When Do Informed Short Sellers Trade?

Evidence from Intraday Data and Implications for Informed Trading Models

Danqi Hu, Charles M. Jones and Xiaoyan Zhang*

June 17, 2021

ABSTRACT

Using intraday 2015-2019 short sale data from CBOE and FINRA, we examine the intraday time patterns and information content of on-exchange and off-exchange shorting. Midday short sales and those near the open strongly and negatively predict the cross-section of stock returns at daily horizons and up to 12 weeks ahead. Short sales near the close are only informative at next-day horizons. We also connect earnings/analyst news to shorting flows. Shorting near the open reacts strongly to and anticipates bad-news releases. Shorting near the close only anticipates post-close firm news. Our evidence supports both Kyle (1985) and Holden and Subrahmanyam (1992).

Keywords: short selling, intraday trading, information content, off-exchange trading, dynamic informed trading models, institutional investors, retail investors.
JEL codes: G11, G12, G14, G23

* Hu is at the Kellogg School of Management at Northwestern University, Jones is at Columbia Business School, and Zhang is at the PBC School of Finance at Tsinghua University. Jones thanks the Norwegian Finance Initiative at Norges Bank for helpful funding. Zhang acknowledges financial support from China NSF Project 71790605. We thank Jonathan Sokobin at FINRA for helpful comments and for pointing us to their time-stamped off-exchange (Trade Reporting Facility) short sale data. We also received helpful comments from workshop participants at Columbia Business School, Microstructure Online Seminars Asia Pacific, Shanghai Jiaotong University, and Tsinghua University.

When Do Informed Short Sellers Trade?

Evidence from Intraday Data and Implications for Informed Trading Models

June 17, 2021

ABSTRACT

Using intraday 2015-2019 short sale data from CBOE and FINRA, we examine the intraday time patterns and information content of on-exchange and off-exchange shorting. Midday short sales and those near the open strongly and negatively predict the cross-section of stock returns at daily horizons and up to 12 weeks ahead. Short sales near the close are only informative at next-day horizons. We also connect earnings/analyst news to shorting flows. Shorting near the open reacts strongly to and anticipates bad-news releases. Shorting near the close only anticipates post-close firm news. Our evidence supports both Kyle (1985) and Holden and Subrahmanyam (1992).

Keywords: short selling, intraday trading, information content, off-exchange trading, dynamic informed trading models, institutional investors, retail investors.
JEL codes: G11, G12, G14, G23

1. Introduction

There are many reasons to sell equity shares short. Market-makers may sell short as part of providing liquidity to share purchasers. Short sellers may be hedging out the risk of a derivatives position; they may have negative information about a particular company that has not yet been incorporated into prices; they may simply believe that a company is overvalued relative to its fundamentals; or they may be pessimistic about a company's prospects. It is also possible that at least some short sellers are predatory or abusive, either intending to drive prices below fundamental value by shorting aggressively, and/or hoping to force the company into a decision that reduces its value (Goldstein and Guembel 2008). Most of the evidence suggests that short sellers help make stock prices more efficient by incorporating negative information into share prices.[1] Almost all the previous literature shows that heavier shorting leads to lower share prices and is associated with worse firm fundamentals in the future,[2] and there is little to no evidence of price reversals following these price declines.

In this paper, we look closely at the behavior of U.S. short sellers throughout the course of the trading day, and we examine the time patterns and information content of on-exchange and off-exchange short sales. We have time-stamped data on all reported off-exchange short sales, and similar on-exchange data from the four CBOE exchanges. The time-stamped intraday data allows

[1] Many empirical studies support this view (e.g, Lamont and Thaler 2003, Mitchell et al. 2002, Bris et al. 2007, and Boehmer and Wu 2013). Meanwhile, journalists and investors noted that short sellers uncovered the Enron fraud and other similar events. Regulators eased various restrictions on shorting, including the so-called uptick rule in 2007, and shorting activity became quite widespread, accounting for as much as 40% of trading volume by the end of 2007 (Diether et al. 2009a).

[2] For example, Asquith et al. (2005) and Boehmer et al. (2008) show that shorts negatively predict future stock returns. Christophe et al. (2004) find that negative earnings surprises are preceded by abnormal short selling, and Christophe et al. (2010) find that shorting predicts future analyst downgrades. Francis et al. (2005) show that short sellers are able to predict downward analyst forecast revisions, Desai et al. (2006) find that short sellers anticipate earnings restatements, and Karpoff and Lou (2010) show that short sellers have advanced knowledge of corporate financial misconduct. Boehmer, Jones, Wu, and Zhang (2020) find that a substantial fraction of the underperformance by heavily shorted firms is due to earnings shortfalls, analyst recommendation downgrades, and downward analyst forecast revisions.

us to investigate whether and how informed short sales concentrate at certain times of day, and it allows us to gauge whether short sellers trade on negative private information differently if it is short-lived vs. long-lived.

When they trade, investors make choices on trading venues: on-exchange or off-exchange. The most important exchanges in the U.S. include the NYSE, Nasdaq and the four exchanges owned by the CBOE. The off-exchange venues include dark pools, crossing networks, and wholesalers. Comparing on-exchange and off-exchange venues, it is generally believed that exchanges provide transparency, while off-exchange venues provide more privacy. The opaque off-exchange venues may be of particular value to short sellers who are trying to take positions gradually and quietly over time, while the lit exchange venues might deter private trades but encourage quick information incorporation. By studying short sales from all off-exchange short sales at FINRA, and all exchange short sales at CBOE exchanges, we provide insights on how different venues are related to short-sellers' intraday trading behavior and their information content. In addition, the off-exchange venues allow sub-penny transaction prices, which allows us to separate short-sales into subgroups related to traders' identity. A novel algorithm, Boehmer, Jones, Zhang, and Zhang (2021) (BJZZ (2011)), allows us to identify off-exchange trades that are almost surely marketable orders submitted by a retail trader and executed by an internalizing or wholesaling broker. Therefore, we are able to compare these short sale executions to the complement of off-exchange executions, which are much more likely to consist of short sales that involve an institution.

Using our time-stamped short sale data, we find distinct differences between short sales that occur shortly after the 09:30:00 open vs. short sales that occur in the middle of the trading day vs. short sales near the 4pm close. We separate intraday shorting into three intervals: near the open (before 10am), midday (between 10am and 3:30pm) and near the close (after 3:30pm). The most important time-of-day finding is that short sales near the open and during the middle of the day, whether on-exchange or off-exchange, anticipate the future cross-section of returns at both short horizons of one day and up to 12 weeks ahead. But the predictive power is much stronger over short horizons. In contrast, short sales between 15:30:00 and 15:59:59 seem to possess useful negative information only about short-term future returns, and the predictive power becomes insignificant over longer horizons.

We also document interesting patterns when we compare on-exchange and off-exchange samples. The Boehmer, Jones, Zhang and Zhang (2021) retail trade identification algorithm conveniently separates off-exchange short sales involving “retail sells” from short sales involving institutions. Notice that while the algorithm can identify the trade direction for retail trades, it cannot identify trade direction as precisely for short sales involving institutions. Similar to the findings on CBOE, off-exchange shorting flow near the open and during the middle of the day has the strongest predictive power for the future cross-section of returns. Interestingly, for off-exchange short sales near the close, those involving institutions barely predict either short- or long-term returns, while those involving retail sells have strong predictions for both short- and long-term returns. This difference suggests that that the purpose of institution-involving short sales near the close might be liquidity or rebalancing needs rather than negative information.

We further examine intraday predictive patterns and confirm that the predictive power for short-term predictions gradually diminishes when we move through the day in 30-minute intervals. For the CBOE short sales, the predictive power of short sales stays significant until the last 30 minutes. Off-exchange shorting flow involving institutions loses its same-day predictive significance 1.5 hours before market close.

Finally, to understand whether the short-sellers' information advantage comes from their information processing skills or access to private information, we connect earnings news releases to shorting flows. In recent years, company news releases mostly happen after trading hours. CBOE shorts quickly react to previously-announced bad news, mostly within the first 30 minutes. In contrast, FINRA institutional shorts and retail shorts both react to public bad news, but with different speeds. The reaction time of FINRA institutional shorts resembles that of CBOE shorts, mostly within the first 30 minutes, while the FINRA shorts involving retail sells take at least one day to react to bad news. Regarding the ability to predict the arrival of future bad news releases, FINRA shorting seems to perform better than COBE shorting, indicating that, compared to FINRA shorting, the predictive power of on-exchange short sales comes more from their quick processing of public news than from their access to private news. Among the three intraday buckets, opening shorts across trading venues and trader identities react strongly to and oftentimes anticipate bad-news releases, whereas shorting near the close is only related to earnings news occasionally released by firms post-close. This gives a strong indication of at least one of the types of fundamental information that informed short sellers may aggressively trade on near the open.

We interpret our results using different dynamic models of informed trading, including Kyle (1985) and Holden and Subrahmanyam (1992), among others. If their private information is obtained overnight and is likely to quickly become stale or public, short sellers tend to take their positions fairly rapidly near the open. If their information is likely to remain private for longer, short sellers trade more gradually throughout the trading day, and this private information is incorporated into price in the following days or weeks, consistent with the dynamic version of Kyle (1985).

Our results on the short-term predictions of the three intraday measures – opening shorts (especially those happening on exchanges that offer immediacy) have the strongest predictive power for both the rest-of-day returns and next-day returns – are consistent with the aggressive trading hypothesis from the Holden and Subrahmanyam (1992) model. Our finding that opening shorting reacts strongly to past news events and anticipates future news events is also supportive for this model. Our results for longer horizons suggest gradual trading over time by short sellers who have considerable monopoly power over their private information along the lines of the dynamic version of Kyle (1985), and this information is only gradually incorporated into prices over the following days and weeks.

There are very few existing papers that examine intraday shorting data. One such paper is Comerton-Forde, Jones, and Putnins (2016), who look at effective spreads and price impacts in the minutes before and immediately after short sales. They mainly distinguish short sales that supply liquidity vs. short sales that demand liquidity, and they focus on pre-short and post-short horizons from one minute to one day. Other relevant papers using intraday shorting data are Boehmer, Jones, and Zhang (2013), who examine the SEC decision during the peak of the financial crisis in September 2008 to place a temporary ban on most shorting in about 800 financial stocks, a trade-by-trade look at 2005 short sales on the INET alternative trading system (ATS) by Chakraborty, Moulton, and Shkilko (2012), and Jain, Jain, and McInish (2012), who study the 2010 introduction of an "alternative uptick rule" that restricts aggressive short sales for up to two days after a stock drops by at least 10% during the trading day. Compared to the existing literature, we believe this is the first paper to distinguish empirically the information content of short sales during various parts of the trading day, to test distinct (and often competing) hypotheses of various informed trading theory models using our intraday short sale data, and to find that short sales at different times of the trading day are informative at different horizons. The findings in this article provide unique insights for investors and regulators to better understand the nature of price discovery due to short-selling. Our results might help investors make better trading decisions, and they might help regulators oversee the market structure that undergirds short selling in the U.S.

The rest of the paper is structured as follows. Section 2 discusses the existing dynamic informed trading models and develops hypothesis for empirical study. We introduce the intraday shorting data as well as our algorithm for partitioning off-exchange short sales into those involving a retail trader vs. those more likely to involve an institution in Section 3. The main empirical results are provided in Section 4. Section 5 concludes.

2. Models of informed trading and hypothesis development

There are many different dynamic theoretical models that aim to capture informed trading over time under various assumptions. Within a setup of insider trading with sequential auctions, Kyle (1985) proposes a monopolist informed trading model. This dynamic model with sequential equilibria implies steady trading by the informed trader over time until the private information is released at a known time T.

