Profiting from Investor Mistakes: Evidence from Suboptimal Option Exercise

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

March 31, 2023

Abstract

The recent retail trading boom in the U.S. options market, driven by an inflow of inexperienced investors, coincides with an increase in call option contracts left suboptimally unexercised. Arbitrageurs exploit these investor mistakes via so-called dividend play trades, which produce (virtually) riskless profits. Using transaction-level data and a new reporting requirement, we identify dividend play trades and document rising profits from this strategy. Puzzlingly, however, arbitrageurs leave 50% of potential profits on the table. Few arbitrageurs exploit each contract, with a median of only one. Finally, our paper details how market participants circumvent a regulation devised to curtail dividend play.

JEL Classification: G4, G5, G11, G12

Keywords: Retail, dividend play, investor mistakes, arbitrage, WallStreetBets

---

*We thank Brad Barber, Su Li, Chester Spatt, Dimitri Vayanos, Ingrid Werner, and seminar participants at Indiana University, London Business School, Stockholm Business School, and the University of Central Florida for their thoughtful comments, Miguel Chumbo for excellent research assistance, and the AQR Asset Management Institute at London Business School for generous support. 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

Over the past several years, financial markets witnessed an unprecedented inflow of retail investors. Lured by new zero-commission brokerage offers, introduced by investing app Robinhood, millions of young and tech-savvy yet inexperienced investors have started trading stocks, options, and cryptocurrencies. At the peak of its popularity in late 2021, Robinhood alone reported 21.3 million monthly active users. It has been widely documented in the literature that retail investors make mistakes.[1] Our paper focuses on retail investor activity in the options market, from which Robinhood made 38% of its 2021 revenue.[2] As options are complex financial instruments, one would expect the new generation of investors to make mistakes in their options transactions.

As a laboratory for the study of the transactions of sophisticated investors capitalizing on retail investor mistakes, we use the so-called dividend play trades, which are pairs trades that produce (virtually) riskless arbitrage profits for market makers and other arbitrageurs. These profits derive from call options left suboptimally unexercised before the underlying stock goes ex-dividend. It may be optimal to exercise an American call option on a dividend-paying stock before maturity, but the decision to do so requires high investor sophistication (e.g., an application of an option-pricing model). We show that an inflow of inexperienced retail investors has boosted potential gains from this strategy. Exploiting the new reporting requirements, we identify dividend play trades and document that, rather than harvesting all the arbitrage profits from suboptimally unexercised call options, market makers and other sophisticated arbitrageurs leave half of the potential gains to writers of these options. This puzzling behavior cannot be explained by arbitrageur costs or constraints. Finally, we document that the number of arbitrageurs participating in dividend play in each contract is quite small, with a median of only one.

We use two alternative approaches to classify retail trades in our data. Our first approach, commonly used in the industry, identifies small trades (up to 10 contracts) as retail.[3] Our second approach is from Bryzgalova, Pavlova, and Sikorskaya (2023) (BPS).[4] BPS exploit a new reporting requirement introduced by the Options Price Reporting Authority

[1]See e.g., Barber and Odean (2001) and Calvet, Campbell, and Sodini (2007). Barber and Odean (2013) provide an excellent overview.

[2]Per the company's quarterly statements, in 2021 Robinhood's total net revenue was \$1.82 billion, and its revenue from options was \$689 million.

[3]For example, Deutsche Bank and Bloomberg rely on small trades to proxy retail participation in options. See `https://markets.businessinsider.com/news/stocks/stock-market-outlook-retail-investing-shorts-trading-options-deutsche-bank-2021-1-1030005344`) and `https://www.bloomberg.com/professional/blog/gamestop-highlights-importance-of-option-related-equity-flows/`, respectively.

[4]See also Ernst and Spatt (2022) and Hendershott, Khan, and Riordan (2022).

(OPRA) in November 2019 as well as a specific institutional feature of the U.S. options market, namely, Single-Leg price Improvement Mechanisms (SLIM), to identify retail trades. They term them SLIM trades. Both measures detect a surge in retail investor trading activity in our sample. We supplement these measures with ticker mentions in `WallStreetBets`, an investing forum popular with the new generation of retail investors, and the Robinhood user count provided by Robintrack. Our final proxy for retail investor interest in a ticker is the internalized volume in underlying equities, reported to the Financial Industry Regulatory Authority (FINRA).

Given the recent surge in retail investor activity, it is important to understand its implications for the 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 market makers and other arbitrageurs who exploit it. The mistake is a failure to exercise in-the-money call options before the underlying stock or exchange-traded fund (ETF) goes ex-dividend when it is optimal to do so.[5] To benefit from it, market makers and other arbitrageurs engage in "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,[6] and it is therefore available only to floor market makers and other floor participants. We exploit the new OPRA trade types to accurately classify such arbitrage trades and study the behavior of arbitrageurs. Due to dividend play, the daily trading volume on the 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. Even for the largest S&P 500 ETF, SPY—the ticker with the most actively traded options in 2021—the last cum-dividend date trading volume is typically 14–53 times larger.[7] In this paper, we often refer to the last cum-dividend dates on which dividend play trade takes place as dividend play dates.

Interestingly, in the data provided by the Options Clearing Corporation (OCC), a large fraction of the increased volume on dividend play dates is attributed to "Customer,"

[5]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 transaction costs outweigh profits from early exercise, exercising an option may not be optimal.

[6]Some exchanges facilitate these strategies by imposing daily fee caps for floor market makers and other floor traders engaging in them. 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 the PHLX options exchange. More than two-thirds of dividend play transactions in our sample are executed on PHLX. Dividend play is not the only strategy that benefits from fee caps. For example, PHLX offers similar fee caps for five other arbitrage strategies.

[7]The lower bound compares the average last cum-dividend date dollar trading volume in call options to an average across all days in our sample, while the upper bound compares to the average volume in a week prior to the last cum-dividend date.

that is, public customers. The Customer flag has been used in the literature to classify retail investor trades.[8] The new OPRA trade flags that we exploit in this paper reveal that most of the trading volume attributable to the dividend play strategy originates on the exchange trading floor. It is highly unlikely that retail investors trade from the floor and, more generally, engage in the dividend play strategy. Most likely, the OCC misclassifies trades of market makers and other sophisticated arbitrageurs as Customer trades. We therefore caution against the use of the Customer classification in identifying retail trades. Trade misclassification could lead to significant errors because the position sizes and trading volume that correspond to dividend play are extraordinarily large.

Expected profits to floor market makers and other arbitrageurs from the dividend play trades have been growing rapidly during the recent retail investor boom. Most of this profit derives from the sheer increase in open interest due to investor inflow, coupled with a higher fraction of options that are left unexercised on dividend play dates. For the latter, we document that in the cross-section, arbitrage activity is higher in contracts with a larger share of retail activity in the preceding week, as measured by SLIM share. Overall, traders engaging in dividend play behave like unconstrained arbitrageurs in that they are able to establish long-short positions of extremely large sizes.

There is, however, one striking pattern that emerges from our examination of dividend play transactions. Market makers and other arbitrageurs harvest only 50% of available profits, leaving the rest on the table.[9] Moreover, if arbitrageurs decide to participate in dividend play in a given contract, they harvest close to all available profits in that contract. Intriguingly, however, there are many contracts, similar in terms of observable characteristics, that do not attract *any* dividend play activity. Even when matching on profitability characteristics, we find that there is a large variation in arbitrageur participation. This is extremely puzzling. Market makers' daily fee on dividend play trades is capped by most exchanges on which dividend play trades take place, so additional fees are not a plausible explanation for leaving money on the table. Furthermore, other trading costs are quite low because dividend play transactions are typically prearranged by pairs of market makers and, in our sample, actual transaction prices are close to the midpoint of the bid-ask spread. We discuss the role of transaction costs in detail in Section 5.3.

We rule out further potential explanations of our money-left-on-the-table puzzle such

[8] Measuring retail investor participation in options had been a long-standing problem. In an influential paper, Pan and Poteshman (2006) propose to use the Customer classification. Their sample runs from 1990 to 2001, and during that time period the dividend play trade represented a much smaller part of the market trading volume than in our sample (see Pool, Stoll, and Whaley (2008)).

[9] Table 14 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.

as capital constraints and limited attention of arbitrageurs. The former is ruled out because market maker exposure is computed at a ticker level, and so the large long and short positions in contracts on the same ticker, required for the dividend play trade, are 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 the trade.[10] Additionally, it is possible that there is a stigma associated with this strategy, since it is frowned upon by the U.S. Securities and Exchange Commission (SEC).

Finally, we take advantage of the granularity of our data to make inferences about the number of arbitrageurs simultaneously engaging in dividend play in each contract. Strikingly, in the majority of the contracts we observe that the entire profit is captured by a single arbitrageur. Across all contracts, the number of arbitrageurs simultaneously participating in dividend play in a given contract rarely exceeds three. This is surprising, given that there are no obvious restrictions on free entry of other arbitrageurs who engage in dividend play in other contracts. Partly because of their sheer size and distinct execution patterns, dividend play trades are easy to detect in transaction-level data, which is available to arbitrageurs in real time from standard datafeeds (e.g., OPRA or data provided by exchanges). Arbitrageurs are therefore fully aware of the presence of other arbitrageurs in a given contract before they decide to engage in dividend play in it. It is unclear why this deters them from competing for profits in the contract. Both this result and the money-left-on-the-table puzzle highlight that forces of arbitrage do not always follow the textbook description in practice.

Our paper offers several policy implications. First, retail brokerages should provide additional tools to help investors make sophisticated financial decisions that often accompany trading in options. For example, retail investors may not know how to compute an expected price of an option right after the underlying stock goes ex-dividend, and providing this valuation or making early exercise automatic, if optimal, would reduce the option exercise mistakes that we document in this paper.

Second, we discuss the difficulties of devising effective regulation in the derivatives 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,[11] which resulted in much lower trading volumes on dividend play dates. However, the recent dramatic increase in options trading by inexperienced retail investors appears to have led to a resurgence of the strategy, whereby arbitrageurs have found a way to circumvent the barriers created by the SEC rule.

[10] For example, a human error in the dividend play strategy inflicted a $10 million loss on Bank of America Merrill Lynch. See https://www.reuters.com/article/us-usa-options-apple-idUSKBN0IQ2FA20141106.

[11] See https://www.sec.gov/rules/sro/occ/2014/34-73438.pdf.

