Closing Auction, Passive Investing, and Stock Prices

Yanbin Wu *

Current draft: November 11, 2019

Abstract

Over the last decade, the volume of market-on-close orders has increased to more than 10% of the entire day's trading volume. This paper investigates this rise and documents four stylized facts: (i) passive investing leads to greater usage of market-on-close orders, consistent with passive fund's motivation for minimizing tracking error; (ii) the price impact from large market-on-close order imbalances is economically large and transitory, leading to short-term price reversal; (iii) a long/short trading strategy exploiting this reversal results in a significant risk-adjusted return of 13.2 basis points per day, consistent with the hypothesis that investors are compensated by providing liquidity to passive funds; and (iv) informed traders also use market-on-close orders, consistent with Admati and Pfleiderer's (1988) prediction that liquidity trades attract informed trades. Overall, the set of findings demonstrates market-on-close order as an important trading channel through which passive investing affects underlying stocks.

JEL Code: G12, G14

Keywords: Closing Auction, Passive Investing, Exchange Traded Funds, Return Predictability

*I am extremely grateful to my committee members Narasimhan Jegadeesh, Jay Shanken, Jeff Busse, Clifton Green, and Ilia Dichev for their invaluable guidance and support. I also thank Tetyana Balyuk, Francisco Barillas, Lawrence Benveniste, Tarun Chordia, Rohan Ganduri, Christoph Herpfer, William Mann, Gonzalo Maturana, and seminar participants at Emory University for helpful comments and suggestions. Address: Goizueta Business School at Emory University, 1300 Clifton Rd, GA 30322. E-mail: yanbin.wu@emory.edu.

1 Introduction

The remarkable growth of passive investing over the last decade has attracted attention among practitioners, policy makers, and academia.[1] There is substantial evidence showing that passive investing has impacted different aspects of the financial markets, such as price discovery (Glosten, Nallareddy, and Zou (2016); Israeli, Lee, and Sridharan (2017)), corporate governance (Schmidt and Fahlenbrach (2017)), and firms' decisions (Appel, Gormley, and Keim (2016); Appel, Gormley, and Keim (2018)). In particular, Ben-David, Franzoni, and Moussawi (2018) show that passive funds, particularly *exchange traded funds* (ETFs), lead to higher nonfundamental volatility of underlying stocks. They argue that the channel through which passive investing affects the constituent stocks is intraday arbitrage activity, since short-term arbitragers would quickly exploit any discrepancy between the net asset value (NAV) of underlying portfolios and the market price of passive funds.[2] Yet, Box, Davis, Evans, and Lynch (2019) find little evidence of this intraday arbitrage trading when they directly examine the order flow between ETFs and underlying portfolios. In this paper, I explore the channel through which passive funds affect underlying stocks.

Trading through closing auction is one possible but overlooked channel for passive investing to affect stock prices. A closing auction is a batch auction that occurs at the end of the trading day for setting the closing price through market-on-close orders. These orders were implemented in the early 2000s by the U.S. exchanges, including NYSE and Nasdaq, to facilitate price discovery and increase liquidity at the end of the day. Unlike normal market orders and limit orders, market-on-close orders are only crossed when the market is closing. Over the past decade, the average trading volume for market-on-close orders has increased dramatically, from 2% of the total daily trading volume in 2010 to nearly 10% in 2018 for S&P 500 stocks (See Figure 1).[3] Put differently, the transaction volume at the closing crosses (which was around \$10 billion in 2018), on average, accounts for more than 10% of the transaction volume in the entire 6.5 trading hours from 09:30 to 16:00.

Given this rise in market-on-close orders and the coincident popularity of passive funds over the past

[1] According to 2018 Investment Company Institute Factbook, the total assets tracking passive indices surpassed \$6.7 trillion in asset under management (AUM) in the U.S. in 2017. Among the passive investing funds in 2017, index mutual funds accounted for \$3.4 trillion, and passive ETFs accounted for \$3.3 trillion.

[2] See also Israeli, Lee, and Sridharan (2017); Broman and Shum (2018); and Da and Shive (2018).