In contrast, many other models predict that informed traders will trade more aggressively to start and less aggressively later on. For example, in Holden and Subrahmanyam (1992), the timing, nature, and duration of private information are the same as in Kyle (1985), but competing informed traders receive the same private signal of value, and as a result they trade aggressively in the first few rounds of trading, leaving little remaining information for later rounds before the information is publicly released at a known time T. Similarly, in one of the formulations studied in Foster and Viswanathan (1990), an informative public signal is released at intermediate times while the market is closed, and this leads informed traders to trade more aggressively early on, carrying less private information forward to future trading sessions. Bernhardt and Miao (2004) examine the related case where private information gradually becomes stale, either because it is publicly released over time or because other traders get correlated signals. They generally find that this leads to more rapid trading on private information initially, with more subdued informed trading later on. Finally, Slezak (1994) introduces a noisy rational expectations model with periodic exchange closures and examines informed trading before and after the closure. He finds that informed trading is always lower pre-closure and greater after the market reopens, mainly because the informed traders receive more information during the closure, and risk-averse uninformed traders who provide liquidity are less willing to do so before the closure.

These models would apply to intraday informed trading if investors typically obtain information while the market is closed and if information is typically released publicly while the market is closed. This is certainly the case for earnings announcements, which are now usually released after the market closes for the day (Lyle, Rigsby, Stephan, and Yohn 2019), and if information is gradually produced, processed or acquired in calendar time, much of this process will also take place while the market is closed.[3] If this describes the timing of information acquisition and public release, these models would predict either more information content in trading near the open or a steady amount of information-based trading over the course of the trading day (Kyle, 1985).

There is also a scenario where short sales late in the day are the most informative about future returns. Suppose for instance that most private information concerns earnings or analyst-

[3] Prior literature has documented a gradual shift in earnings announcement timing from regular trading hours to outside of regular trading hours. Patell and Wolfson (1984) find 67% of their sample in 1976/1977 announce during regular hours, while Lyle, Rigsby, Stephan, and Yohn (2018) document more than 95% of firms announce outside of regular trading hours from 2006 to 2015.

related fundamental information (as found by Boehmer et al. (2020)) is released after the close or the next morning before the open, and suppose that investors obtain or process this information late in the trading day, just before the markets close, with the fundamental information to be released after the close or before markets reopen for trading. Then virtually any model of information-based trading will imply that informed trading near that particular close will be much more predictive about the next day's returns relative to trading at any other time. Thus, we have a variety of empirical hypotheses that may potentially support different dynamic models of informed trading.

3. Data

We start by introducing our intraday time-stamped sample of reported on-exchange and off-exchange transactions in the U.S. that involve a short seller in Section 3.1. In Section 3.2, we describe our algorithm that separates off-exchange short sales into two disjoint groups: one group consists almost exclusively of transactions involving a retail investor interacting with a wholesale market-maker, and the other group consists of short sales that are much more likely to involve an institution.

3.1 Data on short selling

We focus on a recent sample period from January 2, 2015 to December 31, 2019, and we obtain intraday data from the four exchanges of CBOE, and off-exchange data from FINRA. The two largest exchanges in the U.S. are the Nasdaq and the NYSE Group, which charge fees for their proprietary trade-by-trade short sale information. Therefore, we collect publicly available time-stamped shorting data from the 3rd largest exchange group, CBOE, which releases shorting data every night on its website, https://markets.cboe.com/us/equities/market_statistics/short_sale, for all four of its exchanges (BYX, BZX, EDGA, and EDGX). This time-stamped short sale dataset is similar to the consolidated tape of all U.S. equity transactions in that it includes the ticker symbol, trade price, size, and other sale conditions, along with a time stamp to the nearest second. If we take all on-exchange and off-exchange trading volumes, short sales on CBOE exchanges account for about 9.37% of total trading volume, and short sales on FINRA account for about 13.6% of total trading volume.

The FINRA short sale sample consists of all short sale transactions that are executed off-exchange and reported to the consolidated tape. All of the transactions are reported via a FINRA Trade Reporting Facility (TRF), and the time-stamped details of the transaction (mainly price and size) are available to market participants in real time via a Securities Information Processor (SIP). By the start of the next trading day, the transaction details are also included in the daily TAQ database used by many researchers. But users of the real-time SIP transaction feed or the daily TAQ database cannot observe whether the seller in a reported off-exchange transaction is a short seller. By the start of the next trading day, FINRA posts on its website a summary for each ticker symbol of the total reported off-exchange trading volume that day and the number of those reported shares that were sold by a short seller. There is no transaction-by-transaction short sale information included in these daily summaries. Two weeks or so after the end of each month, FINRA posts all off-exchange transactions for that month that involve a short seller, and this trade-by-trade short sale dataset, which is similar to the CBOE dataset, is the one that we use.

There are lags between real transactions and data posting. For CBOE, the data become available on the same night, while FINRA does not post the trade-by-trade time-stamped short sale dataset until the middle of the following month. Due to the publication lags, market participants would not have the time stamps or the individual short sale trade prices and sizes in their information sets right away. Therefore, the cross-sectional predictability that we find from time-stamped trade-by-trade short sales, at intraday horizons for CBOE and short horizons for FINRA, should be interpreted as a measure of short sale information content rather than as a signal that can be used in real time to implement a related trading strategy.

The relative opacity of off-exchange short sales, and the long lag before these time-stamped short sales are publicly posted, indicates that FINRA might be preferred by traders who value privacy. It is well known that informed institutional traders who are trying to execute relatively large parent orders tend to disproportionately use dark pools and crossing networks (at least at first) as venues for child orders in order to pay less for immediacy and also perhaps to reveal as little as possible about their underlying trading intentions (see for example Menkveld, Yueshen and Zhu, 2017). Even though off-exchange trades are reported immediately to the tape, it is very difficult, for example, to reliably sign a trade as buyer- vs. seller-initiated if it takes place as a midpoint cross. An informed short seller may very well have the same desire for opacity, and if the details of her off-exchange short sales will not be revealed until well into the following month, that short seller may attach considerable value to the resulting lack of transparency.

In comparison, there are reasons to think that on-exchange shorting could also be quite informed. For example, Reed, Samadi and Sokobin (2020) argue that informed shorts with short-lived information trade urgently on exchanges to ensure execution, and they find confirming information content evidence in their non-public disaggregated TRF data. (See also Menkveld, Yueshen and Zhu (2017) for a similar ranking within the class of dark pools.)

It is important to note that even with our data, we can observe within about a month and a half every off-exchange short sale in an NMS security (generally securities listed on a major US exchange), but we do not ever observe whether the buyer is covering an existing short position.

Thus, it is impossible for us to reconcile our short sale data with twice-per-month short interest data. We view our data as a flow measure of shorting, and most of our empirical tests analyze whether certain types of short sales, or short sales at certain times of the day, better anticipate future share price declines over specific time horizons.

With the on-exchange and off-exchange trade-by-trade short sale data in hand, we then cross-match to CRSP using CUSIPs and ticker symbols. We retain only common stocks (those with a CRSP share code equal to 10 or 11) and exclude securities such as warrants, preferred shares, American Depositary Receipts, closed-end funds, and REITs. We require a minimum share price greater than $1 per share. From October 2016 through September 2018, the SEC mandated a tick size pilot that changed the minimum tick size for a large number of less active smaller-cap stocks. To minimize the impact of the pilot on our results, we exclude pilot program treatment stocks during the pilot period.[4] Combined with our other filters, this means that none of our sample stocks experience a change in their minimum price increments during the January 2015 through December 2019 sample period.[5] The resulting FINRA sample consists of a daily average of about 2,700 common stocks listed on one of the major exchanges before and after the tick size pilot, and about 1,900 sample stocks during the tick size pilot. The CBOE sample has about 3,000 and 2,200 sample stocks outside and during the tick size pilot period, respectively.

Following many other researchers who focus on shorting flow (going back at least to Boehmer, Jones and Zhang, 2008), we compute daily proportional shorting flow over total volume for stock i on day t as,

[4] In other words, during the tick-size pilot program, we require either that a stock does not meet the criteria and is thus not part of the tick size pilot, or the stock is eligible for the pilot but is randomized to be a control stock during the two-year pilot period.

[5] Stocks with market caps greater than $3 billion are not eligible for the tick size pilot, so during the pilot period our sample is somewhat underweighted in the smaller-cap eligible stocks. However, the regression is mainly identified by cross-sectional variation in the amount of off-exchange shorting, and unreported summary statistics show that off-exchange shorting has very similar cross-sectional properties during the pilot period vs. outside of the pilot period.

$$SS_{it} = \frac{Shares\ Shorted_{it}}{Total\ Share\ Volume_{it}}\ . \tag{1}$$

For CBOE shorting flows, the numerators are the total shares shorted at CBOE exchanges, and the denominators are total CRSP share volumes, which aggregate volume at all venues. For FINRA shorting flows, the numerators are the total shares sold short in FINRA's short-sale transaction files (excluding short-exempt trades that are typically hedging trades or relative-value trades between, say, a stock and an ADR), and the denominator is that stock-day's total off-exchange trading volume reported by FINRA (Wang, Yan and Zheng, 2020).

We compute intraday shorting flow variables in a parallel fashion. First, we divide the 09:30 to 16:00 trading day into 13 half-hour intervals (09:30:00-09:59:59, 10:00:00-10:29:59, and so on up to 15:30:00-15:59:59). Next, we measure shorting flow during a given half hour as the fraction of volume in a given stock on a given day that involves a short seller as follows:

$$SS_{it}^{begintime} = \frac{Shares\ Shorted\ over\ [begintime, begintime{+}30\ minutes)_{it}}{Total\ Share\ Volume_{it}}\ . \tag{2}$$

The first half hour of regular trading and the last half hour of regular trading are somewhat unusual, as there is more trading activity at these times, and all U.S. listing exchanges open and close trading with auctions of some sort. Thus, for most of the paper, we consider the trading day as one of three time buckets: the opening half hour (which we call the *open*), the last half hour of continuous trading before 16:00:00 (which we call the *close*) and the interval from 10:00:00 to 15:29:59 (which we call the *middle* of the trading day, where we aggregate the shorting flow during this five-and-a-half-hour interval).[6] The corresponding intraday measures are defined for each stock i and day t as:

$$SSOPEN_{it} = \frac{Shares\ Shorted\ over\ [9{:}30, 10{:}00)_{it}}{Total\ Share\ Volume_{it}}\ ,$$

[6] For simplicity we exclude from the sample the small number of trading days just before major holidays which have an early 1:00pm stock market close.

$$SSMIDDLE_{it} = \frac{Shares\ Shorted\ over\ [10{:}00,15{:}30)_{it}}{Total\ Share\ Volume_{it}},$$

$$SSCLOSE_{it} = \frac{Shares\ Shorted\ over\ [15{:}30,16{:}00)_{it}}{Total\ Share\ Volume_{it}}. \quad (3)$$

Table 1 provides summary statistics about shorting during different parts of the trading day. All the statistics are computed over days and stocks. There are 3,460,790 observations for the CBOE sample in Panel A, and 3,001,383 observations for the FINRA sample in Panel B. Researchers using recent U.S. short sale data (such as Wang, Yan and Zheng, 2020) find that around 40% of all reported off-exchange share volume during regular trading hours involves a short seller. In Panel A, the CBOE's short-sale market share is around 9.37% of CRSP trading volume. Given that CBOE's total volumes are around 20-25% of CRSP trading volumes, the CBOE's short-sale market share is about 30-45% of its total trading volume share. In Panel B, FINRA short sales average 44% of total FINRA trading volume.

In terms of intraday shorts in Panel A, 0.73% of CRSP trading volume happens during the first half hour (less than 10% of overall CBOE shorts), around 2.24% of CRSP trading volume happens during the last half hour (around 20% of overall CBOE shorts), and 6.40% occurs between 10:00 and 15:30 (accounting for about 70% of overall CBOE shorts). Similar patterns also exist for FINRA shorts in Panel B.

We present the intraday shorting patterns for each 30-minute interval in Figure 1. In Panel A of Figure 1, on-exchange shorting follows a U-shape over the course of the trading day. As seen in Table 1, the 30-minute intervals for the opening, the middle of the day, and near the close account for an average of 0.7%, less than 0.5%, and above 2.0% of daily CRSP share volume, respectively. In Panel C of Figure 1, FINRA intraday shorting flows a similar U-shape over the course of the trading day. In Panel B and Panel D of Figure 1, we plot the time series of average short-selling of opening, middle, and closing half-hours over our sample period for CBOE and

FINRA, respectively. There are no strong time trends in either panels. From Panel B, we observe CBOE shorting flowing dips on the third Friday on every March, June, September and December. This is likely because those trading days are triple-witching days when futures and options expire.