Our paper relates to the literature on optimal option exercise by investors. It has been previously documented that not all American options are exercised rationally (e.g., Poteshman and Serbin (2003)). 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. More generally, the literature has documented a variety of mistakes that retail investors make (see, e.g., Barber and Odean (2001), Calvet, Campbell, and Sodini (2007), and Barber and Odean (2013)).

The closest papers to ours are Hao, Kalay, and Mayhew (2009) and Pool, Stoll, and Whaley (2008) that show how arbitrageurs exploit investor mistakes by engaging in dividend play trades. Our measure of arbitrageur activity in dividend play, based on the new OPRA codes, is more accurate, and our transaction-level data is more granular. This allows us to document puzzling reluctance of market makers and other arbitrageurs to harvest arbitrage profits in certain contracts and to infer that the number of arbitrageurs participating in dividend play trades is very small. Furthermore, Hao, Kalay, and Mayhew (2009) and Pool, Stoll, and Whaley (2008) were written before the 2014 SEC regulation that attempted to curtail dividend play activity. We document how market participants circumvented the regulation. de Silva, Smith, and So (2022) explore another strategy that market makers used in order to profit from retail trading in options, namely, trading around earnings announcements, and document that market makers benefit from retail investor trading, especially during the recent retail investor trading boom.

Our paper is also related to the limits-to-arbitrage literature, such as Gromb and Vayanos (2002) and Shleifer and Vishny (1997). Fardeau (2021) explores arbitrageur behavior in a setting with a limited number of arbitrageurs. In contrast to the studies of strategic arbitrageurs for whom price impact is a primary consideration, in our setting, price impact is not a concern. Buy and sell orders in a dividend play strategy are entered at the same time and are perfectly balanced. The dividend play trade is best compared to an IPO lottery: The windfall gain (from subotimal exercise of a call option contract) is effectively fixed, and it is divided pro-rata between arbitrageurs who participate in the dividend play. Entry of additional arbitrageurs dilutes the share of the gain received by the existing option writers.

The rest of the paper is organized as follows. Section 2 describes our dataset and documents the resurgence of dividend play in the recent data. Section 3 discusses measures of retail trading in options. Section 4 investigates retail investors' failure to exercise options when it is optimal to do so and describes the arbitrageurs' dividend play strategy that exploits these mistakes. Section 5 documents the puzzling behavior of market makers who leave money on the table, and attempts to rationalize this behavior. Section 6 makes several policy recommendations and Section 7 concludes. The Appendix presents technical details and robustness checks.

2 Dividend play

In this section, we establish stylized facts that suggest that the inflow of retail investors has affected the behavior of arbitrageurs in the options market. Specifically, we focus on a specific arbitrage strategy, dividend play, for which we can accurately identify trades of arbitrageurs. The inflow of inexperienced retail investors who do not exercise their options optimally has made this strategy more profitable for the arbitrageurs.

2.1 Dataset

Our options data comes from two main sources: OPRA and OptionMetrics. Transaction-level data from OPRA LiveVol is provided by CBOE. The data ranges from November 4, 2019, to July 31, 2021. It contains all transactions in index, ETF, and equity options on 16 U.S. options exchanges. In our analysis, we exclude index options and focus on single-name options on equities and ETFs. We use daily option price, volume, and open interest data from OptionMetrics, available 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.[12]

Our data-cleaning procedure is as follows. Following the literature, we remove the first 15 and last 10 minutes in the day, canceled trades, trades with nonpositive size or price, and negative spread (the difference between best ask and best bid). We only keep trades for which trade price is above (best bid minus spread) and below (best ask plus spread). We aggregate trades of the same contract with the same quote time, exchange ID, trade price, and trade type into one line.

We complement this data by the daily options trading volume data from the OCC, which is available from the OCC website in open access for the 24 months preceding its publication date. In our tests, we restrict our sample to the same time period, November 4, 2019, to July 31, 2021, as our OPRA sample.

For all stock data, we use CRSP. We obtain dividend history, stock prices and returns, outstanding shares, and rolling monthly volatility of daily returns. We rely on the SecId-PERMNO crosswalk provided by WRDS to link CRSP with OptionMetrics.

[12]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.

To construct a measure of internalized volume in the underlying, we rely on FINRA OTC Transparency data, which reports stock-level trading volume on automated trading systems (ATS), typically referred to as dark pools, and non-ATS, with the latter representing internalized trades. This data is available from April 2016, by security and venue. Securities are split into NMS Tier 1, Tier 2, and OTCE. Details are on the FINRA website: `https://otctransparency.finra.org/otctransparency/AtsIssueData`.[13]

For our WallStreetBets count measure, we download all comments from Daily Discussion and What Are Your Moves Tomorrow threads on the `WallStreetBets` subreddit of `reddit.com`. This data is collected via PRAW, which is a Python API toolkit to access `reddit.com` for the period from October 1, 2019, to June 30, 2021. In particular, we download all the comments (original posts and reactions to them) for each daily Daily Discussion and What Are Your Moves Tomorrow thread.[14]

Finally, we utilize Robintrack portfolio data, which is provided by Robinhood in intraday snapshots and covers the period of May 5, 2018, to August 13, 2020.

2.2 Resurgence of dividend play

Daily trading volume in options on high-dividend stocks in the U.S. exhibits an intriguing seasonality, illustrated in Figure 1 for the UPS case. The spikes in trading volume apparent from the figure occur every quarter, on the last cum-dividend date, that is, the day before UPS pays a dividend. The average daily traded notional for UPS is $125.3 million on the last cum-dividend dates and only $2.5 million on the remaining dates. This pattern is common for options on high dividend paying stocks. Appendix A.3 presents more examples.

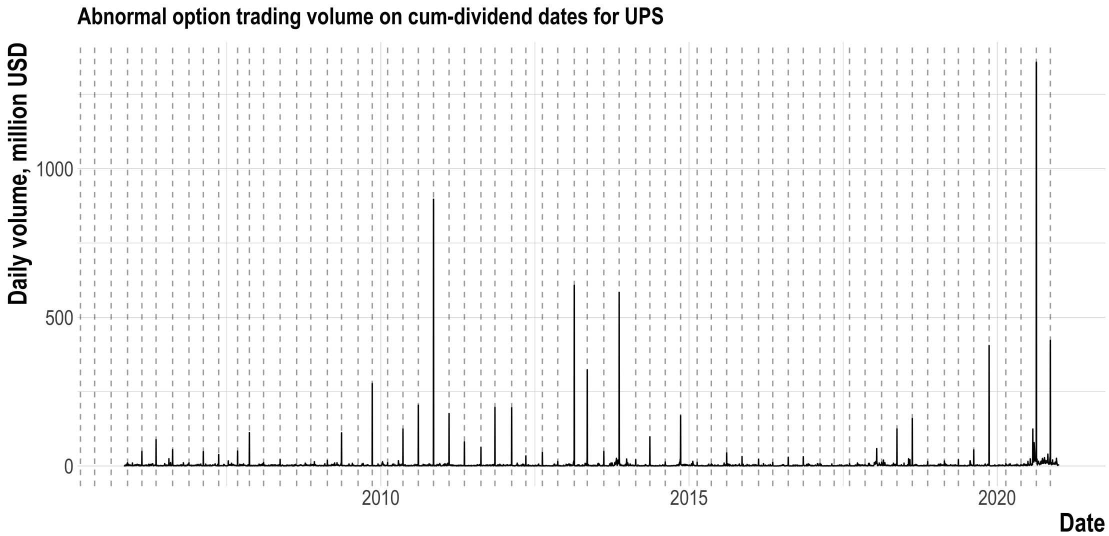
On the last cum-dividend dates, market makers and other arbitrageurs engage in an arbitrage trade known as the dividend play. This strategy is available only for transactions originating from the floor of the exchange,[15] or, in other words, only to the market participants who must be physically located 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 11), which resulted in much lower trading volumes on dividend

[13] See also FINRA's regulatory notice requiring reporting of this data: `https://www.finra.org/rules-guidance/notices/15-48`.

[14] Some dates are missing due to retrieval limitations on `reddit.com`. We interpolate between the neighboring dates to fill in those values.

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

Figure 1: Abnormal trading volume on last 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 the last cum-dividend dates.

play dates. 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.

The goal of the dividend play strategy is to take advantage of inattentive investors who fail to exercise their call options on dividend-paying stocks when it is optimal to do so. It is optimal to exercise a call option if the value of exercising it on the last 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 involves an application of the Black-Scholes-Merton formula or a more sophisticated option pricing method, which is typically difficult for novice retail investors. Alternatively, some retail investors may be unaware of the possibility of early exercise or are simply inattentive.[16] Since a fraction of in-the-money call options remains suboptimally unexercised, the writers of these options would not be asked to deliver the stock and would profit from this inattention. It is a zero-sum game.

If all in-the-money call option contracts on a stock have been exercised on or before the last cum-dividend date, *all* holders of short positions in the same contracts receive a request to deliver the stock. If some contracts are left unexercised, however, the OCC randomly "assigns" short positions that must deliver the stock. The unassigned holders simply hold

[16]There might be other reasons why investors do not exercise, such as costs of unwinding more complex strategies. Hao, Kalay, and Mayhew (2009) show that dividend play profits outweigh such costs in most cases.

on to their options and profit from a capital gain. Market makers and other 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 underlying.[17] They exercise all long positions and deliver on all assigned short positions. Since some fraction of the options remains unassigned (owing to suboptimal exercise by investors), arbitrageurs then 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 a counterparty to each other on their arbitrage positions.

Table 1 illustrates the mechanics of the dividend play strategy by means of an example. Suppose there is 1 call option contract outstanding, and it is optimal to exercise it.[18] Case 1 corresponds to the case when the option is exercised, the holder of the short position is assigned to deliver the underlying, and there is no profit for the 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 for sure.[19] Now consider the entry of a market maker. The market maker attempts to pocket most of the potentially harvestable profit of $500. To do so, the market maker buys and simultaneously sells 100 contracts and exercises all of their long positions. The probability of assignment increases, but, because of the OCC's random assignment, with probability 100/101, the market marker 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 market maker 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 what follows, we detect dividend play activity at a contract level in the full sample and characterize its importance relative to the overall trading volume on the last cum-dividend dates.

[17]The current SEC rule prohibits simultaneous buying and selling of the same contract. See footnote 11.

[18]Appendix A.4 provides another example, in which there are multiple contracts outstanding, some of which are exercised optimally and some are not.

[19]We assume that, as in the data, each option contract is for 100 shares of underlying.

Table 1: 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. market maker (F)	Gain per share (G)	Expected gain orig. option writer (E*G*100)	Expected gain market maker (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**	
Market maker	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 last cum-dividend date, and OI_t stands for the open interest on date t.