[3] The market-on-close trades for each stock per day can be uniquely identified by their sale condition recorded in the Trade and Quote (TAQ) database.

decade, I address the following research questions: First, do daily passive flows lead to the increased usage of market-on-close orders? Second, what are the implications of this rise in the market-on-close orders for underlying stocks and investors? Third, what are the implications of the clustering of trades through market-on-close orders for market microstructure theories (Admati and Pfleiderer (1988); Admati and Pfleiderer (1991))?

There are reasons to believe that passive fund flows lead to the increased usage of market-on-close orders. First, the most important performance metric for a passive fund is the tracking error relative to the benchmark it follows. A market-on-close order allows a passive fund to automatically minimize tracking errors, because the transaction prices of constituents will be exactly the same closing prices that determine the benchmark level.[4] Furthermore, the increasingly large swing of daily flows into passive funds requires that funds deploy their capital quickly and simultaneously into hundreds of underlying baskets, and market-on-close orders could achieve this easily.

For index mutual funds, managers could directly buy or sell stocks using market-on-close orders in the secondary market. For ETFs, the ETF sponsors interact with the secondary market only through authorized participants (APs) to cater to the investors' demand. To do so, the APs will create or redeem ETF shares in exchange for the underlying basket at the end of the day. Similar to the argument by Pan and Zeng (2019) that bond ETF APs may use the ETF creation/redemption process to manage their inventory risk, equity ETF APs could minimize their inventory risk by using the market-on-close orders for the end-of-day creation and redemption settlement. However, from the perspective of transaction cost, passive funds (as institutional investors) do care about *implementation shortfall* (Anand, Irvine, Puckett, and Venkataraman (2012)). Placing market-on-close orders without knowing the execution price or the price impact exposes the passive funds to additional market risk and transaction costs.

I start the empirical analysis by showing that flows from passive funds are strongly related to increased usage of market-on-close orders at the stock level and daily frequency by utilizing the daily ETF flows[5] and

[4]Besides determining the benchmark level, the closing price is also a widely reference price for other financial assets. For mutual funds, the net asset value (NAV) is calculated at the end of the day, based on the closing prices of the underlying securities. This NAV is then used as the share price for purchase/redemption orders received on that day. For many derivative contracts, the settlement is tied to the closing price of the underlying securities or indices on the day of expiration.

[5]Because daily flows from index mutual funds are not available from major data vendors, my analysis focuses on ETF flows. However, the results documented here could be generalized to index mutual funds.

contemporaneous market-on-close volume from TAQ data. The daily ETF flow for a stock is the aggregate of the weighted daily flow from all ETFs that hold the same stock. Under various specifications, a one standard deviation increase in ETF flows, on average, is associated with a 23% of a standard deviation increase in the market-on-close volume.

To further address endogeneity concerns, I use the post-2007 annual Russell 1000 and Russell 2000 index reconstitutions as exogenous shocks to passive institutional holdings (Coles, Heath, and Ringgenberg (2017)).[6] I conduct this test using two-stage least squares specifications in which I first instrument stock-level ETF flows as exogenous variations around 1000/2000 threshold and then test the effects of instrumented ETF flows on market-on-close volume. Consistent with the OLS results, I find that the exogenous increase in passive flows leads to a significant increase in market-on-close volume. Specifically, the first-stage regression shows that the index assignment from the Russell 2000 to the Russell 1000, on average, leads to a decrease of 17% of a standard deviation in ETF flows. The second-stage regression estimates exceed the OLS estimates in magnitude.[7]

Next, I examine the asset pricing implication of this rise of market-on-close order for underlying stocks. An important feature of the closing auction for NYSE exchange is that at 15:45, all market-on-close orders will be aggregated and disseminated to the public and only market-on-close orders that offset the aggregate order imbalances will be accepted from 15:45. This dissemination of order imbalances is somewhat like the "sunshine" trading modeled by Admati and Pfleiderer (1991), since the passive funds preannounce their trading needs. On the one hand, according to the theory suggested by Admati and Pfleiderer (1991), market-on-close orders might have a minimal price impact at the aggregate level, due to the information dissemination