3.2 Separating retail vs. institutional off-exchange trading involving a short seller

According to previous literature, such as Boehmer, Jones and Zhang (2008), retail short-sellers and institutional short-sellers might possess different types of information, might be informed to a different extent, and might trade differently. In this section, we adopt a novel algorithm from Boehmer, Jones, Zhang and Zhang (2021) to separate retail and institutional short sale order flows using FINRA off-exchange data.

For regulatory reasons related to the minimum price increment rules associated with Regulation NMS, there are almost no institutional off-exchange trades that take place at non-midpoint subpenny prices over the past decade or so. As a result, off-exchange trades that are reported at non-midpoint subpenny prices during our sample period are virtually certain to be the result of a marketable retail order handled by an internalizing or wholesaling broker and given a small amount of price improvement.

We thus partition our off-exchange trades involving a short seller into three non-intersecting buckets: one that we categorize as marketable *retail sell* order flow based on *lower-than-midpoint* subpenny trade prices, one that we categorize as marketable *retail buy* order flow based on *higher-than-midpoint* subpenny trade prices and the rest, which consists of executions at round penny and half-penny prices. Here we identify short sales involving "retail sales" using "lower-than-midpoint" subpenny trade prices. For internalizers and wholesalers to attract retail sells, they normally offer a price slightly higher than the round penny (or "lower-than-midpoint"), and the price is normally 0.01 to 0.40 cents above the round penny. In these types of trades, the retail investors are shorting the stocks. Using a parallel logic, we identify "retail buy" by "higher-than-midpoint" subpenny trade prices. For internalizers and wholesalers to attract retail buys, they normally offer a price slightly lower than the round penny (or "higher-than-midpoint") and the price is normally 0.01 to 0.40 cents below the round penny. In these trades, somebody else (most likely a market-maker) is on the short side, while the retail investor is on the buy side.

Most of the off-exchange trades at round and half-penny prices are dark pool and crossing network executions which are predominantly institutional, as well as some remaining retail order flow that is handled by an internalizer or wholesaler and is price improved by a half or whole penny or executes at a round penny because it is not price improved at all. We follow the algorithm from Boehmer, Jones, Zhang and Zhang (2021) and label all round penny and close to half-penny executions as *institutional*,[7] even though that trading volume category surely includes some executed retail order flow. In this case, it is hard to say whether the institution is on the short side or on the long side or on both sides. But the short-sell more likely involves institutions rather than retail investors.

As mentioned in the introduction, for shorting involving "retail sells", the identification algorithm can relatively precisely identify the investor identity, and the trade direction, which potentially indicates negative opinions from these retail investors. In contrast, short sales involving institutions can identify the investor type but often cannot discern the trade direction. That is, the identification of a "short sale involving retail sell" might be a more precise indicator for negative information than a "short sale involving institutions".

[7] BJZZ (2021) label execution prices with subpenny fractions between \$0.0001 and \$0.0040 as retail sells, between \$0.0041 and \$0.0060 as institutional, and between \$0.0061 and \$0.0099 as retail buys. Our results are virtually identical if we adopt a wider definition of retail trades that assigns all trades that are not at a round penny or a half-penny to the retail category based on the amount of price improvement provided.

For each stock i on day t, we define proportional off-exchange shorting involving institutional investors and retail investors respectively as:

$$SS_{inst,it} = \frac{Off\text{-}exchange\ Shares\ Shorted\ Involving\ Institutions_{it}}{Total\ FINRA\ Share\ Volume_{it}},$$

$$SS_{retailsell,it} = \frac{Off\text{-}exchange\ Shares\ Shorted\ Involving\ Retail\ Sell_{it}}{Total\ FINRA\ Share\ Volume_{it}},$$

$$SS_{retailbuy,it} = \frac{Off\text{-}exchange\ Shares\ Shorted\ Involving\ Retail\ Buy_{it}}{Total\ FINRA\ Share\ Volume_{it}}. \qquad (4)$$

Following equation (2) and (3), we further define institutional and retail shorting flows during the three parts of the trading day.

Table 1 Panel C provides summary statistics on the intraday shorting flows involving institutions and individuals. Institutional shorting flow is the majority of the total off-exchange shorting, accounting for 36.04% of FINRA daily trading volume, while shorting flow involving retail investor buys (sells) accounts for about 5.12% (2.76%) of FINRA daily trading volume. Within the day, institutional shorting during the first half hour accounts for 3.18% of total daily FINRA trading volume, while shorting during the middle hours and closing half hour accounts for 25.18% and 7.69%, respectively, of total daily off-exchange trading volume. A similar U-shape pattern is also observed for retail shorting flows. Figure 1 Panel C shows that the U-shape of short-sale volume shares over the trading day is fairly modest for retail order flow and is much more pronounced for off-exchange trades that involve a short seller and are likely to involve an institution. Time trends in the partitioned data over the past five years are similarly modest and correspond closely to the time trends discussed earlier in the paper.

4. Empirical Results

We introduce our benchmark specifications in Section 4.1. Results using CBOE and FINRA data are presented in Section 4.2 and 4.3, respectively. In Section 4.4, we further examine the predictive pattern of shorts involving institutions and retail investors. Section 4.5 provides results for intraday return predictions. We connect short-selling and information in earnings announcements in Section 4.6. We offer our interpretation for the prediction results in Section 4.7.

4.1 Benchmark specifications

To examine the informativeness of shorting flow during our 2015-2019 sample period for future returns, we begin with a simple benchmark regression similar to the one in Boehmer, Jones and Zhang (2008):

$$ret_{i,t+d} = b_0 + b_1 shorting_{it} + b_2{}' control_{it} + e_{i,t+d}. \qquad (5)$$

Here the dependent variable, $ret_{i,t+d}$, is the daily return for stock *i* on day *t*+*d*, with *d*=1 to 5, computed using daily closing bid-ask average prices to avoid bid-ask bounce. In the original Boehmer, Jones and Zhang (2008) specification, the shorting measure $shorting_{i,t}$ is directly one of the SS variables defined in equation (1). In our case, we focus on shorting flow for different intraday hours, such as the open, middle and closing parts of the trading day. To mitigate potential effects of any time trend in shorting prevalence and to reduce the effects of outliers, we adopt a rank transformation, as in Cao and Narayanamoorthy (2012) and Livnat and Mendenhall (2006). That is, for stock *i* on day *t*, we first rank SS_{it} cross-sectionally into 100 groups, from the lowest to the highest. Then we use the rank variable divided by 100 as a new shorting flow variable, RSS_{it}. The regression coefficient on this rank variable can be intuitively interpreted as the effect of changing the shorting variable from the 1st percentile to the 100th percentile, and it is comparable for different shorting flow variables we examine.[8] In our tables, we report the rank variable results; the results based on the original measures are qualitatively similar and are available on request. If

[8] Similarly, 0.5* the regression coefficient of the rank variable captures the effect of moving the shorting variable from the 25th percentile to the 75th percentile. We usually interpret our regression results using such an interquartile move.

the shorting flow variable contains relevant information about the next few days and can predict near term returns $ret_{i,t+d}$, the coefficient b_1 should be negative and significant.

Following previous literature, our $control_{it}$ variables include the following variables measured for stock *i* on day *t*: the log market capitalization at the most recent quarter end, *Lsize*; log book-to-market ratio at the most recent quarter end, *Lbm*; the previous month's daily return volatility following Ang et al. (2006), *Volatility*; previous day's return, *Ret [-1]*; the return over the past six months *Ret [-120,-21]*; the return over the past month *Ret [-20,-2]*; and last month's consolidated trading volume as a fraction of outstanding shares *Turnover*. All estimations in this section are Fama-MacBeth regressions, with one regression estimated per trading day. We conduct inference with Newey-West (1987) standard errors with one lag.

To understand whether the predictive power of short-selling, if any, continues or reverses over longer horizons, we also examine shorts' predictive power over longer horizons using a similar specification:

$$wret_{i,k} = b_0 + b_1 shorting_{it} + b_2^{\prime control_{it}} + e_{i,k}. \quad (6)$$

Here the dependent variable, $wret_{i,k}$, measures the average daily returns over the k-th week after day *t*, with *k* ranging from 1 to 12. For instance, when *k* = 1, $wret_{i,k}$ measures the average daily returns in the first week after date *t* (i.e., from *t*+1 to *t*+5); and when *k* = 12, $wret_{i,12}$ measures the average daily returns in the 12th week after date *t* (i.e., from *t*+56 to *t*+60). Shorting flow over longer horizons is more likely to reflect short sellers' private information about fundamentals rather than information trading on anticipated order flow or other short-horizon microstructure effects. If the predictive power of the shorting variables extends to longer horizons, we expect coefficient b_1 to be negative and significant at that horizon. Equation (6) is estimated every day, and we use eight Newey-West lags to account for the overlap that is present in these daily regressions.[9]

4.2 Predicting returns using on-exchange short flows from CBOE

We test the potentially competing theoretical model hypotheses in Section 2 by partitioning our short sales into SSOPEN, which are short sales near the open (before 10:00:00), short sales during the middle of the trading day (10:00:00 to 15:29:59, denoted as SSMIDDLE), and short sales near the close (15:30:00 to 13:59:59, denoted as SSCLOSE). To examine their different predictive power for future returns, we estimate the Fama-MacBeth regressions defined in equation (5). Now the independent $shorting_{it}$ variable is an intraday shorting measure after the rank transformation: RSSOPEN, RSSMIDDLE and RSSCLOSE, along with their institutional or retail counterparts when we use FINRA data.

We report the estimation results using overall CBOE intraday shorting flow variables in Table 2 Panel A. In regression I, the b_1 coefficient of -8.9 basis points (t-stat = -5.88) means that moving through the cross-sectional shorting distribution near the open from quartile 1 to quartile 3 increases the normalized shorting variable by 0.5 and reduces next-day returns by -8.9 * 0.5 = 4.5 basis points (about 11% annualized). CBOE trades involving a short seller in the middle of the day also significantly predict next-day returns, with a coefficient on shorting prevalence of -0.091 and a t-stat of -7.21. CBOE trades near the close continue to predict next day return with a coefficient of -0.066, and a t-stat of -4.85. In the last regression, we include all three shorting flow variables. All three of them maintain their negative signs and statistical significance. Given possible correlations among the three measures, the magnitudes of the coefficients are slightly smaller than in the first three regressions, and RSSOPEN has the largest magnitude coefficient of

[9] In particular, we use the typical textbook rule of $0.75*T^{1/3}$, where T=1364, to calculate the lag = 8.25 (Andrew 1991). Our results are qualitatively similar if we use five lags.

-0.069. It seems that all CBOE intraday shorting flows can predict next day returns, with somewhat stronger results for short sales near the open.

For the control variables, the coefficients on previous-day and previous-month returns are both negative and strongly significant. The coefficient on the previous-day return is -0.764, indicating an estimated reversal of over half of the previous day's return. This is even more notable, given that we use closing bid-ask average prices throughout the paper to calculate returns, which should limit or eliminate reversals from microstructure sources such as bid-ask bounce. The coefficient on the previous-month return is a smaller -0.179 in the base case regression I, indicating a smaller reversal magnitude, but the coefficient has a strongly significant t-stat of -3.40. Most of the other control variables are insignificant, except for lagged turnover, which is negative and sometimes borderline significant.

To make sure the results are not driven by outliers in the time-series, we present the time-series of Fama-MacBeth regression coefficients in Figure 2 Panel A. There are no large outliers or clear trends over time, so it is unlikely that our coefficients are affected by outliers or time trends.