2.3 Arbitrageur activity in dividend play strategy

We first present our novel measure of arbitrageur activity in the dividend play strategy. Through fee caps, exchanges incentivize dividend play strategies to originate from the physical floor. To construct our measure of dividend play activity, we therefore exploit OPRA new detailed trade type flags 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.1 for a more detailed description.) Other floor trade types, used infrequently in our sample are MLCT, MSFL, SLCN, TLFT, and TLFT. We analyze this measure in detail below and show that it accurately captures dividend play trades in our sample. To our knowledge, this is the most precise measure of arbitrageur activity in the dividend play strategy in the literature, which typically uses trading volume on the last cum-dividend dates in excess of the past average volume.[20]

In our data, we see bursts of simultaneous buy and sell activity in neighboring-strike call option contracts, executed usually within several seconds, all coming from the floor. In an effort to reduce operationally risky dividend play trades, in 2014 the SEC began to forbid the market makers and other arbitrageurs form simultaneously buying and selling the same contract. Market participants have adjusted their trading strategies, and

[20]Strictly speaking, we should define our measure of arbitrageur activity in dividend play as *abnormal* floor volume on the last cum-dividend dates, but our measure is simpler and our results stay virtually the same if we use the abnormal floor volume measure instead.

Table 2: Characteristics of activity on last cum-dividend dates

	Average ticker dollar volume ($ million) on		**Total market dollar volume share (%) on**	
	last cum-dividend date	any other date	last cum-dividend date	any other date
	(1)	(2)	(3)	(4)
Panel A. Option type				
Call	24.9	1.5	92.8	54.3
Put	2.1	1.4	7.2	45.7
Panel B. Moneyness				
In the money	26.8	0.8	80.6	18.3
At the money	4.8	2.0	17.6	70.2
Out of the money	0.6	0.4	1.8	11.4
Panel C. Trade size				
Small	1.5	0.7	5.6	27.5
Large	31.0	2.8	94.4	72.5
Panel D. Floor trade				
Yes	48.2	0.9	76.9	6.5
No	6.1	2.4	23.1	93.5
Panel E. Exchange				
PHLX or BOX	24.6	0.5	79.7	15.0
Any other	5.5	2.3	20.3	85.0

This table compares option trading activity for dividend-paying tickers (2,153 stocks and ETFs) on the last 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" are contracts for which this value is between -0.1 and 0.1, and "out of the money" are 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.

they now simultaneously buy and sell neighboring contracts, which ultimately achieves the same objective. The trades are typically prearranged by pairs of market makers or other arbitrageurs. 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 accurately captures arbitrageur activity in the dividend play strategy.

Table 2 presents descriptive statistics of trading activity on the last cum-dividend compared to any other dates for dividend-paying stocks and ETFs. We see a considerable difference in floor trading volume and volume of large trades on the last cum-dividend dates relative to other dates. Moreover, on the last cum-dividend dates we see colossal spikes in volume on two exchanges that cap fees for the dividend play strategy: PHLX and BOX. Breaking 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 utter size of the dividend play positions is astonishing. Clearly, market participants have been able to circumvent the regulators' attempt to curtail this strategy, and profit-driven incentives of arbitrageurs have pushed the trading volume in dividend play to new highs.

3 Measures of retail investor popularity

We hypothesize that the resurgence of the dividend play strategy is related to the inflow of amateur retail investors who do not exercise their options optimally. BPS show that during the pandemic-fueled retail investor boom, retail trading volume more that 48% of the total market trading volume. To test our hypothesis, we need a measure of retail trading activity in the options market.

3.1 Customer classification as a measure of retail activity

The first candidate measure to use is the Customer trading volume from the OCC, which researchers use to identify retail investor transactions. To shed light on the accuracy of the OCC Customer volume in capturing retail activity, we contrast the OCC data with our measure of arbitrageur activity in the dividend play strategy, namely, floor trades in pertinent call options on the last cum-dividend dates in the OPRA data. We show that in the OCC data a significant fraction of the trading volume in the dividend play strategy is attributed to Customer, that is, a public customer. It is highly unlikely that public customers/retail investors participate in the dividend play strategy, which is a sophisticated strategy that involves very large long and short positions and that is usually executed on the exchange floor.

The following example illustrates the statement above. Table 3 provides the trading volumes from in OPRA and OCC for Microsoft option contracts during a short window that includes the last cum-dividend date on which the dividend play strategy is executed (November 17, 2020). In the OPRA data, we see a massive spike in floor trades, a signature of the dividend play strategy. Notice that the growth in volume in the OCC data on the dividend play date comes from Customer and Market Maker categories. To provide further evidence that the abnormal trading volume on November 17, 2020, reflects dividend play activity, Table 9 in the Appendix breaks the OCC volume by exchange and compares it to a "normal" trading day (Table 10 in the Appendix). We see a clear increase in volume on several exchanges on the last cum-dividend date, in particular, on BOX and PHLX. Not all exchanges in the U.S. facilitate the dividend play strategy, and the fact that the abnormal volume comes from only two exchanges suggests that the dividend play strategy is at work. Again, a large fraction of the increase in volume comes from Customer transactions. These examples easily generalize. Recall from Table 2 that 80% of trading volume on the last cum-dividend dates comes from floor trades. In the OCC data, much of this increased volume is attributed to Customer transactions. We therefore caution researchers against using the Customer flag in the OCC data to classify retail trades and suggest that papers relying on this flag should at least remove dividend play transactions from their sample.

Table 3: Trading volume on normal and dividend play dates: An example

	OPRA				OCC		
Date	Volume (mln contr.)	Volume ($ mln)	Floor (mln contr.)	Floor ($ mln)	Customer (mln contr.)	Total Off-Floor (mln contr.)	Market Maker (mln contr.)
11/09/2020	289,265	131.522	245	0.113	264,635	279,953	327,241
11/10/2020	647,395	236.356	5,824	5.172	497,693	575,918	551,594
11/11/2020	303,515	153.100	638	1.824	224,568	236,950	247,488
11/12/2020	311,166	117.720	11,902	5.565	200,805	232,980	243,654
11/13/2020	252,664	101.164	3,746	2.320	195,420	205,790	214,086
11/16/2020	291,008	151.184	6,540	5.295	187,743	198,631	209,103
11/17/2020	464,769	1,674.898	176,323	1,452.843	268,631	355,234	436,364
11/18/2020	347,618	245.097	25,591	24.742	258,316	273,524	272,092

This table reports the trading volumes in Microsoft option contracts in the OPRA and OCC data on normal days (from 11/09/2020 – 11/16/2020 and 11/18/2020) and on the last cum-dividend date (11/17/2020). For the OPRA data, we list both contract volume (mln contracts) and dollar volume. In the OCC data, the trading volume (mln contracts) is broken by Customer, Total Off-Floor, and Market Maker. In the OCC volume data, each transaction is counted twice. The total OCC market volume adjusted for double-counting is given by the average of Total Off-Floor and Market Maker volumes. The OPRA and OCC volumes (adjusted for double-counting) do not exactly match because of our data filters (see Section 2).

We next turn to defining our two alternative, transaction-level measures of retail trading in options, as well as additional stock-based measures of retail investor popularity.

3.2 Measures of retail investor trading in options

All options trades in the U.S. must be executed on exchanges, and our transaction-level data contains every transaction in the U.S. options exchanges during the sample period. Using this data, we construct two measures of retail investor trading. Our first measure, often used in the industry, is the volume share of small trades (up to 10 contracts), *Small Share*. One could compute it as a frequency share and as a trading volume share. We adopt the latter definition, as it would be more relevant for assessing the influence of retail traders on asset prices. We compute it daily and, if necessary, we aggregate it to a ticker level using traded volumes.

We follow BPS and exploit unique features of our data, which allows us to construct a measure of retail investor trading in options. (Ernst and Spatt (2022) and Hendershott, Khan, and Riordan (2022) propose the same methodology.) We take advantage of the new trade type codes, introduced by OPRA on November 4, 2019, which provide a detailed classification of transaction types. This reporting requirement is significantly more detailed than its predecessors; hence, we can construct our measure starting only from November 4, 2019.

A highly publicized advantage to investors for having their orders routed to a wholesaler by a retail brokerage is that the wholesaler promises a price improvement to the customers, that is, the execution price that is better than the best quoted price, known as National Best Bid and Offer, or NBBO. To meet this commitment, wholesalers often execute retail orders through a "price improvement auction" mechanism offered by options exchanges. The orders entered into an auction typically originate from the same wholesaler. A wholesaler can "bring" a *paired* order to the exchange to be "internalized" as long as the order is exposed to other market participants on that exchange. Market participants ("responders") have a window of time to respond with a better price for the agency order (hence, the name "price improvement auction"), which could lead to the wholesaler losing the trade.[21] In practice, the fees are stacked against responders, and it is often prohibitively expensive to break up one of these paired trades. See BPS for more details. We borrow their acronym SLIM to denote trades that went through a single-leg price improvement mechanism, and use the volume share of these trades, *SLIM Share*, as a measure of retail activity.

3.3 Measures of retail investor popularity of a stock

We follow the literature to construct several measures of retail popularity of a stock. Although these measures are all for stocks, we hypothesize that cash-constrained, risk-loving retail investors who bet on a stock may also trade options on this stock, which is a leveraged bet on the stock.

Our first measure of retail investor popularity, introduced in BPS, is a fraction of stock trading volume that was internalized on private trading platforms, off lit exchanges. We utilize FINRA OTC Transparency data, specifically stock-level non-ATS trading volume, which is internalized trading volume.

Our second measure of popularity is a ticker-level *WallStreetBets mentions*, which is based on `WallStreetBets` subreddit of `reddit.com`. We search for each unique historical ticker from CRSP in all the comments, and then sum the number of mentions for each date. We search only for capitalized tickers, which the reddit audience typically uses. Since we might omit any lowercase mentions, and since we do not cover other threads of the forum (such as occasional megathreads), our measure provides a lower bound for ticker popularity.

[21] Specifically, we use the same OPRA type SLAN, which stands for single-leg non-ISO price improvement auctions. See Appendix A.1 for a description.

Our final measure is Robinhood's breadth of ownership, used in, for example, Welch (2022), based on Robintrack portfolio data. Our measure of Robinhood's breadth of ownership is the number of users holding a stock as of the last intraday snapshot.