[6]Specifically, the Russell 1000 and the Russell 2000 indices are constructed each year based on the end-of-May market capitalization ranks. There are only small differences in market value around the threshold and firms cannot precisely control their rankings, such that being assigned to the left or right of the cutoff is as good as random. Before 2007, index assignment followed a simple threshold: Stocks ranked from 1 to 1,000 were assigned to the Russell 1000 index, while stocks ranked from 1,001 to 3,000 were assigned to the Russell 2000 index. Starting in 2007, the Russell implemented a new assignment procedure. After the initial market capitalization breakpoints are determined, existing members are reviewed to determine whether they fall within the the accumulated 5% market cap range around the new breakpoints. This creates an upper band and lower band around the Russell 1000/2000 index threshold (Coles, Heath, and Ringgenberg (2017)). Because of the value-weighted difference for the Russell 1000 and the Russell 2000 (Chang, Hong, and Liskovich (2015)), stocks that move to a different index will result in significant changes in the portfolio weights, which in turn alters passive ownership and passive flows around the threshold.

[7]The IV estimate captures the local average treatment effect (LATE), which only focuses on stocks that are potential index switchers. The effect is larger than the OLS effect since these stocks change status drastically from receiving significant weight (stocks that are at the top of the Russell 2000) when replicating passive indices to being neglected (stocks that are at the bottom of the Russell 1000), while the OLS effect captures the effect for average stock.

and the information content of these sunshine trades. On the other hand, these orders may exhibit a price impact because the aggregated market-on-close orders placed by all passive fund managers are so substantial that providing immediacy and liquidity from the opposite side cannot be guaranteed, much like the price impact of block trades examined by Chan and Lakonishok (1993) and Chan and Lakonishok (1995). However, the price impact of large market-on-close orders is different from that of the block trades placed by other institutional investors, because this price impact is quantified by the price movement during the "advertising" period from 15:45 to 16:00, instead of the price movement after the transaction.

By leveraging proprietary NYSE closing auction data that provide the directions of order imbalances, I investigate the magnitude of the price impact cross-sectionally for both buying and selling market-on-close orders. This analysis suggests that the price impact (defined as the return in the last 15 minutes) on stocks in the cross-sectional top quintile of market-on-close order volume and with a buying direction is significantly positive: 10.1 basis points (bps) more than stocks with low market-on-close order volume. Similarly, the price impact for stocks with high selling market-on-close order volume is significantly negative (10.8 bps) compared to that of stocks with low market-on-close volume.[8] The counterfactual analysis based on the various price impact estimates from the Breen, Hodrick, and Korajczyk (2002) specification, the Glosten and Harris (1988) specification, effective spread, and quoted spread (Korajczyk and Sadka (2004)) demonstrates that the price impact from the market-on-close orders is relatively large.

To further confirm that the price movement for the last 15 minutes is due to the dissemination of closing auction information, I conduct two falsification tests: one using the 15-minute return from 15:15 to 15:30 and the other using the 15-minute return from 15:30 to 15:45 as dependent variables. The results show that the price impact during these periods is marginally significant but economically smaller (0.5 bps) than that during the last 15 minutes. The statistical significance is consistent with two economic interpretations. First, it is possible that information about the order imbalances is leaked before 15:45, consistent with the practice at NYSE that floor traders are notified of large market-on-close imbalanced stocks around 15:00. Second, the market-on-close trades capture only part of the overall anticipated passive flows, and the trading activity

[8] The subperiod analysis shows that this price impact is higher in the first half of the sample, suggesting that, over time, the market incorporates the information from the closing auction efficiently. As an additional robustness check, I also use the market-on-close trading volume, retrieved from TAQ data, and I show that stocks with high market-on-close trading volume move more during the closing auction period.

during the day through market or limit orders may move the stock price in the same direction.