What if we extend the horizons from 1 day to 12 weeks? These results are reported in Table 2 Panel B. Take RSSOPEN as an example. The coefficients quickly drop by about half after the first day, from -0.089 to -0.0421, and then stay relatively stable around -0.025 over the next 12 weeks. For each of the 12 weeks, the coefficients are negative and are statistically significant in 9 cases. That is, CBOE RSSOPEN seems to be able to predict returns over the next 12 weeks, even though the magnitude is smaller than that of the next day predictions. In other words, RSSOPEN can predict both short and long-term information, while the predictive power is much stronger over the short term. The pattern of RSSMIDDLE is quite similar to RSSOPEN, but with slightly smaller coefficients. In contrast, the RSSCLOSE variable loses its statistical significance after the second day. Even though the coefficients stay around -0.010 over the next 12 weeks, only 1 of these 12 weeks is statistically significant.

To summarize, the evidence from the CBOE exchanges indicates that all three intraday shorting flow variables can predict returns for the next day but the predictive power is strongest for shorting near the open, consistent with the aggressive trading hypothesis from the Holden and Subrahmanyam (1992) model.[10] The greater predictability of opening shorts versus the weaker predictability of the closing shorts also provides support for Slezak (1994). RSSOPEN and RSSMIDDLE have significant predictive power for future returns over the next 12 weeks. The long-term predictive power of shorting near the open and in the middle of the day is consistent with Kyle (1985) model.

4.3 Predicting returns using off-exchange short flows from FINRA

Compared with the CBOE on-exchange shorting flows, the FINRA data are different in two aspects. First, the off-exchange data, due to its opacity, might be preferred by different types of investors and might display different patterns of predictive power. Second, we can separate the off-exchange short-sales into institutional trades and retail trades and study each group's behavior separately.

We report the estimation results using overall off-exchange intraday shorting flow variables in Table 3 Panel A. In regression I, short sales near the open are the most informed about next-day returns. The b_1 coefficient of -6.2 basis points (t-stat = -5.20) means that moving through the cross-sectional shorting distribution near the open from quartile 1 to quartile 3 increases the normalized shorting variable by 0.5 and reduces next-day returns by -6.2 * 0.5 = 3.1 basis points

[10] Foster and Viswanathan (1990) and Bernhardt and Miao (2004) also have similar predictions that informed traders will trade more aggressively to start and less aggressively later on. We use the Holden and Subrahmanyam (1992) model as the representing model of similar models.

(about 7.8% annualized). Off-exchange trades involving a short seller in the middle of the day also significantly predict next-day returns, with a coefficient on shorting prevalence of -0.027 and a t-stat of -3.20. Surprisingly, off-exchange trades near the close have a coefficient of 0.025, with a significant t-statistic of 2.42. That is, higher shorting at the close actually is followed by a higher next day return, which is opposite to the prediction that shorts are negatively informed about the cross-section of future returns. When the three intraday shorting flow variables are included together in the last regression, all maintain their original signs and significance, except the coefficients are slightly smaller. If we compare the coefficients with those in Table 2 Panel A, the shorting flow near the open still has the strongest predictive power, but the off-exchange version has a smaller magnitude than its on-exchange counterpart. This suggests that aggressive short sales near the open value the immediacy offered by exchanges as opposed to off-exchange venues. In addition, the predictive power of FINRA shorting near the close is opposite to that of analogous shorting on CBOE exchanges, and we further examine this pattern using different groups of investors later in the paper.[11]

Table 3 Panel B presents the predictability of shorting during different parts of the trading day for returns at longer horizons. For RSSOPEN, when we extend the window from 1 week to 12 weeks, the coefficients stay around -0.020, and 10 out 12 coefficients are statistically significant. For RSSMIDDLE, the coefficients are slightly smaller, mostly around -0.015, and they are significant in 9 out of 12 weekly horizons. For RSSCLOSE, as mentioned above, the coefficient is positive and significant for the 1-day horizon, but it quickly turns negative. For one to 12-week horizons, all RSSCLOSE coefficients are negative, but only one is statistically significant. In comparison with the results in Panel B of Table 2, the predictive power of off-exchange RSSOPEN

[11] As we will see in later decompositions, the positive predictability for future returns of off-exchange shorting flow is mainly driven by short sales at the close that involve a retail buy.

and RSSMIDDLE is weaker but exhibit similar patterns, while the predictive power of RSSCLOSE is mostly insignificant. That is, if we combine all investor groups, the results from FINRA are mostly consistent with those from CBOE, in the sense that the long term predictive power of the open and midday is consistent with the Kyle (1985) model, while the short term prediction of the three intraday measures and their different predictive powers are consistent with the aggressive trading hypothesis from the Holden and Subrahmanyam (1992) model.

4.4 Predicting returns using FINRA short flows from different investor groups

In this subsection, we use our retail trade identification algorithm to distinguish between off-exchange short sale return predictability involving retail traders and off-exchange short sale return predictability that is more likely to involve institutional traders. We construct rank variables for institutional, retail sell and retail buy shorting flows, similar to RSS_{it}. The results are reported in Table 4. Panels A to C present results on institutions, retail sells, and retail buys, respectively. As discussed in earlier sections, among "shorts sales involving institutions", "short sales involving retail sell" and "short sales involving retail buys", the identification algorithm can relatively precisely identify investor identity, and trade directions for the latter two, while for the first one, it can generally only identify investor identity reliably. That is, the identification of "short-sell involving retail sell (or buy)" might contain more directional information than "short-sell involving institutions".

From the first row of Panel A, we find that short sales involving institutions are strongly informed and predict price declines over the near future, which is in line with the findings of Boehmer, Jones, and Zhang (2008). The estimated b_1 coefficient of -0.054 in the first row and first column means that an increase in short selling from quartile 1 to quartile 3, for example, results in next-day returns that are -5.4 bps * 0.5 = 2.7 basis points lower for short sales involving institutions (with a t-statistic of 5.21), or about 6.8% annualized. For the intraday measures, the open shorting flow coefficient is -0.071 (t-stat=-5.82), the middle of the day shorting flow coefficient is -0.071 (t-stat=-5.82), and the shorting flow coefficient near the close is 0.007 (t-stat=0.65). That is, off-exchange shorting involving institutions near the open and during the middle of the trading day are much more predictive for returns over the next day. If we extend to longer horizons to up to 12 weeks, the overall shorting flow involving institutions can predict returns up to 5 weeks. Again, shorting near the open and during the middle of the day are much more predictive for future returns, over the longer term, than shorting flow near the close. In fact, the RSSCLOSE variable is never significant. It is possible that the off-exchange trades near the close are dominated by large institutions passively rebalancing their portfolios at the end of the day, and thus do not predict future returns because they are not information-based trades.

For the retail sell RSS measures in Panel B of Table 4, we find significant negative predictive power of retail sells for future stock returns. For the 1-day horizon in the first row, the coefficient is -0.038 with a t-statistic of -4.28. This coefficient is smaller than that in Panel A, indicating the predictive power of retail sells is weaker than those from institutions, which is again consistent with Boehmer et al (2008). For the intraday measures predicting the next-day returns, the open shorting flow coefficient is -0.065 (t-stat=-4.93), the midday shorting flow coefficient is -0.034 (t-stat=3.72), and the close shorting flow coefficient is -0.019 (t-stat=-1.84). Again, shorting near the open is more predictive for next day returns than shorting at other times of day, consistent with Holden and Subrahmanyam (1992). When we turn to longer horizons, it is interesting to find that shorts in all three time buckets involving retail sells all contain information for longer horizons and they all can predict returns over the next 12 weeks. This long-term predictability of shorting involving retail sells throughout the day implies that retail short sales may have private information that is of monopolistic nature and behave like the dynamic version of Kyle (1985). If we compare the magnitudes and statistical significance, retail short sales near the open dominate the other two parts of the day.

Finally, Panel C shows results for off-exchange shorting flow involving retail buys. As mentioned earlier, in most buyer-initiated retail shorting that we identify with our algorithm, these trades are likely to be a retail purchaser submitting a marketable buy order that interacts with a wholesaler who happens to be short selling and provides a small amount of price improvement relative to the National Best Offer. If that retail buy is informed, on average the next-day return would be positive, not negative. In the first row for 1-day horizon, the coefficient on daily retail buys is 0.026 is with a t-statistic of 2.48, consistent with the finding in BJZZ (2021). The coefficient is negative for shorts near the open but positive and significant for the other two parts of the trading day. Referring back to the positive coefficient of RSSCLOSE for next day return in Table 3, it is likely driven by the fact that the institution close is positive and insignificant, retail sell at close is negative and insignificant, while the retail buy at close is positive and highly significant. At longer horizons, most of the coefficients on retail buy shorting flow variables turn negative for future returns and become statistically significant for horizons longer than 7 weeks. That is to say, these retail buy trades actually predict returns in the wrong way over longer horizons, which indicates that retail buys might have an information disadvantage over the long run. Since our focus are mostly on short sale patterns, rather than buying behavior, for the rest of the paper we focus mainly on short sales involving an institution as well as retail short sells.

With detailed data on different groups involving short sales, we find that those involving institutions have strong and negative predictive power for future returns over both the short term and long term, especially for shorting flows near the open and during the middle of the trading day.

Shorting flows involving retail sells predict returns negatively over the short and long term, no matter whether we use open, middle or close, which indicates these investors are truly informed about the cross-section of future returns, and also that the measure is probably more precise than the analogous measure for institutions. Shorting flows involving retail buys predict returns positively over the short term, which still indicates an information advantage given that they are buyer-initiated orders, but they predict returns negatively over the longer term, implying an information disadvantage at these horizons.

4.5 Using intraday shorting flow to predict intraday returns

In this section, we directly investigate the intraday predictive patterns of short-selling. In particular, the returns at intraday horizons help us understand how quickly share prices incorporate the information short sellers trade on. We construct empirical estimates of same-day predictability by dividing the trading day into 30-minute intervals, conducting our predictive regressions on future same-day returns, and assessing the results and significance for short sales at varying times of day. Our specification is:

$$ret_{i,j+2,12} = b_0 + b_1 shorting_{ij} + b_2{}' control_{ij} + e_{i,j+1,12}. \qquad (7)$$

That is, we use shorting variable from 30-minute interval *j* to predict the average 30-minute return from time j+2 (that is, starting 30 minutes after the shorting measurement) until 15:30 (30-minute interval 12). We skip one 30-minute interval between shorting at j, and return at j+2, so that any short-term temporary price impact from microstructure effects that is directly associated with the short sales has time to dissipate. We exclude the bucket from 15:30 to the close in order to avoid any unusual trading activity that might be associated with the closing auction and associated price moves.

The estimation results for each half-hour interval are reported in Table 5, and we present separate results for CBOE, FINRA institutions and FINRA retail sells. The first half-hour interval of short sales consists of short sales that take place before 10:00:00, and this interval coincides with our earlier definition of short sales near the *open*. The results for this interval show that same-day returns are much lower after 10:00:00 for stocks that are heavily shorted in the first half-hour of trading. Taking CBOE as example, the estimated coefficient of -0.018 means that 30-minute returns for the rest of the day are on average -1.8 bps * 0.5 = 0.9 basis points lower when moving from shorting quartile 1 to quartile 3, with a strongly significant t-statistic of -15.55. This coefficient reflects an average half-hour return, and since there are 10 half-hour buckets that we use for these particular same-day predicted returns, it also implies a rest-of-the-same-day return that is 0.9 * 10 = 9 basis points lower moving from quartile 1 of shorting prevalence to quartile 3. This magnitude is double that reported in Table 2 Panel A for next day return, indicating that the first half hour has quite strong predictive power for the rest-of-same-day return. The results for shorting near the *open* are consistent with short-horizon private information that is quickly incorporated into price. Other half-hour buckets from morning activity show qualitatively similar results with uniform statistical significance, though the coefficient magnitudes decline more or less monotonically toward zero as the trading day passes, which is consistent with Holden and Subrahmanyam (1992)'s prediction that informed traders will trade more aggressively to start and less aggressively later on.

For FINRA institution and FINRA retail sell shorting flow variables, the patterns are similar, but the magnitudes are smaller. That is, the immediate price reaction is larger on exchange rather than off exchange. It is interesting to notice that for FINRA shorts involving institutions, the coefficient becomes insignificant starting from the shorting interval 13:00-13:30, while the coefficient for FINRA retail sell only becomes insignificant for shorting interval 14:00-14:30. This pattern suggests that the retail sells, even at later trading hours, still contain significant negative information, while the short sales involving institutions do not.