4 Retail investor presence and failure to exercise

4.1 Failure to exercise and dividend play profits

In this section, we compute exploitable profits from a dividend play strategy. Some of these profits 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: (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

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

where $t-1$ is the day before the last 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 is 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 last 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 last cum-dividend day (OI_{t-1}) and fraction not exercised (f_t) are available from OptionMetrics. In rational and frictionless markets, we expect that $f_t = 0$ if $EEV > 0$.

The early exercise value is model-based, and we rely on the Black-Scholes-Merton option pricing formula to compute it.[22] 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: $S - K - c_{ex}$.[23] The details of the computation of c_{ex}

[22] To ensure 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 hold in that sample and are available upon request.

[23] Note that this definition is from Pool, Stoll, and Whaley (2008), and it is equivalent to the definition in

are in Appendix A.5.

In the following analyses, we restrict our sample to call option 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.6, and Table 12 in the Appendix presents the descriptive statistics for our dividend play sample.

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

$$Y_{c,t} = \beta_1 \times share^{SLIM}_{c,t} + \beta_2 \times share^{small}_{c,t} + \boldsymbol{\gamma}' X_{c,t} + \alpha_{i,t} + \varepsilon_{c,t} \tag{2}$$

where, for each contract c on the last 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 (1). $share^{SLIM}_{c,t}$ is the average dollar volume share in SLIM trades (one of our measures of retail activity) over one trading week before the last cum-dividend date t, and $share^{small}_{c,t-h}$ (another measure of retail activity) defined as the average dollar volume share of trades up to 10 contracts over one trading week before the last cum-dividend date t. In some specifications we also use ticker-level measures of retail investor popularity such as *Internalized volume in underlying*, which 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, and *WSB mentions, log* (the logarithm of the number of times that ticker i was mentioned on the `WallStreetBets` forum on date t). Our vector of controls $X_{c,t}$ includes the following contract-level variables: log OI, EEV, log dollar trading volume, relative spread, implied volatility, moneyness, and days to expiration.[24] Our specification also includes the ticker by date fixed effects $\alpha_{i,t}$. Our measures of retail investor trading are computed over one trading week before the last cum-date because the new generation of retail investors has a strong preference for options expiring within a week (as documented in BPS). Table 15 in Appendix A.9 presents an alternative specification in which we measure retail trading over two weeks preceding the last cum-dividend date.

Panel A of Table 4 reports the results of the regression in (2), with the fraction of open interest unexercised as the outcome variable. We find that there is a strong positive relationship between retail investor trading and the fraction of options that were suboptimally not exercised on the last cum-dividend day. We measure retail investor trading in two different ways, and both variables prove to be strong predictors of failures to exercise the option. A one standard deviation increase in the share of SLIM or small trades in the contract in

Hao, Kalay, and Mayhew (2009). The latter uses dividend instead: $Dividend - c_{ex} + S_{ex} - K$.

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

the week preceding the last cum-date increases the fraction unexercised by approximately 1 percentage point, depending on the specification. This result is robust, and the magnitudes of the coefficients of interest do not change much as we relax the specification of fixed effects and use ticker-level controls instead (columns (1)–(3) and (5)–(6)).

Another measure of retail investor activity that strongly predicts the fraction left unexercised is the internalized volume volume in the underlying stock, measured over the preceding trading week. Although this is a measure of retail investor trading in the underlying stock, it is conceivable that it is correlated with retail investor trading in options on that stock. A call option is a leveraged position in a stock, and we hypothesize that some risk-loving retail investors may wish to trade options rather than the stock. We find that a one standard deviation increase in the internalized volume share increases the fraction of options left unexercised by 1.6 percentage points (column (6)). Using another measure of retail trading discussed in Section 3.3, WallStreetBets mentions, we also find a positive and significant relationship between retail investor interest and the fraction unexercised. These results are consistent with BPS, who also document that the new generation of retail investors is less likely to exercise options optimally.

One may argue that an alternative explanation for our findings is that the failures to exercise the options may be driven by transaction costs that make exercise impractical. To rule out this explanation, we restrict the sample to the top EEV tercile, the most profitable contracts to exercise (column (5)). We find that the size of the effect increases significantly relative to our base case, implying that investor mistakes are a more likely driver of our findings. Another possible alternative explanation is that investors hold the call options in our sample as part of a sophisticated strategy, and exercising the option breaks one leg of the strategy. Although this is possible and we do see mentions of a number of options strategies on `WallStreetBets`, we believe that the new generation of retail investors that drives our results are financial novices and relatively few of them engage in options strategies. Furthermore, to engage in such strategies, investors must qualify for a certain level of investment proficiency that is required by investing platforms.

Panel B of Table 4 considers the regression in (2), but with potential profits as the outcome variable. It reveals that the coefficients on all measures of retail activity are positive and significant: A one standard deviation increase in SLIM Share corresponds to around $4,200 higher profit *per contract*. In other words, the higher the retail activity in a contract in a week preceding the last cum-dividend date, the more profitable it is for arbitrageurs to engage in a 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.

Table 4: 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	5.946***		5.791***	6.724***	5.792***	6.447***
	(3.48)		(3.37)	(2.66)	(3.38)	(3.58)
Small Share		4.438**	4.137**	10.365***	4.768**	4.569**
		(2.17)	(2.02)	(3.69)	(2.46)	(2.05)
Internalized volume in underlying						31.763*
						(1.88)
WSB mentions, log						0.528*
						(1.85)
Observations	21,105	21,105	21,105	6,942	21,105	19,134
Adjusted R-squared	0.237	0.237	0.238	0.345	0.214	0.213
Panel B. Potential profits, log U.S. dollar						
SLIM Share	1.465***		1.435***	1.701***	1.531***	1.564***
	(7.24)		(7.09)	(5.22)	(7.54)	(7.58)
Small Share		0.944***	0.873***	2.394***	0.887***	0.870***
		(3.27)	(3.08)	(4.09)	(3.27)	(3.04)
Internalized volume in underlying						2.991*
						(1.72)
WSB mentions, log						0.027
						(0.73)
Observations	21,105	21,105	21,105	6,942	21,105	19,134
Adjusted R-squared	0.307	0.304	0.308	0.314	0.286	0.292
Sample	All	All	All	Top EEV tercile	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 (2) 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 last 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 the `WallStreetBets` forum. In Panel B, contract controls include log dollar trading volume, relative spread, IV, moneyness, and 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. Standard errors are clustered by ticker and date. Robust t-statistics are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

5 Money left on the table: A puzzle

In this section, we show that arbitrageurs engaging in the dividend play strategy leave money on the table by failing to capture arbitrage profits in some call option contracts. We explore the determinants of this puzzling behavior and present suggestive evidence that arbitrageurs may behave non-competitively.

5.1 Case study

November 11, 2020, was the last cum-dividend date for UPS, a high-dividend paying stock, and a number of call options on UPS were deeply in-the-money and optimal to exercise on that day. Table 5 focuses on a pair of such contracts, both expiring on November 20, 2020.

We first compare the trading volume in the contracts on November 11, 2020. Notice that

Table 5: Case study of arbitrageur activity: Two UPS call options on the last cum-dividend date

	Strike	EEV	OI (t-1)	Moneyness	Spread	Fraction unexercised	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

the trading volume in Contract 2 exceeds that in Contract 1 by two orders of magnitude. Notice also that Contract 2 has most of the orders from the trading floor on that day, while Contract 1 has zero. We also see characteristic bursts of floor orders in the transaction-level data for Contract 2. This means that market makers (or other arbitrageurs) engaging in a dividend play trade exploited Contract 2 but not Contract 1.

Why did the arbitrageurs leave money on the table in Contract 1? The contract had a high EEV and a large fraction unexercised. Using equation (1) to compute the arbitrageur's forgone profits from not participating in Contract 1, we arrive at $1,945 \times 0.76 \times 0.29 \times 100 \approx \$42,900$, a significant sum.[25]

Trading costs do not explain the market participants' reluctance to trade Contract 1. First, exchanges offer daily fee caps for the dividend play strategy, and so if market makers (or other arbitrageurs) entered in Contract 2, they should have also entered in Contract 1. Second, contract bid-ask spreads in Table 4 are very similar. In the regression framework that follows, we further control for the option contract's liquidity and show that trading costs do not explain why arbitrageurs forgo profitable opportunities.

It is puzzling that arbitrageurs fully exploited the arbitrage opportunity in Contract 2 but not in Contract 1. In the following section, we show that this is the general pattern in our sample. The unexploited profit in Contract 1 accrued to the writer of this contract, who could be a market maker or perhaps a retail investor. The latter is less likely because retail brokerages take an automated action to close short positions that have dividend risk on behalf of their clients.[26] Appendix A.10 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 own profit. But it is

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

[26]Since each option contract is for 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.

puzzling that other market makers or arbitrageurs would not wish to exploit Contract 1 and reap arbitrage profits.

Table 14 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. We aggregate our data to the ticker level and 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 \$51 million, whereas the total amount of forgone profit stands at \$67 million. For a virtually riskless arbitrage strategy, the amount of money left on the table is striking.

In the full sample, the values of harvested and forgone profits are \$96 million and \$97 million, respectively; hence, the total potential arbitrageur profit from the dividend play strategy is around \$200 million. To put the total amount of profits from this strategy into perspective, it is useful to compare these numbers to market maker profits from other strategies that exploit retail investor mistakes. de Silva, Smith, and So (2022) study retail investor trading in options around earnings announcements and find that this trading amounts to a wealth transfer of \$360 million from retail investors to market makers in their sample that runs from 2010 to 2021.

Table 14 does not reveal any particular pattern in harvested and 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.

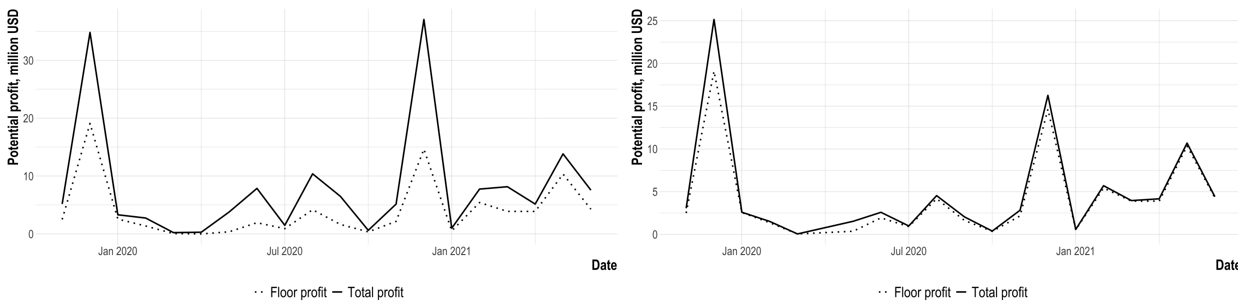
5.2 Drivers of arbitrageur activity

To examine drivers of arbitrage activity, we start by contrasting potential and harvested profits from the dividend play strategy on the last cum-dividend dates in our sample. Figure 2 presents potential profit of the dividend play strategy in all outstanding contracts, computed using equation (1), and profits harvested by floor traders. It emerges from Panel (a), that a large fraction of potential profit, about 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, market makers and other arbitrageurs selectively exploit profitable contracts, capturing almost 100% of exploitable gains, but forgo arbitrage profits in contracts they do not enter.