Next, I investigate whether this price impact is short lived, since the ETF flows are mostly demand-driven and unrelated to fundamental shocks for a given stock. In addition, the fact that large order imbalances require immediacy at the end of the day suggest that liquidity providers should be compensated for providing liquidity to passive funds (See Nagel (2012) and Duffie (2010)). I use Fama and MacBeth (1973) regressions at a daily frequency to analyze the cross-sectional relation between return reversals and market-on-close trading volume. The cross-sectional results show that stocks with high market-on-close trading volume experience significant reversals on the following day. Results that decompose close-to-close returns into overnight returns and intraday returns further suggest that this reversal is not only persistent during the overnight period but also continuous during the following open-to-close trading period. The results are robust across different stock size quintiles and different exchanges. Furthermore, the results are not subsumed by order imbalance; Chordia and Subrahmanyam (2004) find that daily order imbalance predicts next day returns. This suggests that the daily reversal pattern observed from market-on-close trades is not just a proxy for the entire day trading activities.

To rule out the possibility that my results are driven by market microstructure bias,[9] such as a bid-ask spread bounce, I use the midpoint of the quoted bid and ask prices to calculate the daily return. My results still hold. Using NYSE order imbalance data, I find that the price reversals for high market-on-close trading volume are symmetric for buying orders and selling orders. This further confirms that the price impact is temporary and unrelated to fundamental information. The findings provide additional evidence on how the price dynamics are influenced by institutional trades at the intraday and overnight level, consistent with the recent studies (Heston, Korajczyk, and Sadka (2010); Bogousslavsky (2018); Gao, Han, Li, and Zhou (2018); and Lou, Polk, and Skouras (2018)).[10]

Given such a strong short-term reversal, a long/short trading strategy yields significant profits both economically and statistically. During the whole sample period, a daily rebalanced portfolio that (a) buys stocks

[9]Nonsynchronous trading bias (Scholes and Williams (1977)) cannot be the cause of return reversals here since the last trade for a given day is the market-on-close trade, which is the same across all stocks.

[10]My findings, however, differ from these studies in several ways. First, the analysis focuses on the reversals of individual stocks instead of return continuation at the portfolio level. Heston, Korajczyk, and Sadka (2010) examine the return continuation at half-hour intervals that are exact multiples of a trading day and Gao, Han, Li, and Zhou (2018) focus on the market intraday momentum, in which the first half-hour market return predicts the last half-hour market return. Second, the interpretation of my findings is consistent with the price pressure from passive demands. Lou, Polk, and Skouras (2018) find that the overnight and intraday patterns at a monthly frequency are consistent with the tug-of-war/clientèle effect, while Bogousslavsky (2018) examines intraday return patterns across different anomalies.

with high selling closing order imbalances and (b) sells stocks with high buying closing imbalances results in a risk-adjusted return of 13.2 bps per day, or a 33.26% annualized return with a Sharpe ratio of 3.13. Furthermore, both the long and short legs of the trading strategy yield a similar magnitude of abnormal returns, consistent with the cross-sectional regression findings.[11]

Taken together, this reversal pattern is consistent with the empirical finding that stocks with higher ETF ownership exhibit a higher volatility and more deviation from a random walk (Coles, Heath, and Ringgenberg (2017) and Ben-David, Franzoni, and Moussawi (2018)). To illustrate the contribution of market-on-close orders on the volatility, I repeat the analysis using S&P 500 stocks by regressing monthly volatility calculated from daily return on ETF ownership and the results are consistent with Ben-David, Franzoni, and Moussawi (2018) that higher ETF ownership leads to higher volatility. Then, instead of using close-to-close return to calculate the daily return and monthly volatility, I use open to 15:45 daily return to purge the impact of the closing auction. The results suggest that impact of ETF ownership on the volatility of stocks shrinks by 60% compared to the original coefficient estimates, shedding light on the importance of the market-on-close orders for stocks with different ETF ownership.

Last, besides passive funds, informed traders might use market-on-close orders to take advantage of market-depth and camouflage their trading intention. Early market microstructure models (Admati and Pfleiderer (1988); Admati and Pfleiderer (1991)) predict that uninformed traders who have discretion over the timing of their trades will optimally cluster their trades to minimize their price impact. These early models also predict that informed traders will herd with uninformed traders to camouflage their trades. Nevertheless, herding is difficult to be observed. The growth of market-on-close orders allows me to empirically test these predictions directly.