4.6 Short-Sellers' information advantage vs. earnings news

In previous sections, we establish that short sellers who trade near the opening bell possess short-term and long-term information that is reliably incorporated into price in the future. It is important to understand the nature of this information, and whether the short-sellers' predictive power for future returns comes from better access to private information or better skills for processing public information. Given that our study is at the firm level, the most important firm level information events are earnings news releases, such as firms' quarterly earnings announcements, and analyst recommendation changes. In this section, we investigate how short-sellers' intraday trading behavior is related to these earnings news releases.

We obtain data on earnings news releases from the IBES database, also with a 2015-2019 sample period. As mentioned earlier, these news releases mostly happen after trading hours. So when short-sellers trade during trading hours, they could be processing public news released after the previous day's close, or they could have access to private news not yet released, and they anticipate a news release after the same day's close. In the latter case, before the news is publicly released, we consider it *private* news.

Here we establish the relation between the intraday shorting flows and earnings news, either in the sense of processing news releases from the previous day or predicting news releases in the future. In particular, for the public news processing hypothesis, we estimate the following specification:

$$RSS_{i,t} = b_0 + b_1 pre\_bad_{i,t-1} + b_2' control_{ij} + e_{i,t}. \qquad (8)$$

Here $pre\_bad_{i,t-1}$=1 if and only if there is a bad earnings or management guidance or analyst-related information event on day t-1 that is released to the public after market close on day t-1, but before the market opens on day t. We define an after-hour news release to be "bad" if it generates negative market adjusted returns over the next day. Depending on whether the news is released by the firm or by the analyst, we further define two parallel variables, $pre\_badfirm$ and $pre\_badanalyst$. If short-sellers effectively process the bad news release the day before, we expect the coefficient b_1 to be significant and positive, because bad news should lead to more short-selling as short-sellers process the negative news. Among open, middle and close, following predictions from Foster and Viswanathan (1995), the open should be the most related intraday shorting flow measure to the previous day after-hour news releases, because it is the closest to the previous day news release.

In parallel, to examine whether short-selling predicts arrival of future negative news, we estimate the following specification:

$$RSS_{i,t} = b_0 + b_1 post\_bad_{i,t+1} + b_2{}' control_{ij} + e_{i,t}. \qquad (9)$$

Here $post\_bad_{i,t+1}$=1 if and only if there is a bad earnings event or management guidance or analyst-related information event on day t that is released to the public after market close on day t, but before the market opens on day t+1. Again, we define an after-hour news release to be "bad" if it generates negative market adjusted returns over the next day. If short-sellers can predict the forthcoming bad news, we expect the coefficient b_1 to be significant and positive. That is, if short-sellers have an information advantage and anticipate the arrival of bad news, they would increase shorting before the bad news release. Among open, middle and close, if we believe that the likelihood of receiving a signal increases as the signal release time approaches, then shorting at the close should be the most related intraday measure to the same-day after-hours news release, as it is closest to the information release time. To estimate equation (8) and (9), we estimate Fama-MacBeth regressions and adjust standard errors with one lag.

The estimation results are presented in Table 6. For each panel, the first three rows present estimation results for equation (8), and the last three rows for equation (9). In terms of processing previous day news releases, taking the first row in Panel A as an example, shorting near the open is positively and significantly related to previous day after-hours bad news releases, as expected. More negative news leads to an immediate increase in shorting. Interestingly, shorting flow during the middle of the day and near the close are both negatively related to the bad news release, which is likely driven by intraday return reversal patterns. Between bad news release by firms and by analysts, they also are positively related to open, but negatively related to middle and close, all indicating that negative news releases are traded aggressively and quickly incorporated by shorting flow near the next open. When we turn to Panel B, where we use shorting flow from FINRA institutions, we find similar patterns. Interestingly, for FINRA retail sells in Panel C, the bad news release from the previous trading day is positively related to retail sells at all three times of day, implying that it takes retail sells a fairly long time to adjust their trading and incorporate the negative information.

For predicting the arrival of negative news releases, in the last three rows of Panel A, we find the coefficients on post_bad and post_badanalyst is significant and positive for RSSOPEN, indicating that CBOE shorting near the open seems to be informed about upcoming bad news released post close, in particularly the analyst-issued news. Post_badfirm is significant and positive only for RSSCLOSE but not for other shorting variables, which suggests that CBOE shorting is not informed about bad news to be issued by the firm post-close until the very end of regular trading hours.

When we turn to FINRA institution shorting in Panel B, we find their predictive power for upcoming bad news is stronger. For post-close bad news releases, the coefficient on post_bad for both RSSOPEN and RSSCLOSE is positive and significant, suggesting that the institutions trading off-exchange are likely informed about the bad news that will arrive after the market closes. Interestingly, when we look at analyst-issued vs. firm-issued news releases in the last two rows, the analyst-related bad news that will be released post-close is anticipated and incorporated by off-exchange intuitions trading near the open (i.e., at least six hours later than when the shorting takes place), whereas firm-issued post-close bad news is significantly associated with institutions' trading near the close.

The pattern for retail sells in last three rows of Panel C is different from those in Panel A and B. We actually find that retail sells in all three parts of the trading day are significantly and positively related to future bad news releases (in the fourth row), indicating some of these retail investors probably are insiders or have access to some other private firm-specific information. It is also likely that the identification algorithm more precisely captures retail investors' shorting than it does for institutions. Regarding analyst-issued vs. issuer-based news, we find results for retail short sales are similar to the results of shorting involving institutions in the sense that post-close analyst-issued news seems to be more related to shorting near the open than the close, whereas post-close firm-issued news is more related to shorting near the close than the open.

Overall, the evidence in this section is quite interesting. First, shorting flow near the open, across trading venues and trader types, reacts strongly to the bad news released during the previous overnight period and oftentimes anticipates bad news to be released on the same day post-close, providing further support for the models of Holden and Subrahmanyam (1992) and Foster and Viswanathan (1990) in which informed traders trade fairly aggressively early on, competing with each other before the signals they receive become public or stale. Shorting in the middle of day is mostly negatively related to bad news events issued during the previous overnight period, with retail shorting flow on FINRA being an exception, indicating slow processing of negative public news. Shorting near the close predicts only the arrival of the post-close firm-issued bad news, which suggests that as the release time approaches, the likelihood that short sellers become informed of firm-issued news increases.

4.7 Interpretation of results

How should we interpret the empirical results economically, and also in light of the theoretical models? So far we have two main sets of results. First, shorting near the open and during the middle of the day has strong predictability for short horizons, but it quickly decreases after one day, while maintaining its negative sign and statistical significance. Second, shorting near the close has mixed signs for predicting next day return, and does not predict returns at longer horizons except shorting near the close that involves retail sells.

We believe the above patterns are consistent with a sorting model, where informed short-sellers trade in a way that reflects the horizon of their private and public information as well as the competition they face in trading on that information. When information is long-lived, short sellers act as though they do not face much competition from other informed traders. Short sellers trade gradually over the trading day and over several trading days. As a result, this information is only gradually incorporated into share prices. These empirical results are consistent with the dynamic version of Kyle (1985).

In contrast, when information is short-lived, the empirical results are more consistent with a Holden and Subrahmanyam (1992) setup, where competing informed traders trade quite aggressively in early trading rounds and much less aggressively later on, as the information release time approaches or correlated signals become stale. As a result, most of the private information is incorporated into prices quite quickly. In particular, our strong same-day and next-day predictability associated with shorting near the open and in the middle of the day suggests that if the private information is obtained by institutions while the market is closed, those institutions trade aggressively on that information near the open once trading resumes, incorporating that information quickly into price, and the information is expected to be released publicly within two trading days.

Shorting near the close, whether on-exchange or off-exchange, does not seem to predict future returns at longer horizons. The only closing shorting flow that has long-term predictability is the shorting flow near the close that involves retail sells. It does predict the cross-section at shorter horizons, but with mixed signs when we consider different venues. This probably suggests that there are reasons other than information that lead short-sellers to trade, and short-sellers may choose different venues depending on their identities and nature of their information.

5. Conclusions

In this study, we consider a recent five-year sample of U.S. on-exchange and off-exchange trading by short sellers throughout the trading day, and we find differential information content of shorts as we vary the time of the short sale during the trading day and as we vary the horizon of the measured cross-sectional return predictability. We can also determine whether the short sale transaction is likely to involve an institution vs. a retail trader and find very different results for these two groups of short sales. We find that during our recent sample period, on-exchange short sales are more informed about the cross-section of future returns than off-exchange shorting flows. In terms of different trading hours, whether on-exchange or off-exchange, short sales before 10am and during the middle of the day strongly predict the next-day return in the expected direction. The predictive power decreases quickly but remains significant for horizons up to 12 weeks. In contrast, shorting flows near the close mostly predict returns only over short horizons.

These results are consistent with several dynamic theory models of informed trading. For example, our next-day return predictability is consistent with a model like Holden and Subrahmanyam (1992) or Foster and Viswanathan (1990) where multiple informed short sellers receive a signal of value while the market is closed. These informed traders begin to trade as soon as the market reopens the next morning and compete aggressively with each other, thereby incorporating most of their private information into price early in the first trading day. This means that shorting near the open, especially shorting on exchanges that offer immediacy, has the strongest predictability for rest-of day returns and next-day returns. Longer-horizon return predictability is present for short sales near the open and in the middle of the day, and this finding is more consistent with a dynamic model such as Kyle (1985), where a single informed trader gradually trades on her information over time.

When we examine earnings news releases, such as earning announcements, manager guidance and analyst recommendation changes, we find that opening shorting across trading venues and trader types reacts strongly to bad news disclosed during the previous overnight period and anticipates most bad news to be released post-close, whereas shorting near the close is associated only with earnings news occasionally issued by firms post-close. This gives a strong indication of at least one of the types of fundamental information that informed short sellers may aggressively trade on near the open.

It is also worth noting that we find no evidence of manipulative behavior by short sellers. All of our predictability results die off as the horizon increases. We never observe price reversals following heavy shorting. Of course, there are certain models of strategic complementarities where reversals would not be observed following abusive or predatory shorting. But our findings significantly narrow the scope for any problematic shorting behavior of this sort.

One caveat of all theory models considered here is that they are all symmetric, in the sense that positive and negative signals of value should have similar effect magnitudes. However, our short sale measures are limited to negative information. Going forward, we think it would be productive to identify similar measures of positive private information to gauge whether the empirical results line up symmetrically behind these theory models, or whether there is something special about negative information and the investors who obtain and trade on it.

Lastly, our sample period is 2015 to 2019, a relatively stable period for capital markets. Starting in 2020, the Covid-19 pandemic and a new generation of retail investors are quickly changing the landscape of trading and the price discovery process. It would be interesting to understand how short-sellers, information and trading interact with each other during this new era. We leave this question to future researchers.

REFERENCES

Andrews, D.W.K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59: 817-858.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. Journal of Finance 61: 259-299.

Asquith, P., P. Pathak, and J. R. Ritter. 2005. Short interest, institutional ownership, and stock returns. Journal of Financial Economics 78: 243-276.

Bernhardt, D. and J. Miao. 2004. Informed trading when information becomes stale. Journal of Finance 59:339-390.

Boehmer, E., C. M. Jones, and X. Zhang. 2008. Which shorts are informed? Journal of Finance 63: 491-527.

Boehmer, E., C. M. Jones, and X. Zhang. 2013. Shackling short sellers: The 2008 shorting ban. Review of Financial Studies 26:1363-1400.

Boehmer, E., C. M. Jones, J. Wu, and X. Zhang, 2020. What do short sellers know? Review of Finance 24: 1203–1235.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang. 2021. Tracking retail investor activity. Journal of Finance, forthcoming.

Boehmer, E., and E. Kelley. 2009. Institutional investors and the informational efficiency of prices. Review of Financial Studies 22: 3563-3594.

Boehmer, E., and J. Wu. 2013. Short selling and the price discovery process. Review of Financial Studies 26: 287-322.

Bris, A., W. Goetzmann, and N. Zhu. 2007. Efficiency and the bear: Short sales and markets around the world. The Journal of Finance, 62(3):1029-1079.