The tendency of market makers and other arbitrageurs to leave money on the table

Figure 2: 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 depicts the potential profit from the dividend play strategy, and the dashed plot depicts the profit harvested by floor traders (arbitrageurs).

in some profitable arbitrage opportunities is puzzling. In what follows, we try to understand the features on the contracts into which market participants are likely to enter.

The total exploitable profit is a calculated quantity, not known with certainty on cum-dividend dates. The determinants of potential profit from a contract, however, are well-understood (see equation (1)), and projected fraction suboptimally unexercised is one of them. As shown in Table 4, this fraction is increasing in retail investor popularity. We therefore examine whether market makers (and other 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:

$$share^{floor}_{c,t} = \beta_1 \times share^{SLIM}_{c,t} + \beta_2 \times share^{Small}_{c,t} + \boldsymbol{\gamma}' X_{c,t} + \alpha_{i,t} + \varepsilon_{c,t}, \tag{3}$$

where the regressors are as those in our previous specification (2) and the outcome variable is now the share of floor trades, which are predominantly market maker dividend play trades, in contract c on date t. Table 6 reports the results of the regression.

Table 6 reveals that arbitrageur activity, as measured by floor trading share, is positively related to both measures of retail investor trading over the preceding week. This means that market makers and other arbitrageurs are aware of and do exploit suboptimal exercise strategies of retail investors by engaging in more dividend play trades in contracts that have experienced elevated retail investor activity. The effects of SLIM and Small Shares are statistically significant across all specifications. The magnitudes of the effect can be interpreted as follows. A one standard deviation increase in SLIM (Small) Share increases the share of floor trading by about 0.048*100*0.16=1 (0.288*100*0.18=5) percentage point(s).

Table 6: Arbitrageur activity and retail investor popularity

	Floor trading share on last cum-date				**D(floor share > 0)**	**Floor trading volume, log**
	(1)	(2)	(3)	(4)	(5)	(6)
SLIM Share	0.048***		0.037**	0.057*	0.036*	0.547***
	(2.59)		(2.11)	(1.87)	(1.89)	(3.43)
Small Share		0.288***	0.286***	0.123*	0.301***	4.025***
		(6.41)	(6.39)	(1.66)	(6.53)	(8.53)
Observations	21,105	21,105	21,105	6,942	21,105	21,105
Adjusted R-squared	0.403	0.408	0.408	0.473	0.397	0.475
Sample	All	All	All	Top EEV tercile	All	All
Contract controls	Y	Y	Y	Y	Y	Y

This table reports estimates of (3) in our dividend play sample. 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 last 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 last cum-dividend date. Contract controls include log OI, EEV, log dollar trading volume, relative spread, IV, moneyness, and 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, * p<0.1.

The magnitude is similar for the extensive margin (column (5)): A one standard deviation increase in SLIM (Small) Share increases the probability of floor entry by around a half (five) percentage points. If, rather than the floor trading volume share, we use floor trading volume in a contract on the last cum-dividend date as an outcome variable, we again see strong effects of retail investor participation in the contract (column (6)). All of these effects are highly statistically significant.

If we restrict our sample to the most profitable contracts (column (4)), the relationship between retail trading and arbitrageur activity weakens and becomes marginally significant. This is surprising, given that our earlier analysis documents that retail investors do not exercise more profitable contracts more (column (4) in Table 4), and therefore market makers and other arbitrageurs seem to leave money on the table in more profitable dividend play trades.

To ascertain the robustness of our results, we pursue an alternative empirical strategy, based on propensity score matching. Such empirical exercise brings us closer to the case study presented in Table 5. Matching is a natural strategy in our setup because the set of characteristics on which one should match options to keep the expected profitability constant is well understood. We again study the relationship between the floor trading share on cum-dividend date and retail popularity. However, here we isolate contracts with high retail popularity (top decile of SLIM Share or Small Share of 100%)[27] and construct the control

[27]Results are robust to the choice of percentiles. Small Share of 100% corresponds to the 57$^{\text{th}}$ percentile in our sample.

group of contracts matched on profitability characteristics from contracts with low retail popularity. In the basic set of characteristics, we use open interest, early exercise value, and moneyness. We also report results with the characteristics extended to relative spread and underlying price. The corresponding covariate balance plots are presented in Appendix A.11.

Table 7: Arbitrageur activity and retail popularity: Matched contracts

	Floor trading share on dividend play date							
		Matched		OLS		Matched		OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D(SLIM Share top decile)	0.035***	0.039***	0.032***	0.019*				
	(2.82)	(3.67)	(3.10)	(1.79)				
D(Small Share = 1)					0.033***	0.035***	0.042***	0.033***
					(3.99)	(5.03)	(6.19)	(4.82)
Observations	21,105	21,105	21,105	21,105	21,105	21,105	21,105	21,105
No. neighbors	1	10	10		1	10	10	
Short controls	Y	Y	Y	Y	Y	Y	Y	Y
Extended controls	N	N	Y	Y	N	N	Y	Y

This table reports the results of propensity score matching estimation and OLS estimates for the same set of contract characteristics in our dividend play sample. Columns (1)–(3) and (5)–(7) report average treatment effects (ATE). SLIM and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the last cum-dividend date. Short controls include log OI, EEV, and moneyness. Extended controls include relative spread and underlying price. Robust z-statistics are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7 generalizes our case study in Table 5 to the sample of all matched contracts from our dividend play sample.[28] We again see that contracts that had experienced a larger volume of retail trading in the week preceding the dividend play date are more targeted by the arbitrageurs. The magnitudes are very similar to those in Table 6. We vary controls and numbers of neighbors across our specifications in Table 7, and the magnitudes of the coefficients do not vary much. The coefficients are also statistically indistinguishable from the OLS estimates from the same specification (columns (4) and (8)), which offers further evidence that our results are robust.

5.3 Possible explanations

We now explore possible explanations for the puzzling tendency of dividend play arbitrageurs to leave money on the table. First, there exist dividend-play specific fee caps on

[28] In fact, we can use matching on the same profitability characteristics to study arbitrageur entry. In Appendix A.12, we show that across the propensity score spectrum, there exist contracts with both zero and positive floor volume. (As shown before, 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 well, thereby exacerbating the puzzle.

PHLX and, more recently, BOX.[29] 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 usually requires two participating parties, it is highly likely that they would 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 analysis above, we always control for contract liquidity or match on contract relative spread. It is therefore unlikely that the contracts in which market makers and other arbitrageurs do not engage in dividend play are systematically less liquid.

Another potential explanation is that arbitrageurs' capital constraints bind. However, most regulatory requirements typically involve netted positions, which are relatively low given the symmetric and fully hedged nature of the strategy. It is therefore not clear why capital constraints may bind unless they bind due 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.[30] Still, such explanations cannot produce the variation in floor trader activity within and across tickers that we document: Table 14 illustrates that there are many profitable contracts in which floor traders do not participate at all.

Finally, it has been documented that even sophisticated market players exhibit limits to attention (Kacperczyk, 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. However, we proceeded to test this hypothesis more formally. To do so, 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 mattered for floor trader activity. These additional results are available upon request.

One alternative explanation that we cannot rule out is that some profits are left unexploited because of the stigma and reputational costs associated with the dividend play

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

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

strategy. The SEC has clearly signaled its disapproval of the strategy in its 2014 rule aimed at making the strategy impractical. (See footnote 11.) Reputational costs could 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.

5.4 How many arbitrageurs engage in dividend play?

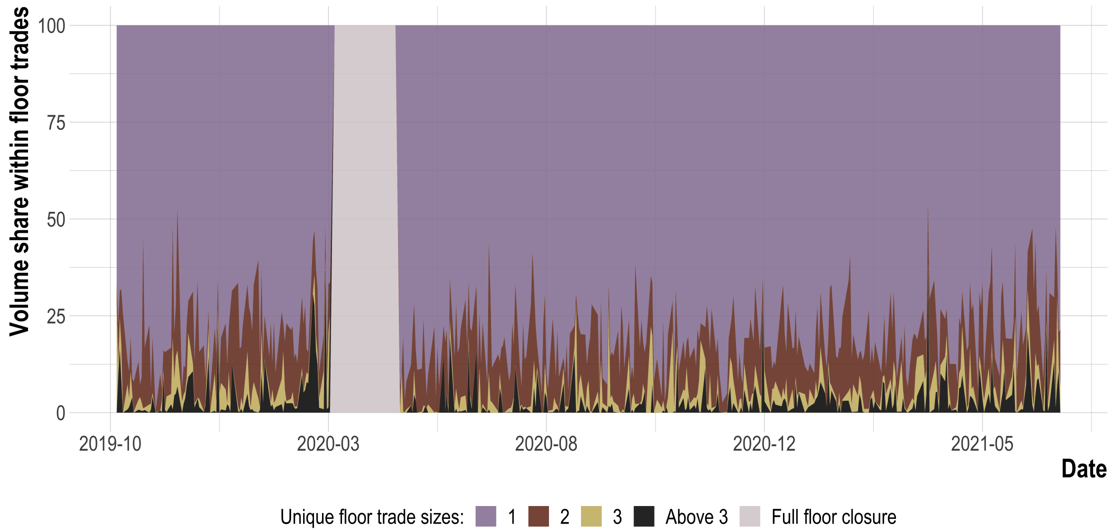
This subsection provides suggestive evidence for the number of pairs of arbitrageurs simultaneously engaging in a dividend play strategy. An advantage of our granular, transaction-level data is that we can estimate the number of arbitrageurs exploiting each contract, which prior literature has not done. As mentioned previously, dividend play trades in our data appear as a sequence of trades (normally five or more), executed within several seconds from the floor of an options exchange. There are two simultaneously executed legs of this trade by each arbitrageur: One leg is a sequence of buys of one call option contract and the other a sequence of sells of a contract with a neighboring strike. Trade sizes within each sequence of trades are typically the same, but they differ across sequences. Differences in trade sizes across dividend play trade sequences are likely due to execution preferences of individual arbitrageurs. We therefore use the number of unique trade sizes within each sequence of trades that we classify as dividend play in a given contract as a proxy for the number of arbitrageurs engaging in dividend play in that contract. Since each dividend play trade is normally executed by a pair of arbitrageurs, our measure identifies the number of such pairs.