Specifically, I conduct two analyses. My first analysis examines the daily return reversal for a subset of stocks with a high probability of informed market-on-close trades, based on the contradictory trading directions from both ETF flows and market-on-close imbalance sides. A weaker and insignificant return reversal among these stocks provides supportive evidence that the informed trades might use market-on-close orders because the price impact is permanent due to private information (Nagel (2012)). The second analysis

[11] The results are robust for different trading strategies and for different subsample periods. However, the risk-adjusted return for the second half period is lower than the risk-adjusted return for the first half period, suggesting that investors may become aware of this pattern and exploit it.

focuses on the trading pattern around earnings announcements. I find abnormal market-on-close trading volume that does not derive from ETF flows. I also find that informed investors might use market-on-close orders as a way of exiting positions, consistent with the " buy the rumor and sell the news" pattern documented in Kaniel, Liu, Saar, and Titman (2012) and Kadan, Michaely, and Moulton (2018). The overall evidence here suggests that informed traders could also use market-on-close orders as an alternative venue for trading.

This paper contributes to the literature in several ways. First, a growing literature explores the impact of passive investing on asset prices. In particular, recent studies of ETF ownership highlight the unintended consequences for underlying securities, such as the increase of nonfundamental volatility (Ben-David, Franzoni, and Moussawi (2018)), increased co-movement in returns (Da and Shive (2018)), increased commonality in liquidity (Agarwal, Hanouna, Moussawi, and Stahel (2019)), more deviation from the random walk (Coles, Heath, and Ringgenberg (2017); Ben-David, Franzoni, and Moussawi (2018)), and reduction in information efficiency (Israeli, Lee, and Sridharan (2017)). Brown, Davies, and Ringgenberg (2018) show that ETF flows at a monthly frequency can predict future stock returns, especially for leveraged ETFs.

However, most of these studies provide indirect evidence that how ETFs impact underlying assets is through creation and redemption throughout the day or through arbitrage activities exploited by hedge funds and high-frequency traders.[12] Box, Davis, Evans, and Lynch (2019) examine the intraday arbitrage opportunities between ETFs and their underlying portfolios and find little evidence that mispricing events alter the direction of constituent order flows. To my knowledge, this paper is the first to show that, aside from arbitrage activities, the price impact of large market-on-close orders is one important but overlooked mechanism through which ETFs influence underlying assets. I demonstrate that this price impact is economically and statistically significant and gives rise to a new arbitrage opportunity that allows for the exploitation of the short-term reversal. This short-term reversal is also consistent with the hypothesis that investors are compensated by providing the liquidity to passive investors (Nagel (2012) and Duffie (2010)).

Second, this paper contributes to the burgeoning literature on the price impact of institutional flow-driven trades. For instance, both Coval and Stafford (2007) and Frazzini and Lamont (2008) analyze the long-term return reversal pattern subsequent to mutual fund flow-induced trading while Lou (2012) focuses on

[12] As Ben-David, Franzoni, and Moussawi (2018) state, "The evidence on the role of arbitrage trades is admittedly indirect. Further research should delve deeper into the channels linking ETF ownership to stock volatility."

the long-term return continuation pattern due to price pressure. Ben-Rephael, Kandel, and Wohl (2011) use data from aggregate daily mutual fund flows in Israel, and find that these flows create a temporary price pressure that is reversed within 10 trading days. Besides return predictability, studies (e.g., Greenwood and Thesmar (2011); Antón and Polk (2014); Akbas, Armstrong, Sorescu, and Subrahmanyam (2015)) examine other consequences of institutional flow-driven trades, such as co-movement and mispricing. More recently, Etula, Rinne, Suominen, and Vaittinen (2019) show that month-end price pressure originated from monthly payment cycle leads to price reversal at the beginning of month. Different from previous papers that focus on flows from actively managed mutual funds or hedge funds at the quarterly or monthly frequency, this paper analyzes the impact of flows that originate from passive investing at the daily frequency and shows that the impact of "dumb money" from passive investing is economically large and plays an increasingly important role in the asset price dynamics observed in recent years.