Cao, S. S., and G. S. Narayanamoorthy. 2012. Earnings volatility, post–earnings announcement drift, and trading frictions. Journal of Accounting Research 50: 41-74.

Christophe, S. E., M. G. Ferri, and J. J. Angel. 2004. Short-selling prior to earnings announcements. Journal of Finance 59: 1845-1876.

Christophe, S. E., M. G. Ferri, and J. Hsieh. 2010. Informed trading before analyst downgrades: Evidence from short sellers. Journal of Financial Economics 95: 85-106.

Comerton-Forde, C., C.M. Jones, and T. J. Putnins. 2016. A tale of two types: Shorting at close range. Journal of Financial Economics 121:546-568.

Desai, H., S. Krishnamurthy, and K. Venkataraman. 2006. Do short sellers target firms with poor earnings quality? Evidence from earnings restatements. Review of Accounting Studies 11: 71-90.

Diamond, D. W., and R. E. Verrecchia. 1987. Constraints on short-selling and asset price adjustment to private information. Journal of Financial Economics 18: 277-311.

Diether, K. B., K. Lee, and I. M. Werner 2009a. It's SHO time! Short-sale price tests and market quality. Journal of Finance 64 (1): 37-73.

Diether, K. B., K. Lee, and I. M. Werner. 2009b. Short-sale strategies and return predictability. Review of Financial Studies 22(2): 575-607.

Engelberg, J., A. V. Reed, and M. Ringgenberg. 2012. How are shorts informed? Short sellers, news, and information processing. Journal of Financial Economics 105: 260-278.

Francis, J., M. Venkatachalam, and Y. Zhang. 2005. Do short sellers convey information aboutchanges in fundamentals or risk? Working paper, Duke University.

Foster, F.D. and S. Viswanathan. 1990. A theory of the interday variations in volume, variance, and trading costs in securities markets. Review of Financial Studies 3:593-624.

Goldstein, I. and A. Guembel. 2008. Manipulation and the allocational role of prices. The Review of Economic Studies 75:133-164.

Holden, C.W. and A. Subrahmanyam. 1992. Long-lived private information and imperfect competition. Journal of Finance 47: 247-270.

Irvine, P., M. Lipson, and A. Puckett. 2007. Tipping. Review of Financial Studies 20: 741-768.

Jain, C., P. Jain, and T.H. McInish. 2012. Short Selling: The Impact of SEC Rule 201 of 2010. Financial Review 47:37-64.

Livnat, J., and R.R. Mendenhall. 2006. Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. Journal of Accounting Research 44: 177-205.

Lyle, M.R., R. Christopher, A. Stephan, and T. L. Yohn. 2019. The Speed of the Market Reaction to Pre-Open versus Post-Close Earnings Announcements. Working paper.

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

Karpoff, Jonathan M. and Lou, Xiaoxia. 2010. Short Sellers and Financial Misconduct. Journal of Finance 65, 1879-1913.

Khandani, A. E., and A. W. Lo. 2011. What happened to the quants in August 2007? Evidence from factors and transactions data. Journal of Financial Markets 14 (1): 1-46.

Kyle, A.S. 1985. Continuous auctions and insider trading. Econometrica 53: 1315-1335.

Lamont, O.A. and R.H. Thaler. 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. Journal of Political Economy 111(2):.227-268.

Lehmann, B.N., 1990. Fads, martingales, and market efficiency. The Quarterly Journal of Economics, 105(1): 1-28.

Lo, A.W. and A.C. MacKinlay. 1990. An econometric analysis of nonsynchronous trading. Journal of Econometrics, 45(1-2): 181-211.

Menkveld, A.S., B.Z. Yueshen, H. Zhu. Miller, E. M. 2017. Shades of Darkness: A Pecking Order of Trading Venues. Journal of Financial Economics 124: 503-534.

Mitchell, M., T. Pulvino, and E. Stafford. 2002. Limited arbitrage in equity markets. Journal of Finance 57: 551-584.

Newey, W.K. and K.D. West. 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review 28: 777-787.

Reed, A.V., M. Samadi, and J. S.Sokobin. 2020. Shorting in Broad Daylight: Short Sales and Venue Choice. *Journal of Financial and Quantitative Analysis* 55: 2246-2269.

Roll, R. 1988. R-squared. *Journal of Finance* 43: 541-566.

Slezak, S.L. 1994. A Theory of the Dynamics of Security Returns around Market Closures. *Journal of Finance* 49:1163-1211.

Tetlock, P. C. 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23: 3520-3557.

Wang, X., X. S. Yan, and L. Zheng. 2020. Shorting flows, public disclosure, and market efficiency. *Journal of Financial Economics 135*: 191-212.

Table 1. Summary Statistics

This table presents summary statistics for short sale volume at different times of the day, along with our control variables. Our sample period is Jan 2015 to December 2019, and our sample firms are common stocks listed on NYSE, NYSE MKT, or Nasdaq with a share price of at least $1. We exclude all treated firms during the 2016-2018 tick size pilot program. Panels A and B present variables related to short sale volume from CBOE exchanges and FINRA, respectively. The variable S*SOPEN* (or $SSOPEN_{inst}$, $SSOPEN_{retailbuy}$ or $SSOPEN_{retailsell}$) is short volume (involving institutional or retail buy and sell orders, respectively) from 09:30:00 to 09:59:59 divided by FINRA total trading volume (for FINRA shoring flows) or by CRSP total trading voume (for CBOE shorting flows) in that stock on that day. Similarly, *SSCLOSE* (or $SSCLOSE_{inst}$, $SSCLOSE_{retailbuy}$ or $SSCLOSE_{retailsell}$) is short volume from 15:30 to 16:00 and *SSMIDDLE* is short volume during the rest of the trading day.

	N	Mean	Std	Median	Q1	Q3
Panel A. CBOE shorting						
SS	3,460,790	0.0937	0.0540	0.0880	0.0610	0.1180
SSOPEN	3,460,790	0.0073	0.0140	0.0040	0.0010	0.0090
SSMIDDLE	3,460,790	0.0640	0.0440	0.0580	0.0380	0.0810
SSCLOSE	3,460,790	0.0224	0.0210	0.0200	0.0110	0.0300
Panel B. FINRA shorting						
SS	3,001,383	0.4399	0.1820	0.4350	0.3090	0.5640
SSOPEN	3,001,383	0.0397	0.0530	0.0240	0.0090	0.0510
SSMIDDLE	3,001,383	0.3132	0.1520	0.3000	0.2040	0.4060
SSCLOSE	3,001,383	0.0869	0.0760	0.0690	0.0380	0.1140
Panel C. FINRA shorting by investor groups						
SS_{inst}	3,001,383	0.3604	0.1680	0.3480	0.2390	0.4710
$SSOPEN_{Inst}$	3,001,383	0.0318	0.0430	0.0200	0.0070	0.0410
$SSMIDDLE_{inst}$	3,001,383	0.2518	0.1360	0.2360	0.1550	0.3310
$SSCLOSE_{inst}$	3,001,383	0.0769	0.0690	0.0600	0.0320	0.1020
$SS_{retailsell}$	3,001,383	0.0276	0.0360	0.0190	0.0090	0.0340
$SSOPEN_{retailsell}$	3,001,383	0.0028	0.0110	0.0000	0.0000	0.0020
$SSMIDDLE_{retailsell}$	3,001,383	0.0213	0.0320	0.0140	0.0060	0.0260
$SSCLOSE_{retailsell}$	3,001,383	0.0035	0.0110	0.0010	0.0000	0.0040
$SS_{retailbuy}$	3,001,383	0.0512	0.0620	0.0340	0.0160	0.0630
$SSOPEN_{retailbuy}$	3,001,383	0.0051	0.0160	0.0010	0.0000	0.0040
$SSMIDDLE_{retailbuy}$	3,001,383	0.0396	0.0530	0.0250	0.0110	0.0490
$SSCLOSE_{retailsbuy}$	3,001,383	0.0065	0.0180	0.0030	0.0000	0.0070

Table 2. Predicting future returns using intraday shorting flows from CBOE exchanges

Table 2 presents Fama-MacBeth regression coefficients of future returns on previous day short flow variables, as in equation (5) and (6). *RSSOPEN (RSSMIDDLE/RSCLOSE)* is a rank variable for shorting flow variable *SSOPEN (SSMIDDLE/SSCLOSE)*, computed as shares shorted on CBOE exchanges over the first half hour of the trading day (middle 4.5 hours/last half hour) over daily CRSP share volume. To compute the rank, we sort all stocks by shorting flow variables into 100 groups, and assign the rank number to the variable. The RSS is computed as rank variables divided by 100. *Ret [-1]* is the return on the previous day using closing bid-ask averages. *Ret [-20, -2]* is the cumulative daily return in the [-20, -2] window. *Ret [-120, -21]* is the cumulative daily return over the [-120, -21] window. *Turnover* is the monthly average of daily turnover measured at the most recent month end. *Volatility* is the monthly volatility of daily return at the most recent month end. *Lsize* is log of market capitalization and *Lbtm* is the log of book to market ratio, both at the most recent quarter end. To account for potential serial correlation in the coefficients, the standard errors of the time-series are adjusted using Newey-West (1987) with eight lags for weekly returns and one lag for daily returns. All regression coefficients are multiplied by 100 for presentation purposes.

Panel A. predict next day return

	I		II		III		IV	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Intercept	0.031	1.05	0.051	1.75	0.037	1.28	0.056	1.95
RSSOPEN	-0.089	-5.88					-0.069	-4.89
RSSMIDDLE			-0.091	-7.21			-0.066	-6.12
RSSCLOSE					-0.066	-4.85	-0.042	-3.27
Ret [-1]	-0.764	-3.86	-0.771	-3.89	-0.787	-3.97	-0.751	-3.80
Ret [-20,-2]	-0.179	-3.40	-0.181	-3.44	-0.183	-3.47	-0.174	-3.32
Ret [-120,-21]	-0.013	-0.55	-0.013	-0.55	-0.011	-0.49	-0.013	-0.57
Turnover	0.000	-1.32	0.000	-1.37	0.000	-1.58	0.000	-0.99
Volatility	0.110	0.27	0.066	0.17	-0.062	-0.15	0.094	0.24
Lsize	0.007	1.92	0.004	1.26	0.005	1.57	0.009	2.74
Lbm	0.007	1.09	0.008	1.15	0.009	1.36	0.008	1.26
Adj.R2	0.04		0.04		0.04		0.04	

Panel B. Predict daily and weekly returns

	RSS		RSSOPEN		RSSMIDDLE		RSSCLOSE	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Day 1	-0.106	-8.18	-0.089	-5.88	-0.091	-7.21	-0.066	-4.85
Day 2	-0.033	-2.61	-0.041	-2.68	-0.031	-2.52	-0.025	-1.91
Day 3	-0.013	-1.04	-0.020	-1.35	-0.015	-1.16	-0.015	-1.19
Day 4	-0.016	-1.32	-0.026	-1.72	-0.019	-1.54	-0.014	-1.01
Day 5	-0.021	-1.70	-0.021	-1.39	-0.025	-2.07	-0.011	-0.87
week 1	-0.038	-3.70	-0.039	-3.16	-0.036	-3.63	-0.026	-2.43
week 2	-0.009	-0.95	-0.023	-1.90	-0.008	-0.91	-0.012	-1.14
week 3	-0.019	-2.08	-0.023	-1.83	-0.019	-2.11	-0.018	-1.77
week 4	-0.029	-3.02	-0.030	-2.32	-0.029	-3.06	-0.018	-1.79
week 5	-0.025	-2.76	-0.026	-2.12	-0.025	-2.68	-0.018	-1.94
week 6	-0.024	-2.47	-0.025	-2.05	-0.025	-2.56	-0.012	-1.22
week 7	-0.022	-2.31	-0.037	-3.05	-0.025	-2.67	-0.009	-0.93
week 8	-0.024	-2.61	-0.034	-2.65	-0.026	-2.75	-0.014	-1.44
week 9	-0.011	-1.15	-0.032	-2.58	-0.013	-1.33	-0.010	-0.89
week 10	-0.022	-2.01	-0.028	-1.97	-0.025	-2.36	-0.010	-1.01
week 11	-0.023	-2.17	-0.037	-2.75	-0.025	-2.31	-0.013	-1.13
week 12	-0.020	-1.92	-0.022	-1.59	-0.022	-2.14	-0.013	-1.20

Table 3. Predicting future returns using intraday shorting flows from FINRA

Table 3 presents Fama-MacBeth regression coefficients of future returns on previous day short flow variables, as in equation (5) and (6). *RSSOPEN (RSSMIDDLE/RSCLOSE)* is a rank variable for shorting flow variable *SSOPEN (SSMIDDLE/SSCLOSE),* computed as shares shorted over the first half hour (middle 4.5 hours/last half hour) over FINRA daily share volume. To compute the rank, we sort all stocks by shorting flow variables into 100 groups, and assign the rank number to the variable. The RSS is computed as rank variables divided by 100. *Ret [-1]* is the return on the previous day using closing bid-ask averages. *Ret [-20, -2]* is the cumulative daily return in the [-20, -2] window. *Ret [-120, -21]* is the cumulative daily return over the [-120, -21] window. *Turnover* is the monthly average of daily turnover measured at the most recent month end. *Volatility* is the monthly volatility of daily return at the most recent month end. *Lsize* is log of market capitalization and *Lbtm* is the log of book to market ratio, both at the most recent quarter end. To account for potential serial correlation in the coefficients, the standard errors of the time-series are adjusted using Newey-West (1987) with eight lags for weekly returns and one lag for daily returns. All regression coefficients are multiplied by 100 for presentation purposes.