Figure 3 plots a percentage split of dividend play trades by the number of pairs of arbitrageurs exploiting each contract. It is striking that in the majority of contracts we observe only one pair of arbitrageurs, implying that all of the arbitrage profits in a given contract accrue to a single party. Only in a small fraction of contracts is the number of arbitrageur pairs above three. Since there are no obvious impediments to free entry of other arbitrageurs who regularly engage in dividend play, it is puzzling that so few arbitrageurs participate in each contract.

Dividend play trades are easy to detect in transaction-level data, which is available to arbitrageurs in real time from standard datafeeds (e.g., OPRA or those provided by exchanges). Arbitrageurs are therefore fully aware of the presence of other arbitrageurs in a given contract before they decide to engage in a dividend play trade in it. According to textbook theories, however, this should not deter them from entering and competing for profits in the same contract.

With our data, we are unable to pinpoint the exact reason for the small number

Figure 3: Number of pairs of arbitrageurs in dividend play trades



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

of arbitrageurs participating in each contract; we can only document this empirical fact. It would be interesting to shed more light on it, as this setting might provide a useful laboratory for testing various theories of limits to arbitrage in which the number of arbitrageurs is finite. It is clearly also a possibility that our measure of the number or arbitrageurs is inaccurate and that multiple arbitrageurs use trades of identical size.

The gray shaded area in Figure 3 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 dividend play dates during the closures is the same as on any other day, which provides additional validation of the measure. Even when PHLX floor was closed but ARCA and BOX floors were open, the mean trading volume on dividend play was an order of magnitude lower.

6 Discussion and policy implications

It is apparent from our analysis that the new generation of investors, while tech-savvy and active on investing forums, is still lacking in financial education that is required to trade options. These investors are likely to make mistakes in sophisticated financial decisions such as early exercise of an option, and these mistakes generate transfers from retail investors to arbitrageurs or market makers.

It is not clear whether retail investing platforms have the appropriate incentives to prevent their customers from making trading mistakes. The question of optimal options exercise requires knowledge of option pricing models, which retail investors are likely to lack. 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 on last cum-dividend dates, when it is optimal to do so, a default option for investors, from which they can opt out if they wish. This simple enhancement would prevent retail investor losses of around $200 million in our sample.[31] More generally, the literature on investor protection (e.g., Barbu (2022), Bhattacharya, Illanes, and Padi (2019), Egan (2019), Heimer and Simsek (2019), Célérier and Vallée (2017), and Campbell, Jackson, Madrian, and Tufano (2011)) has long been concerned about complex investment products and incentives of intermediaries. The complexity of options contracts from the viewpoint of an average retail investor and the potentially misaligned incentives of intermediaries call for enhancements to investor protection on retail trading platforms.

Finally, our paper highlights the difficulties of devising effective regulation in financial markets. Two previous papers on dividend play, Hao, Kalay, and Mayhew (2009) and Pool, Stoll, and Whaley (2008), written before the 2014 SEC regulation aimed at curtailing the dividend play strategy, argue that arbitrageurs engage in dividend play by simultaneously opening long and short positions in the *same* call option contract. The 2014 regulation affected clearing rules, and it was no longer beneficial for arbitrageurs to long and short the same contract. In our transaction-level data, we clearly see that after the regulation, dividend play arbitrageurs simultaneously long and short contracts with *neighboring* strikes. For the purposes of computing regulatory risk exposures, these positions are simply netted to zero. There is some economic risk exposure resulting from this trade because the deltas of the long and short positions are not exactly the same, but the short-horizon nature of the trade and the ability of arbitrageurs to delta hedge their economic exposure, if desired, still makes it a worthwhile strategy, leading to its resurgence in the data.

The dividend play arbitrage trade does not serve any particularly useful purpose in financial markets. It neither improves liquidity nor aids price discovery. It simply reallocates profits from option writers to arbitrageurs. In the process of doing so, however, it dramatically inflates trading volume and hence skews various important statistics that market participants rely on, while exposing arbitrageurs and potentially exchanges to large

[31] While $200 million is the total potential profit to be exploited by dividend play arbitrageurs, a large fraction of it is likely to come from retail investors, as sophisticated investors are likely to exercise their options optimally.

operational risk. By documenting the mechanics of how arbitrageurs circumvent the SEC regulation, we hope to help policymakers devise a new, more effective approach to curtailing the dividend play strategy.

7 Conclusion

The new and inexperienced retail investors who entered the options market during the pandemic are more likely to suboptimally exercise their options. These mistakes can be exploited by arbitrageurs. We utilize newly introduced transaction flags to isolate one specific trading strategy executed predominantly by market makers (and other arbitrageurs): the dividend play. This strategy capitalizes on early exercise mistakes and produces (virtually) riskless arbitrage profits for market participants. We document that market makers and other arbitrageurs engage in this strategy, amplifying tenfold or more the daily trading volume in affected tickers. However, puzzlingly, they do not take part in dividend play in some contracts, forgoing large profits. We rule out a number of explanations for this puzzling behavior and document that very few pairs of arbitrageurs engage in the dividend play strategy in a given contract on the same day, with a median of just one pair. Future research may be able to shed light on why dividend play arbitrageurs leave money on the table.

The resurgence of dividend play also has implications for research that uses recent data for single-name options and ETFs. The size of the trade has become so large that it significantly skews trading volume on days when it takes place and volume shares of other trading activities. For example, trading volume shares of both small trades and SLIM trades in a particular dividend-paying stock or ETF in our sample are significantly lower on dividend trade dates relative to other days. A researcher using these measures to gauge retail presence in options may incorrectly interpret these drops as falls in retail activity, while in fact they are fully explained by the dividend play arbitrage. We therefore advise researchers, who are working on applications that are unrelated to dividend play, to exclude dividend play activity from their data. This can be done either by dropping trades executed from the floor on the last cum-dividend dates or simply by dropping the last cum-dividend dates altogether.

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, and Terrance Odean, 2013, The Behavior of Individual Investors, vol. 2 . chap. 22, pp. 1533–1570 (Elsevier).

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

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.

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.

Bryzgalova, Svetlana, Anna Pavlova, and Taisiya Sikorskaya, 2023, Retail Trading in Options and the Rise of the Big Three Wholesalers, *The Journal of Finance,* forthcoming.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, 2007, Down or out: Assessing the welfare costs of household investment mistakes, *Journal of Political Economy* 115, 707–747.

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

Célérier, Claire, and Boris Vallée, 2017, Catering to Investors Through Security Design: Headline Rate and Complexity, *The Quarterly Journal of Economics* 132, 1469–1508.

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.

de Silva, Tim, Kevin Smith, and Eric C. So, 2022, Losing is Optional: Retail Option Trading and Earnings Announcement Volatility, working paper, available at SSRN: https://ssrn.com/abstract=4050165.

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*.

Fardeau, Vincent, 2021, Arbitrage with Financial Constraints and Market Power, working paper, available at SSRN: https://ssrn.com/abstract_id=3630914.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Hao, Jia, Avner Kalay, and Stewart Mayhew, 2009, Ex-dividend Arbitrage in Option Markets, *The Review of Financial Studies* 23, 271–303.

Heimer, Rawley, and Alp Simsek, 2019, Should Retail Investors' Leverage be Limited?, *Journal of Financial Economics* 132, 1–21.

Hendershott, Terrence J., Saad Khan, and Ryan Riordan, 2022, Option Auctions, working paper, available at SSRN: https://ssrn.com/abstract=4110516.

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.

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

Pan, Jun, and Allen M. Poteshman, 2006, The Information in Option Volume for Future Stock Prices, *The Review of Financial Studies* 19, 871–908.

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.

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

Shleifer, Andrei, and Robert W. Vishny, 1997, The Limits of Arbitrage, *The Journal of Finance* 52, 35–55.

Welch, Ivo, 2022, The Wisdom of the Robinhood Crowd, *The Journal of Finance* 77, 1489–1527.

A Appendix

A.1 OPRA trade types

The following table 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 8: 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 8: 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 8: 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.2 Customer classification in the OCC data

The extant literature has relied on a Customer flag, offered by some options data providers, to identify retail trading activity. The following tables provide an additional example that in the OCC data a significnt fraction of the trading volume in the dividend play strategy is attributed to Customer transactions.

Table 9: Trading volume on a dividend play date: An example

Trading volume on last cum-div. date – 11/17/2020 (mln contracts)						
Ticker	Exchange	Customer	Firm	Total Off-Floor	Market Maker	Total
MSFT	AMEX	9,974	10,326	20,300	12,800	33,100
	BOX	9,094	7,110	16,204	44,722	60,926
	CBOE	38,285	10,827	49,112	30,318	79,430
	EMLD	8,952	126	9,078	8,600	17,678
	EDGX	6,400	280	6,680	6,910	13,590
	CFE	21,019	313	21,332	20,334	41,666
	ISE	17,590	877	18,467	18,707	37,174
	MCRY	2,056	25	2,081	2,321	4,402
	MIAX	6,285	234	6,519	7,173	13,692
	ARCA	13,072	1,460	14,532	14,524	29,056
	NSDQ	16,405	2,471	18,876	35,944	54,820
	MPRL	5,098	252	5,350	6,874	12,224
	SML	75	43	118	308	426
	C2	2,816	575	3,391	4,993	8,384
	PHLX	97,749	48,067	145,816	194,666	340,482
	BATS	13,761	3,617	17,378	27,170	44,548
Ticker Total		268,631	86,603	355,234	436,364	791,598

This table reports the trading volumes in Microsoft option contracts, as reported by the OCC on the last cum-dividend date, 11/17/2020. The trading volume (million contracts) is broken by Customer, Firm, Total Off-Floor, and Market Maker. There are large spikes in volume on AMEX and ISE relative to a normal day, reflecting the dividend play activity on that day. The reported OCC figures double-count each transaction. To reconcile the numbers with the OPRA volume, one needs to divide them by two.