Third, my paper complements research on price manipulation at close and the effect of closing auction on end-of-day price discovery and liquidity. Price manipulation at close has been examined by previous studies such as Hillion and Suominen (2004), Comerton-Forde and Putniņš (2011), and Ben-David, Franzoni, Landier, and Moussawi (2013). The adoption of closing auction is an attempt to alleviate such manipulation. Yet, relatively few studies examine the closing auction after its adoption in the U.S. stock market.[13] Mayhew, McCormick, and Spatt (2009) utilize floor trading data from the first quarter of 2005 to study the role of the specialists during the closing auction after its adoption by the NYSE. My paper demonstrates the increasing importance of closing auctions on financial markets, especially due to passive investing. Additionally, my findings shed light on the trade-off between minimizing tracking errors and better execution even for passive funds and have policy implications for the design of closing auctions.

The remainder of this paper is organized as follows. Section 2 reviews the institutional background of closing auctions and passive investing. Section 3 describes the data and the construction of the variables. Section 4 provides empirical analysis. Section 5 concludes.

[13] Before its adoption in the U.S stock market, market-on-close orders are used for the settlement of option expirations. For instance, Cushing and Madhavan (2000) investigate the stock returns following market-on-close order imbalance, especially on the days when options expire.

2 Institutional Background

2.1 Closing auction and close order types

The closing auction is the last event of the trading day across all major exchanges and is designed to determine the closing price for each stock. The closing price is crucial, as it is the most widely published reference price for all equity-linked products, such as mutual funds, exchange trade products (ETPs). During the trading day, a stock can be traded on any exchange. But at the close, it reverts to the exchange where it is listed. This means that stocks listed on the NYSE will have dramatic closing auction volume on the NYSE at closing, and the process is similar for Nasdaq-listed stocks.[14]

Three order types are allowed in this process: *market-on-close* (MOC) orders, *limit-on-close* (LOC) orders and *imbalance-only* (IO) orders (also referred to as *closing offset* (CO) orders in the Nasdaq market). A MOC order represents interest that must trade in the closing auction, irrespective of the price. Like a usual limit order, an LOC order is an order to buy or sell a stock if the closing price is at or better than the bidder's limit price. Finally, an IO order is a limit-order type that offsets daily order imbalances at the market close. At the NYSE, MOC and LOC orders can be entered from 7:00 until 15:45 and cannot be canceled except in the case of an error.[15] Starting from 15:45, the NYSE electronically publishes a rundown of open interest on each stock, including information such as imbalance quantity, paired quantity, imbalance side, and indicated match prices.[16] At 16:00, auctions are run for all orders, and closing prices are published shortly thereafter.

[Figure 1 here]

To illustrate the growing importance of closing auction trades, I present descriptive statistics for the S&P 500 and all stocks traded on the NYSE in Figure 1. For S&P 500 stocks, the daily average fraction of a stock's closing auction trades over total shares traded has risen almost threefold, from 3.5% in 2010 to 10%

[14] https://www.wsj.com/articles/goldman-cashes-in-on-passive-investing-boom-with-big-4-p-m-trade-1535295600.

[15] Different exchanges have different rules about their cutoff times: 15:50 (Nasdaq) or 15:59 (NYSE Arca). There have been several rules changes that relate to the closing auction process, with the most recent change on March 1, 2010. With this change, the time of dissemination changed then from 15:40 to 15:45, and exchange systems automatically and electronically tracked MOC and LOC interest, as opposed to the designated market maker (DMM). After this change, the frequency of order imbalance dissemination reduced from 15 seconds to 5 seconds.

[16] See NYSE rule 123C for details about the closing auction procedure: http://wallstreet.cch.com/nysetools/PlatformViewer.asp?SelectedNode=chp_1_3&manual=/nyse/rules/nyse-rules/.

in 2018. This represents, on average, 203 million shares and $14.9 billion traded daily through the closing auction (based on the first six months of 2018). The statistics for all stocks traded on the NYSE show a similar picture.