Panel A. predict next day return

	I		II		III		IV	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Intercept	0.065	2.03	0.065	2.02	0.044	1.39	0.063	2.02
RSSOPEN	-0.062	-5.20					-0.058	-4.96
RSSMIDDLE			-0.027	-3.20			-0.025	-3.16
RSSCLOSE					0.025	2.42	0.033	3.14
Ret [-1]	-0.440	-2.01	-0.453	-2.06	-0.463	-2.11	-0.435	-1.98
Ret [-20,-2]	-0.149	-2.73	-0.154	-2.81	-0.154	-2.81	-0.150	-2.75
Ret [-120,-21]	0.005	0.22	0.004	0.17	0.006	0.23	0.005	0.21
Turnover	-0.506	-1.67	-0.559	-1.84	-0.588	-1.94	-0.503	-1.67
Volatility	-0.262	-0.64	-0.329	-0.79	-0.314	-0.75	-0.223	-0.55
Lsize	0.001	0.43	0.000	-0.15	-0.001	-0.37	0.001	0.21
Lbm	0.003	0.46	0.004	0.54	0.004	0.54	0.003	0.37
Adj.R2	0.04		0.04		0.04		0.04	

Panel B. Predict daily and weekly returns

	RSS		RSSOPEN		RSSMIDDLE		RSSCLOSE	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Day 1	-0.024	-2.77	-0.062	-5.20	-0.027	-3.20	0.025	2.42
Day 2	-0.018	-2.01	-0.024	-1.93	-0.014	-1.66	-0.015	-1.50
Day 3	-0.001	-0.14	-0.028	-2.37	0.007	0.82	-0.013	-1.35
Day 4	-0.003	-0.33	-0.015	-1.28	-0.003	-0.31	-0.008	-0.82
Day 5	-0.006	-0.72	-0.011	-0.92	-0.005	-0.64	-0.005	-0.50
week 1	-0.010	-1.72	-0.028	-2.98	-0.008	-1.51	-0.003	-0.44
week 2	-0.012	-2.02	-0.015	-1.65	-0.009	-1.66	-0.008	-1.11
week 3	-0.016	-2.78	-0.022	-2.35	-0.012	-2.18	-0.011	-1.52
week 4	-0.019	-3.28	-0.021	-2.22	-0.017	-3.16	-0.014	-1.99
week 5	-0.018	-3.58	-0.020	-2.18	-0.018	-3.81	-0.009	-1.37
week 6	-0.013	-2.42	-0.020	-2.22	-0.015	-2.76	-0.002	-0.27
week 7	-0.016	-2.89	-0.032	-3.55	-0.014	-2.70	-0.008	-1.29
week 8	-0.009	-1.64	-0.026	-2.81	-0.009	-1.76	-0.006	-0.85
week 9	-0.012	-2.10	-0.028	-3.05	-0.015	-2.63	-0.005	-0.64
week 10	-0.012	-2.16	-0.026	-2.69	-0.013	-2.42	-0.005	-0.75
week 11	-0.013	-1.92	-0.025	-2.56	-0.013	-2.06	-0.011	-1.51
week 12	-0.017	-2.53	-0.019	-1.83	-0.017	-2.88	-0.013	-1.85

Table 4. Predicting future returns using different groups' intraday FINRA shorting flows

Table 4 presents Fama-MacBeth regression coefficients of future returns on previous day shorting flow variables, as in equation (5) and (6). Panel A, B, and C present results for shorting flow involving institutions, retail sells and retail buys, respectively. *RSSOPEN (RSSMIDDLE/RSCLOSE)* is a rank variable for shorting flow variable *SSOPEN (SSMIDDLE/SSCLOSE),* computed as shares shorted over the first half hour (middle 4.5 hours/last half hour) over daily FINRA share volume. To compute the rank, we sort all stocks by shorting flow variables into 100 groups, and assign the rank number to the variable. The RSS is computed as rank variables divided by 100. *Ret [-1]* is the return on the previous day using closing bid-ask averages. *Ret [-20, -2]* is the cumulative daily return in the [-20, -2] window. *Ret [-120, -21]* is the cumulative daily return over the [-120, -21] window. *Turnover* is the monthly average of daily turnover measured at the most recent month end. *Volatility* is the monthly volatility of daily return at the most recent month end. *Lsize* is log of market capitalization and *Lbtm* is the log of book to market ratio, both at the most recent quarter end. To account for potential serial correlation in the coefficients, the standard errors of the time-series are adjusted using Newey-West (1987) with eight lags for weekly returns and one lag for daily returns. All regression coefficients are multiplied by 100 for presentation purposes.

Panel A. Using intraday FINRA shorting flow involving *institutions* to predict future return

	RSS		RSSOPEN		RSSMIDDLE		RSSCLOSE	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Day 1	-0.054	-5.21	-0.071	-5.82	-0.071	-5.82	0.007	0.65
Day 2	-0.026	-2.55	-0.025	-1.98	-0.026	-2.61	-0.011	-1.08
Day 3	-0.006	-0.60	-0.030	-2.56	0.001	0.13	-0.010	-0.93
Day 4	-0.009	-0.91	-0.019	-1.57	-0.007	-0.71	-0.009	-0.80
Day 5	-0.012	-1.24	-0.008	-0.66	-0.015	-1.56	-0.002	-0.17
week 1	-0.022	-2.81	-0.030	-3.29	-0.020	-3.00	-0.005	-0.60
week 2	-0.011	-1.52	-0.016	-1.73	-0.011	-1.63	-0.009	-1.03
week 3	-0.017	-2.15	-0.022	-2.42	-0.013	-1.87	-0.013	-1.58
week 4	-0.020	-2.52	-0.022	-2.42	-0.018	-2.62	-0.015	-1.78
week 5	-0.017	-2.41	-0.017	-1.87	-0.019	-3.04	-0.009	-1.25
week 6	-0.010	-1.28	-0.016	-1.78	-0.013	-1.93	-0.001	-0.08
week 7	-0.012	-1.63	-0.028	-3.14	-0.011	-1.67	-0.007	-0.86
week 8	-0.008	-1.01	-0.027	-3.04	-0.008	-1.09	-0.005	-0.58
week 9	-0.008	-0.98	-0.026	-2.97	-0.010	-1.39	-0.003	-0.33
week 10	-0.009	-1.24	-0.027	-2.75	-0.013	-1.87	-0.002	-0.22
week 11	-0.012	-1.52	-0.025	-2.44	-0.014	-1.94	-0.009	-1.01
week 12	-0.017	-2.11	-0.018	-1.71	-0.018	-2.50	-0.013	-1.68

Panel B. Using intraday FINRA shorting flow involving *retail sells* to predict future returns

	RSS		RSSOPEN		RSSMIDDLE		RSSCLOSE	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Day 1	-0.038	-4.28	-0.065	-4.93	-0.034	-3.72	-0.019	-1.84
Day 2	-0.008	-0.88	-0.026	-2.00	-0.003	-0.37	-0.038	-3.85
Day 3	-0.003	-0.44	-0.027	-2.12	-0.001	-0.16	-0.023	-2.24
Day 4	-0.006	-0.64	-0.027	-1.95	-0.004	-0.46	-0.023	-2.40
Day 5	-0.009	-1.12	-0.025	-1.89	-0.010	-1.17	-0.017	-1.82
week 1	-0.013	-2.00	-0.034	-3.08	-0.011	-1.64	-0.024	-3.34
week 2	-0.010	-1.65	-0.026	-2.44	-0.009	-1.39	-0.013	-1.99
week 3	-0.012	-2.11	-0.023	-2.15	-0.011	-1.79	-0.017	-2.71
week 4	-0.013	-2.19	-0.025	-2.30	-0.014	-2.42	-0.013	-2.04
week 5	-0.017	-2.99	-0.024	-2.35	-0.017	-2.93	-0.016	-2.47
week 6	-0.010	-1.55	-0.024	-2.30	-0.008	-1.28	-0.012	-1.79
week 7	-0.020	-3.22	-0.034	-3.38	-0.018	-2.93	-0.023	-3.66
week 8	-0.008	-1.25	-0.023	-2.11	-0.009	-1.40	-0.016	-2.40
week 9	-0.017	-2.97	-0.030	-2.90	-0.017	-2.97	-0.022	-3.40
week 10	-0.008	-1.46	-0.027	-2.61	-0.006	-1.07	-0.025	-3.72
week 11	-0.012	-1.87	-0.031	-2.78	-0.013	-2.04	-0.016	-2.18
week 12	-0.012	-1.90	-0.019	-1.69	-0.014	-2.19	-0.019	-2.73

Panel C. Using intraday FINRA shorting flow involving *retail buys* to predict future returns

	RSS		RSSOPEN		RSSMIDDLE		RSSCLOSE	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Day 1	0.026	2.48	-0.033	-2.24	0.026	2.47	0.039	3.45
Day 2	0.007	0.66	-0.014	-0.92	0.003	0.24	-0.012	-1.10
Day 3	0.011	1.03	-0.023	-1.59	0.012	1.10	-0.006	-0.60
Day 4	0.005	0.44	-0.019	-1.26	0.006	0.56	-0.002	-0.19
Day 5	0.011	1.07	-0.025	-1.77	0.011	1.02	-0.015	-1.39
week 1	0.012	1.26	-0.023	-1.77	0.011	1.20	0.001	0.09
week 2	-0.004	-0.44	-0.016	-1.26	-0.003	-0.36	-0.014	-1.69
week 3	-0.007	-0.72	-0.024	-1.87	-0.006	-0.69	-0.014	-1.67
week 4	-0.012	-1.43	-0.020	-1.57	-0.013	-1.56	-0.018	-2.37
week 5	-0.011	-1.19	-0.027	-2.21	-0.011	-1.24	-0.013	-1.75
week 6	-0.016	-1.83	-0.026	-2.13	-0.016	-1.85	-0.015	-1.85
week 7	-0.023	-2.73	-0.040	-3.28	-0.022	-2.60	-0.026	-3.56
week 8	-0.017	-1.98	-0.036	-2.84	-0.018	-2.20	-0.022	-2.93
week 9	-0.022	-2.58	-0.035	-2.84	-0.022	-2.58	-0.025	-3.28
week 10	-0.024	-2.78	-0.034	-2.71	-0.024	-2.74	-0.031	-4.03
week 11	-0.017	-1.94	-0.030	-2.30	-0.018	-2.05	-0.027	-3.22
week 12	-0.020	-2.17	-0.028	-2.07	-0.021	-2.39	-0.020	-2.51

Table 5. Predicting future same-day returns

Table 5 presents Fama-MacBeth regressions based on shorting prevalence during 30-minute intraday intervals to predict average half-hour same-day returns that start 30 minutes after the measurement of shorting prevalence and end at 15:30. Stated times in the first column are the time at the beginning of the interval, and all shorting variables are for the subset of off-exchange shorting that is likely to involve an institution. Shorting variables are normalized to be between zero and one as in previous specifications. Time-series standard errors are Newey-West (1987) with one lag. Control variables are the same as in prior tables but are not displayed here. All regression coefficients are multiplied by 100 for presentation.

