-ATS OTC volume. Based on FINRA OTC Transparency and CRSP data for the underlying securities in our dividend play sample.

A.37 Big Three share and floor trading, continuous share

Table A31: Arbitrageur activity and market concentration

	Floor trading share on cum-date			
	(1)	(2)	(3)	(4)
D(SLIM imbalance)	-0.033***		-0.019	
	(-3.12)		(-1.23)	
D(SLIM buy imbalance)		-0.017***		-0.010
		(-3.45)		(-1.03)
Big Three share in equity volume			-0.074	-0.217
			(-0.20)	(-0.64)
Big Three share in equity volume × D(SLIM trade imbalance)			-0.282*	
			(-1.67)	
Big Three share in equity volume × D(SLIM buy trade imbalance)				-0.146
				(-1.16)
SLIM Share	0.069***	0.031**	0.067***	0.025*
	(4.54)	(2.31)	(4.10)	(1.75)
Observations	41,737	41,737	40,713	40,713
Adjusted R-squared	0.415	0.415	0.375	0.374
FE	Ticker*Date	Ticker*Date	Ticker and Date	Ticker and Date
Contract controls	Y	Y	Y	Y
Ticker controls	N	N	Y	Y

This table further explains floor trader activity in our dividend play sample. Big Three share is the total share of the Big Three wholesalers' non-ATS OTC volume in the total stock trading volume over the past trading week. $D(SLIM\,imbalance) = 1$ if there was an order imbalance in SLIM trades over the past trading week. $D(SLIM\,buy\,imbalance) = 1$ if there was a positive order imbalance in SLIM trades over the past trading week. Contract controls include SLIM Share, log OI, EEV, log dollar trading volume, relative spread, IV, moneyness, and days to expiration. Ticker controls include: underlying price, underlying volatility, underlying relative bid-ask spread, underlying market cap. t-statistics are based on standard errors clustered by ticker and date (in parentheses). *** p<0.01, ** p<0.05, and * p<0.1

A.38 Big Three share and floor trading, matching approach

Similar to Table A30 in the main text, we employ a matching approach to study the importance of concentration in PFOF market for the floor trading share on cum-dividend date. For matching, we use the same characteristics as in the main text. The corresponding covariate balance plot is presented in Figure A13 below.

Table A32: Arbitrageur activity and market concentration: Matched contracts

	Floor trading share on cum-date			
	Matched			OLS
	(1)	(2)	(3)	(4)
D(Big Three share > 10%)	-0.055***	-0.050***	-0.036***	-0.053***
	(-6.48)	(-7.00)	(-4.36)	(-7.31)
Observations	41,737	41,737	41,737	41,737
No. neighbors	1	10	10	
Short controls	Y	Y	Y	Y
Extended controls	N	N	Y	Y

This table reports the results of propensity score matching estimation and OLS estimates for the same set of contracts as in our dividend play sample (contracts expiring in more than two days). Columns (1)–(3) report ATE. Big Three share is the total share of the Big Three wholesalers' non-ATS OTC volume in the total stock trading volume over the past trading week. Short controls include: log OI, EEV, moneyness. Extended controls include contract spread, underlying price, and underlying spread. Robust z-statistics in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

Figure A13: Covariate balance for the top decile of Big-Three share dummy in Table A32