Table 10: Trading volume on a normal day: An example

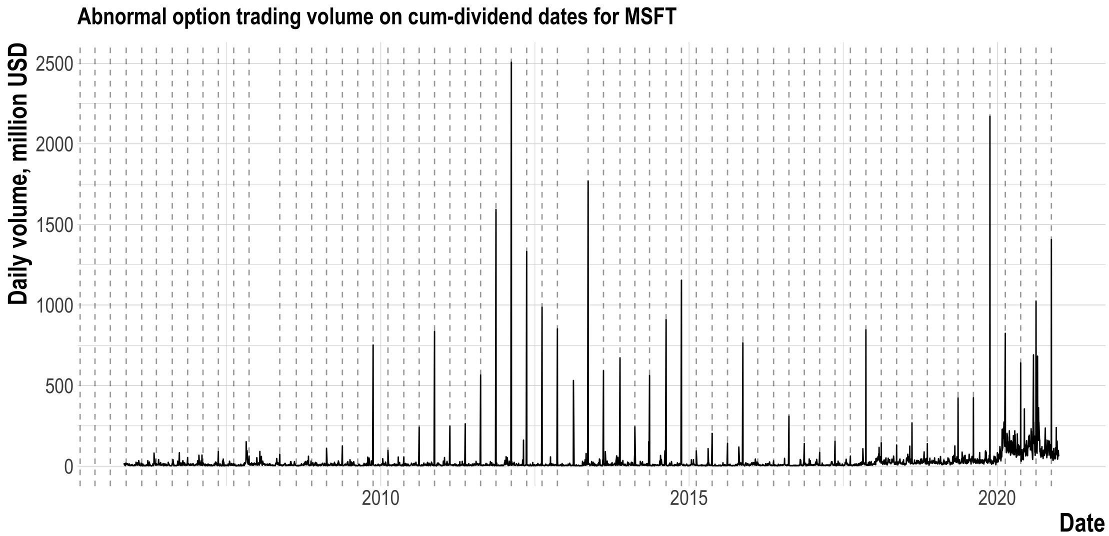
Trading volume on a normal day – 11/12/2020 (mln contracts)						
Ticker	Exchange	Customer	Firm	Total Off-Floor	Market Maker	Total
MSFT	AMEX	12,736	612	13,348	14,610	27,958
	BOX	2,505	79	2,584	4,282	6,866
	CBOE	20,208	7,839	28,047	28,473	56,520
	EMLD	13,744	27	13,771	13,159	26,930
	EDGX	7,745	1,874	9,619	6,715	16,334
	CFE	20,146	201	20,347	18,719	39,066
	ISE	16,318	3,355	19,673	19,037	38,710
	MCRY	2,342	0	2,342	2,244	4,586
	MIAX	6,502	339	6,841	6,859	13,700
	ARCA	22,946	6,262	29,208	15,832	45,040
	NSDQ	20,571	1,343	21,914	46,410	68,324
	MPRL	5,118	291	5,409	6,557	11,966
	SML	338	1	339	371	710
	C2	3,440	764	4,204	5,708	9,912
	PHLX	29,200	5,892	35,092	23,792	58,884
	BATS	16,936	3,296	20,232	30,876	51,108
Ticker Total		200,795	32,175	232,970	243,644	476,614

This table reports the trading volumes in Microsoft option contracts, as reported by the OCC on a normal day, 11/12/2020. The trading volume (million contracts) is broken by Customer, Firm, Total Off-Floor, and Market Maker. There are large spikes in volume on AMEX and ISE relative to a normal day, reflecting the dividend play activity on that day. The reported OCC figures double-count each transaction. To reconcile the numbers with the OPRA volume, one needs to divide them by two.

A.3 Abnormal trading volume on cum-dividend dates: Further examples

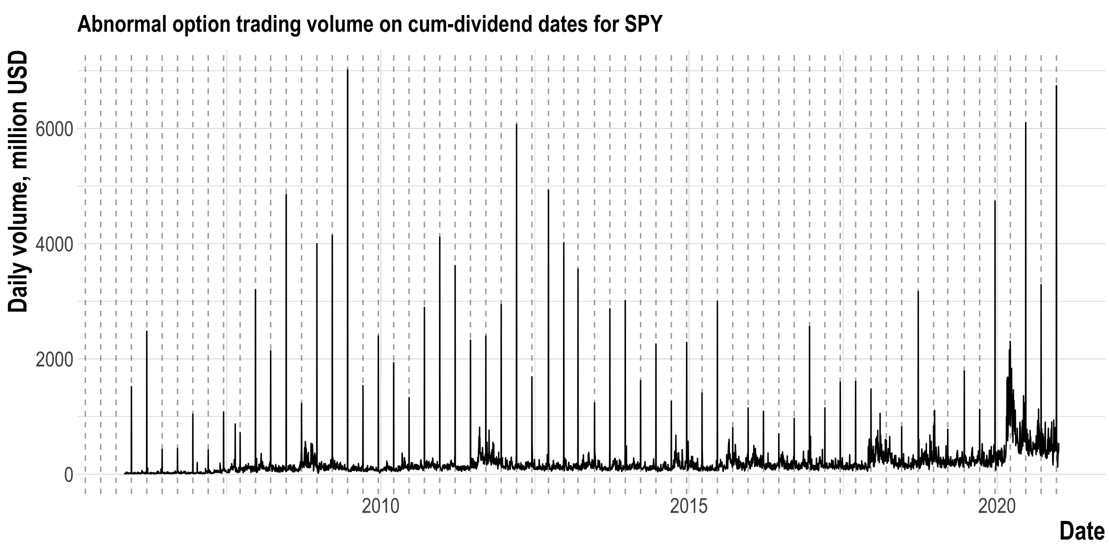
This appendix contains two further examples of abnormal trading volume on cum-dividend dates. The figures below plot daily trading volume of options on Microsoft, MSFT, and on the largest S&P 500 ETF, SPY.

Figure 4: Abnormal trading volume on last cum-dividend dates for MSFT



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 the last cum-dividend dates.

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



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 the last cum-dividend dates.

A.4 Dividend play: Another example

Table 11 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 and 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 market makers. The market makers 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 market makers' short positions remain unassigned and hence yield a gain. In our example, market makers 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, market makers simply increase the number of contracts they buy and sell.

Table 11: Dividend play: Another Example

	OI_{t-1}	New positions(t)	Available for ex.	No. exercised	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		
Market makers	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 last 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.5 Dividend play: Technical details

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

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

where S_{ex} is the expected price after the stock goes ex-dividend, that is, price at close on the last 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[32]; r is the interpolated maturity-specific interest rate provided by OptionMetrics (annualized %); and $Dividend_{ex}$ is the expected dividend after the ex-date.[33]

A.6 Dividend play sample: Data filters and calculated variables

We use our dataset described in Section 2.1 together with the following filters to arrive at the final dividend play sample. We include all call option contracts with $EEV > 0$. Furthermore, since our valuation might be imperfect, we add a market-based filter of the optimality of exercise: We keep only contracts with a decline in open interest on the last cum-dividend date.[34] By implication, we have only contracts with non-zero open interest on the last cum-dividend date and on the day before that.

Following the early papers on dividend play, we remove contracts with no trading volume on the last cum-dividend date. Additionally, we remove contracts expiring immediately after the ex-dividend date.[35]

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 last cum-dividend date.[36] For both SLIM and Small Share, we compute a one-week moving

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

[33] 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.

[34] This is consitent with Hao, Kalay, and Mayhew (2009).

[35] The last filter does not change results significantly.

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

average and use its lagged value on the cum-dividend 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 the trade as $0.5(underlying\,bid + underlying\,ask)/strike - 1$.[37]

Table 12: Dividend play sample descriptive statistics

	Mean	Median	St. Dev.	p1	p99
Fraction of OI unexercised, %	23.38	6.98	31.02	0.00	99.32
Floor trades volume share on last cum-date	0.56	0.85	0.47	0.00	1.00
D(floor share > 0)	0.61	1.00	0.49	0.00	1.00
SLIM Share	0.15	0.15	0.16	0.00	0.65
Small Share	0.85	0.88	0.18	0.27	1.00
Internalized volume in underlying (share)	0.17	0.16	0.05	0.07	0.30
WSB mentions	8.20	0.33	36.15	0.00	232.67
OI, log	4.97	4.89	1.99	0.69	9.80
Early exercise value (EEV), $	0.49	0.30	0.59	0.00	2.79
Market EEV, $	0.08	0.02	0.39	-0.41	1.12
Dollar potential profit	7,059.9	215.4	57,949.3	0.00	121,549.1
Dollar volume, log	4.16	4.09	1.40	1.17	7.85
Relative spread	0.07	0.04	0.09	0.00	0.46
Implied volatility, annualized	0.41	0.36	0.28	0.07	1.29
Moneyness	5.56	3.99	5.57	0.47	29.27
Days to expiry	67	21	124	4	625

This table reports descriptive statistics for all contracts in the dividend play sample (21,997 observations). SLIM and small share are the contract-level volume shares of SLIM and small trades, respectively, averaged over one trading week before the last cum-dividend date. The share of internalized volume in the total trading volume in the underlying is computed as a ticker-level non-ATS OTC share of volume executed in the non-ATS OTC space relative to the total trading volume, averaged over one trading week before the last cum-dividend date. WSB mentions is the number of underlying ticker mentions on the `WallStreetBets` forum, averaged over one trading week before the last cum-dividend date. Relative spread is the option 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 is measured as $(Midpoint\,Price - Strike)/Strike$.

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

A.7 OTC trading volume, by venue

Table 13: Top 15 venues in the non-ATS OTC market in the U.S.

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 venues in terms of their total OTC non-ATS volume in 11/2019–06/2021. Based on FINRA OTC Transparency data.