		CBOE		FINRA Institution		FINRA Retail Sell	
shorting hours	Return hours	coef.	tstat	coef.	tstat	coef.	tstat
9:30-10:00	10:30-15:30	-0.018	-15.55	-0.015	-16.63	-0.008	-8.78
10:00-10:30	11:00-15:30	-0.014	-11.63	-0.011	-12.79	-0.007	-7.88
10:30-11:00	11:30-15:30	-0.015	-12.11	-0.011	-12.14	-0.004	-4.94
11:00-11:30	12:00-15:30	-0.013	-10.69	-0.009	-9.24	-0.004	-4.87
11:30-12:00	12:30-15:30	-0.013	-9.79	-0.009	-9.00	-0.004	-4.34
12:00-12:30	13:00-15:30	-0.012	-8.92	-0.005	-5.22	-0.004	-4.40
12:30-13:00	13:30-15:30	-0.011	-7.44	-0.003	-2.92	-0.002	-2.23
13:00-13:30	14:00-15:30	-0.009	-5.00	-0.002	-1.46	-0.003	-1.97
13:30-14:00	14:30-15:30	-0.008	-4.00	-0.002	-1.11	-0.004	-2.90
14:00-14:30	15:00-15:30	-0.008	-2.80	0.001	0.52	-0.001	-0.27

Table 6. Intraday shorting flows and analyst/earnings information events

This table presents the relationship between shorting flow and past and future earnings news. The dependent variables, *RSS (RSSOPEN/RSSMIDDLE/RSSCLOSE)* are rank variables for shorting flow variables, computed as shares shorted over share trading volume. To compute the rank, we sort all stocks by shorting flow variables into 100 groups, and assign the rank number to the variable. The independent variable, pre_bad, is equal to 1 if and only if there is a bad earnings-related news release that is released to the public after the market closes on day t-1, but before the market opens on day t, and zero otherwise. The independent variable, post_bad, is equal to 1 if and only if there is a bad earnings-related news release that is released to the public after the market closes on day t but before the market opens on day t+1, and zero otherwise. We also add "firm" and "analyst" to indicate whether the news is issued by the firm or the news is about analyst recommendation changes. Results from Fama-MacBeth regressions are presented and time-series standard errors are Newey-West (1987) with one lag.

Panel A. CBOE shorting flow and analyst/earnings information events

	RSSOPEN		RSSMIDDLE		RSSCLOSE	
VARIABLES	coef.	tstat	coef.	tstat	coef.	tstat
pre_bad	14.502	71.89	-2.468	-12.36	-10.081	-53.14
pre_badfirm	16.655	46.11	-4.664	-14.37	-13.381	-42.15
pre_badanalyst	12.814	59.41	-1.913	-7.81	-8.947	-38.62
post_bad	0.862	2.86	-1.568	-5.03	-0.786	-2.40
post_badfirm	-1.527	-3.79	-3.401	-8.22	1.147	2.77
post_badanalyst	4.185	11.32	0.268	0.68	-4.312	-10.27

Panel B. FINRA institution shorting flow and analyst/earnings information events

	RSSOPEN		RSSMIDDLE		RSSCLOSE	
VARIABLES	coef.	tstat	coef.	tstat	coef.	tstat
pre_bad	17.509	89.74	-0.671	-3.31	-8.334	-46.65
pre_badfirm	21.013	57.38	-1.985	-5.44	-11.511	-34.73
pre_badanalyst	15.100	65.99	-0.978	-4.16	-7.388	-34.15
post_bad	1.925	5.89	-0.736	-2.29	2.990	8.11
post_badfirm	0.766	1.83	-1.381	-3.57	7.545	16.48
post_badanalyst	3.753	9.20	-0.412	-0.98	-3.404	-7.67

Panel C. FINRA retail sell shorting flow and analyst/earnings information events

	RSSOPEN		RSSMIDDLE		RSSCLOSE	
VARIABLES	coef.	tstat	coef.	tstat	coef.	tstat
pre_bad	17.055	91.97	6.269	34.37	4.774	25.59
pre_badfirm	20.713	55.25	10.773	30.76	7.453	22.42
pre_badanalyst	15.066	67.64	4.197	19.68	3.169	15.20
post_bad	4.167	14.04	3.393	10.13	7.997	22.47
post_badfirm	4.289	10.77	4.668	10.07	12.630	26.28
post_badanalyst	4.684	12.53	2.264	5.72	2.076	5.30

Figure 1. Intraday Shorting Patterns

Panel A of Figure 1 presents the average CBOE short volume during each half hour of the trading day. Panel B presents the average intraday shorting flows on CBOE exchanges on each date in our sample. Panel C presents the average off-exchange short volume, and off-exchange short volume that involves institutions or retail investors, during each half hour of the trading day. Panel D presents the average off-exchange intraday shorting flows on each date in our sample. *Shortall or Short* includes all off-exchange trades with a short seller. *Short_inst* consists of off-exchange short sales that are classified as involving an institution. S*hort_retail* consists of off-exchange short sales that are classified as involving a retail investor. *Open* (*middle*/*close*) indicates the shorting flow during 09:30:00 to 09:59:59, 10:00:00 to 15:29:29, and 15:30:00 to 15:59:59, respectively. All off-exchange short volumes are scaled by the total off-exchange trading volume of the day. All CBOE short volumes are scaled by the total CRSP trading volume of the day.

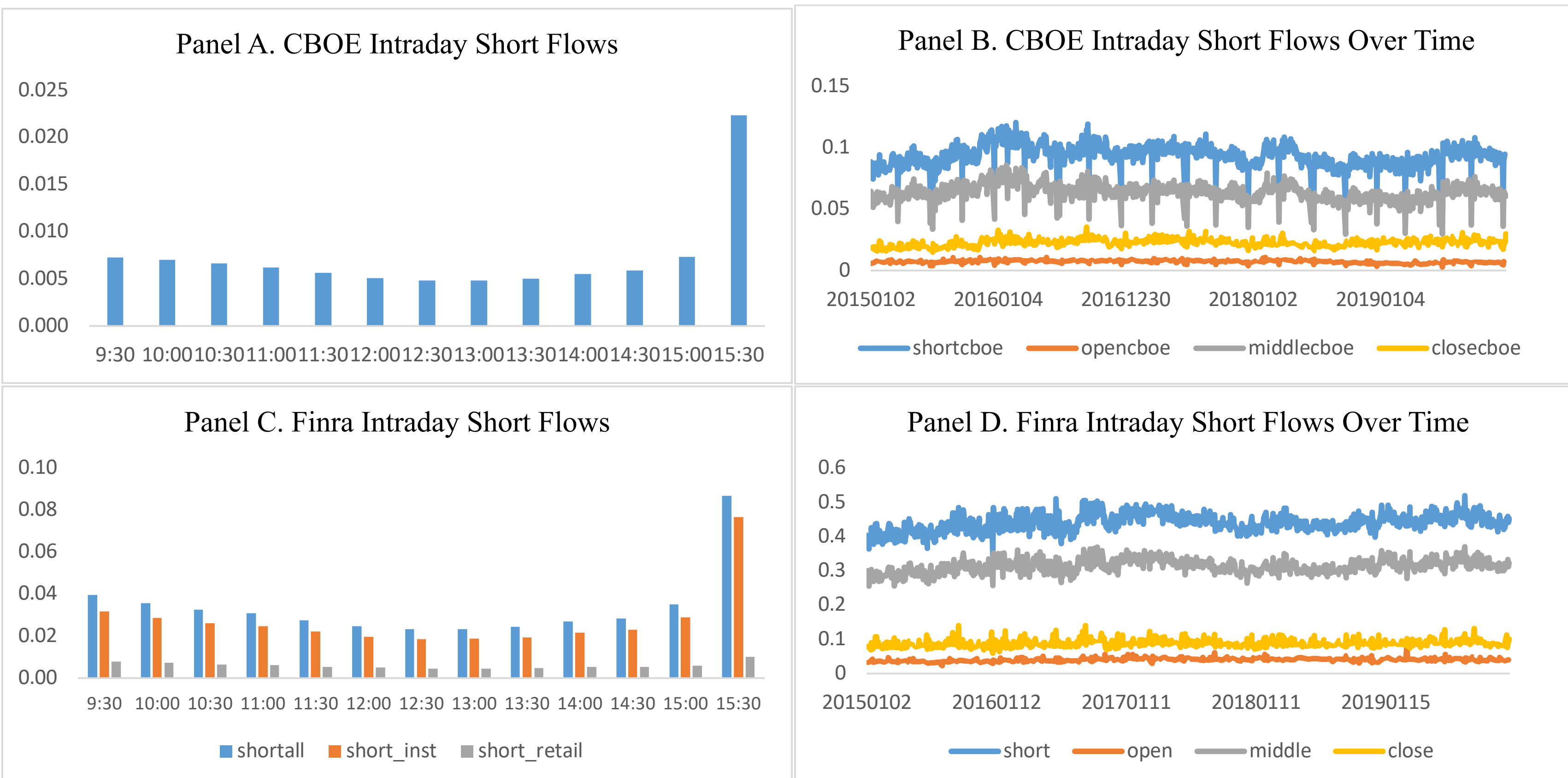
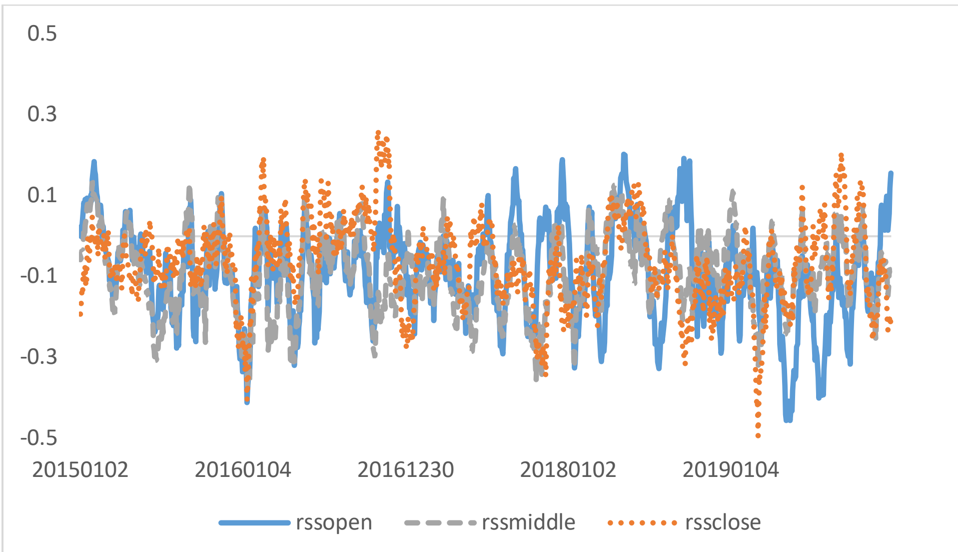


Figure 2. Time-series of Fama-MacBeth regression coefficients

Panel A of Figure 2 presents daily CBOE Fama-Macbeth regression coefficients (multiplied by 100 for presentation) for next-day predictive regressions during our sample period. Panel B and C presents off-exchange Fama-Macbeth regression coefficients. The variable of interest R*SSOPEN*, R*SSMIDDLE* or R*SSCLOSE* indicates the ranked shorting flow from 9:30 am to 9:59 am, 10:00 am to 15:29 pm, or 15:30 pm to 16:00 pm, respectively. To smooth the time-series, here we report the 20-day moving average of the coefficients.

Panel A. CBOE intraday shorts



Panel B. Off-exchange intraday shorts