A.8 Dividend play profits by ticker

Table 14: Dividend play profits by ticker

Ticker	**Profit, USD** Harvested	**Profit, USD** Forgone	**No. contracts** Fully harvested	**No. contracts** Partly harvested	**No. contracts** Forgone	**Traded volume (contracts)**
Ticker	(1)	(2)	(3)	(4)	(5)	(6)
SPY	2,395,306.0	23,600,000.0	228	38	991	722,404
AAPL	3,464,514.0	8,591,433.0	302	129	238	849,464
EEM	10,800,000.0	4,907,643.0	142	5	46	5,266,442
IWM	1,484,462.0	3,528,205.0	51	3	162	512,940
EFA	2,436,712.0	3,324,776.0	97	7	25	1,087,908
XLE	1,794,247.0	2,844,799.0	171	7	91	447,875
VALE	1,913,317.0	2,737,224.0	66	6	11	1,876,400
QQQ	105,585.1	2,039,488.0	23	2	208	27,750
EWZ	3,374,398.0	1,330,285.0	87	1	31	1,327,432
KO	440,711.1	1,097,762.0	76	23	66	322,120
HYG	36,411.0	912,840.1	11	4	57	63,710
SAN	-	753,484.1	0	0	11	-
HD	936,386.6	674,275.4	95	21	96	197,887
COST	1,703.4	658,710.4	11	4	43	1,207
XLF	478,434.8	620,190.3	57	5	57	344,130
IBM	684,029.4	567,198.3	116	116	36	383,594
BHP	176,163.5	553,367.0	33	4	12	57,055
DIA	156,767.9	539,148.9	60	8	146	17,401
ET	620,329.4	529,149.1	51	12	48	574,990
QCOM	749,520.4	497,014.1	68	16	49	426,659
GOLD	313,229.0	449,155.9	27	4	32	68,580
VIAC	1,770,252.0	420,720.0	97	1	51	437,395
XOM	8,123,625.0	404,302.2	242	82	62	1,734,910
XLI	10,843.2	401,816.1	12	1	15	17,123
RIO	27,621.1	375,591.7	16	4	5	57,782
XLP	116,856.2	370,016.7	16	0	32	15,990
T	2,700,908.0	369,322.1	155	35	47	2,381,173
JPM	883,655.1	365,889.3	80	32	37	1,096,394
CVX	623,103.9	320,866.8	234	86	91	419,757
FXI	877,237.6	309,056.5	77	4	18	1,242,431
GILD	355,846.2	308,605.1	65	23	43	280,310
MRO	-	307,556.9	0	0	23	-
NVDA	-	283,368.6	0	0	57	-
BP	339,841.9	277,599.5	99	28	52	209,456
DIS	836,205.3	273,563.9	41	5	1	503,899
PGR	661,473.3	263,863.8	12	5	20	56,496
MPC	561,419.5	251,516.0	105	47	64	418,852
TGT	90,064.5	241,550.0	77	29	66	65,810
DOW	138,056.3	231,155.0	32	13	61	92,369
PRU	68,932.0	224,236.2	33	29	18	85,664
Total	**50,548,168.7**	**66,756,745.0**	**3,165**	**839**	**3,219**	**23,691,759**

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 in 11/2019–06/2021. Total dividend play potential profits are computed as in equation (1). To compute Harvested profits, we multiply the total profits by the floor volume share on the last 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 over 99% in Fully harvested contracts and 69% in Partly harvested contracts.

A.9 Retail trading measured over a longer window

In this appendix, we redefine our retail trading measures. Rather than measuring shares of retail trading in the dollar trading volume in options over one trading week preceding the last cum-dividend date, we measure them over two trading weeks. Tables 15 and 16 are the analogs of Tables 4 and 6, respectiviely, but with the redefined measures of retail trading.

Table 15: Suboptimal exercise and retail investor popularity

	Dividend play profitability feature					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Fraction of OI unexercised, %						
SLIM Share	7.460***		7.280***	7.689**	7.633***	7.628***
	(3.46)		(3.36)	(2.34)	(3.57)	(3.41)
Small Share		5.120**	4.767**	12.893***	5.818***	5.041**
		(2.23)	(2.06)	(3.94)	(2.62)	(2.00)
Internalized volume in underlying						25.559
						(1.19)
WSB mentions, log						0.557*
						(1.69)
Observations	21,105	21,105	21,105	6,942	21,105	19,134
Adjusted R-squared	0.238	0.237	0.238	0.345	0.214	0.214
Panel B. Potential profits, log U.S. dollar						
SLIM Share	1.970***		1.954***	2.238***	2.078***	2.040***
	(7.56)		(7.52)	(5.46)	(8.10)	(7.59)
Small Share		0.568	0.485	2.504***	0.520	0.531
		(1.61)	(1.40)	(3.49)	(1.62)	(1.47)
Internalized volume in underlying						1.388
						(0.69)
WSB mentions, log						0.036
						(0.87)
Observations	21,105	21,105	21,105	6,942	21,105	19,134
Adjusted R-squared	0.315	0.311	0.315	0.318	0.293	0.299
Sample	All	All	All	Top EEV tercile	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 (2) in our dividend play sample. SLIM Share and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over two trading weeks before the last 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 two trading weeks before the cum-dividend date. WSB mentions, log, is the logarithm of total mentions of the ticker on the `WallStreetBets` forum. In Panel B, contract controls include log dollar trading volume, relative spread, IV, moneyness, and 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. Standard errors are clustered by ticker and date. Robust t-statistics are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 16: Arbitrageur activity and retail investor popularity: Retail trading measured over a longer window

	Floor trading share on last cum-date				**D(floor share > 0)**	**Floor trading volume, log**
	(1)	(2)	(3)	(4)	(5)	(6)
SLIM Share	0.051*		0.038	0.057	0.033	0.604**
	(1.82)		(1.45)	(1.21)	(1.22)	(2.57)
Small Share		0.338***	0.336***	0.126	0.351***	4.848***
		(6.05)	(6.07)	(1.22)	(6.24)	(8.26)
Observations	21,105	21,105	21,105	6,942	21,105	21,105
Adjusted R-squared	0.404	0.410	0.410	0.473	0.399	0.479
Sample	All	All	All	Top EEV tercile	All	All
Contract controls	Y	Y	Y	Y	Y	Y

This table reports estimates of (3) in our dividend play sample. SLIM Share and Small Share are the contract-level volume shares of SLIM and small trades, respectively, averaged over two trading weeks before the last cum-dividend date. Contract controls include log OI, EEV, log dollar trading volume, relative spread, IV, moneyness, and days to expiration. All regressions include ticker by date fixed effects. Robust t-statistics are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.10 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 the last cum-dividend dates undertaken by retail brokerages. The example is from the Robinhood Terms and Conditions.

Figure 6: 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.11 Covariate balance for matching

Figure 7: Covariate balance for SLIM share in Table 7

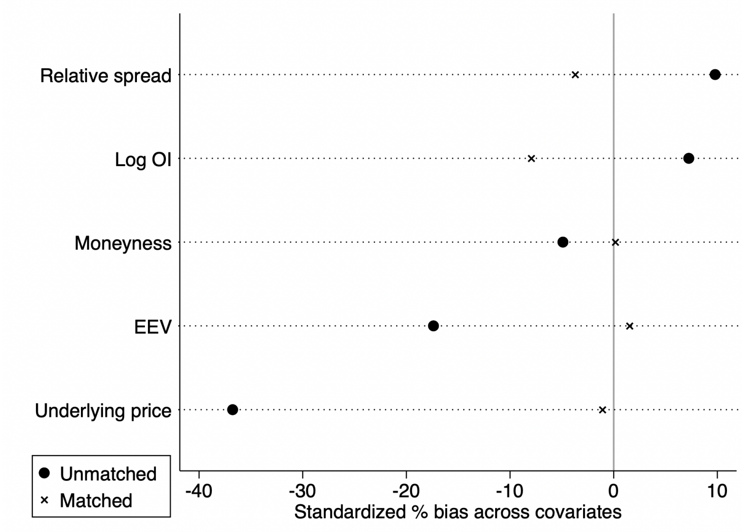
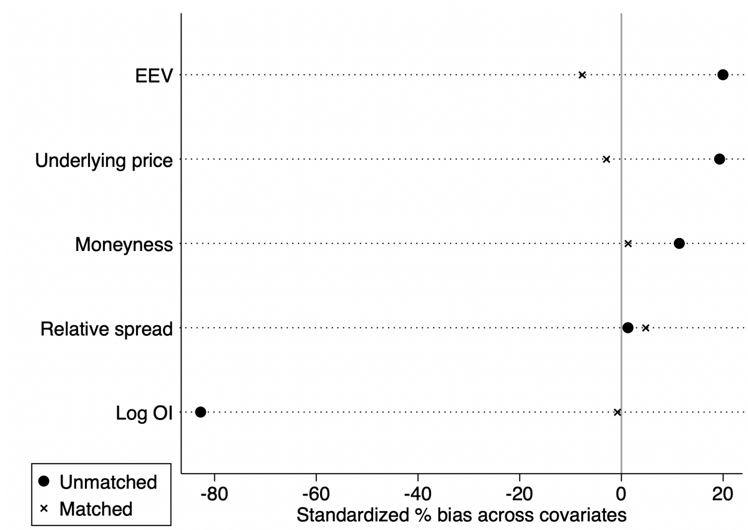
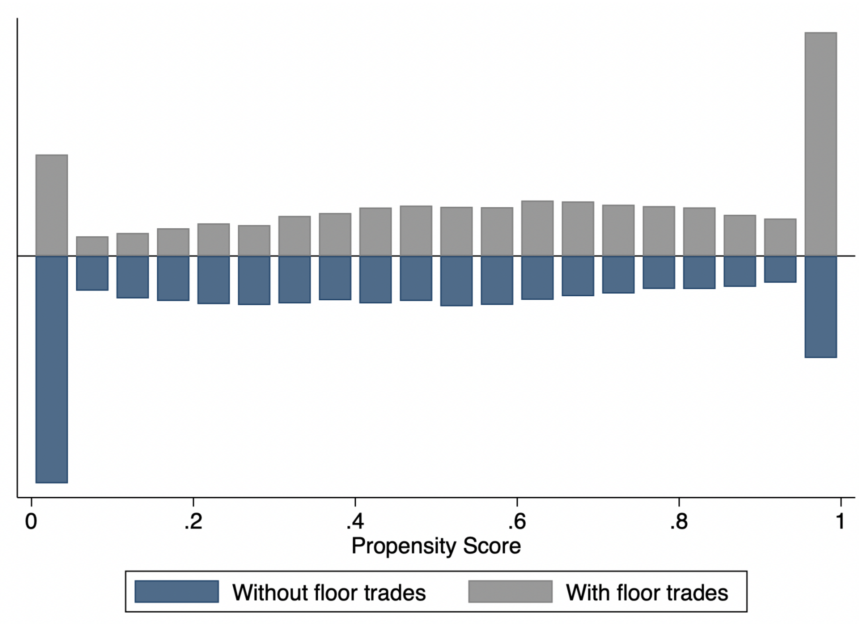


Figure 8: Covariate balance for Small share in Table 7



A.12 Dividend play puzzle in matched contracts

Figure 9: 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 full set of controls: log OI, EEV, log trading volume, relative spread, IV, moneyness, days to expiry, underlying price, underlying volatility, underlying relative bid-ask spread, and underlying market cap. We report the balance tests in Appendix A.11.

Figure 10: Covariate balance for Floor share in Figure 9