2.2 Passive investing: Index mutual funds and ETFs

Passive investing is designed to track the performance of a market index. To do this, the fund manager purchases all or a representative sample of the securities in the index through the form of index mutual funds and ETFs.[17] While index mutual funds were introduced in the 1970s, the first index ETFs were offered in 1993 as an SPDR investment vehicle, tracking S&P 500. Passive investing has grown remarkably over recent years in terms of total net assets under management as well as the number of ETFs being offered. According to the 2018 Investment Company Fact Book, the total assets of these funds (both index mutual funds and ETFs) reached $6.7 trillion by the end of 2017, and these funds comprised 35% of the total net assets in long-term funds, compared to 15% in 2007. In particular, the total net assets for index mutual funds accounted for $3.4 trillion, up from $619 billion in 2005, while index ETFs accounted for the remaining $3.3 trillion.

Even though both index mutual funds and ETFs are designed for passive investing, there are two major differences between these two types of funds. First, ETFs, like stocks, are listed on an exchange, and investors can buy and sell them throughout the day at market-determined prices. However, index mutual funds can be traded only once per day at the closing price.

Another unique feature of ETFs is the creation–redemption mechanism that operates between the fund and authorized participants (APs). An ETF does not interact directly with the secondary market. Instead, APs play a key role in the primary market for ETF shares because APs are the only investors allowed to interact directly with the fund. An AP can create (redeem) ETFs shares by transferring to (receiving from) the ETF the underlying securities. Creation and redemption follows a predefined procedure specified in the authorized participant agreement or the authorized participant handbook. This procedure determines how large predefined blocks (often 50,000 ETF shares or creation units) are exchanged at designated times (usually

[17] Most ETFs are standalone and differ from mutual funds in terms of structure. Vanguard ETFs are not standalone ETFs, and they are merely another share class of the Vanguard open-end index mutual fund https://personal.vanguard.com/pdf/icrsc.pdf.

end-of-day) and at designated prices (typically closing prices) between APs and funds. [18]

Creation and redemption at the end of the trading day can be viewed as two components that originate from the two different roles of the APs. First, creations and redemptions are partly the results of arbitrage activity. Specifically, if an ETF trades at a discount relative to its basket of stocks, APs have an incentive to buy ETF units in the market and short-sell the underlying basket during intraday trading. Near the end of the trading day, the APs then deliver the ETF shares they bought to the ETF sponsor, and the APs receive the underlying basket to cover the short position. In the case of a premium, the APs short-sell the ETF unit and buy the underlying basket during intraday trading. At the end of the trading day, APs deliver the underlying basket, and receive ETF shares from the ETF sponsor to maintain zero inventory risk. Figure A.2 illustrates examples of these two scenarios. The profit of arbitrage activity is locked in during the intraday trading session, since the end-of-day creation or redemption is based on the closing NAV of the ETF and on the closing prices of underlying securities. This means there is no arbitrage for this end-of-day process.

At the end of the trading day, the value of the creation (or redemption) basket equals the value of the creation unit based on the ETF's NAV at the end of the day on which the transaction was initiated. Under both scenarios, APs might participate in the closing auction for the underlying securities to minimize mismatching risk or inventory risk. Therefore, similar to index mutual funds, inflows and outflows captured by the ETF shares outstanding are also closely related to the closing auction.

[Figure 1 here]

Another part of creation and redemption derives from an investor's demand for exposure; that is, an AP's activity can be viewed as a technology for adjusting the shares outstanding of the ETF in response to the demand for the exposure provided. For instance, if an institutional investor (or an aggregation of retail investors) seeks a large block of a particular ETF's shares, the investor may turn to an AP to facilitate a creation. The buyer delivers either cash or securities to the AP, who in turn delivers the basket of securities to the ETF sponsor. The ETF sponsor then issues ETF shares to the AP (i.e., a creation) to give to the buyer.

[18] See appendix A.1 for one example of an Authorized Participants agreement. There could be a cash adjustment if an AP creates or redeems in cash.

3 Data

For ETF and daily flow data, I first use the Bloomberg terminal to identify all ETFs that are traded on the major U.S. exchanges. Following previous studies on ETFs, leveraged and inverse-leveraged ETFs are excluded.[19] The information on ETFs from Bloomberg contains the total shares outstanding at the end of the day, the ticker symbol, CUSIP, and the calculated daily flows in dollars. I then match the ETFs with the Center for Research in Security Prices (CRSP) database for all securities that have a share code of 73 by symbol and by CUSIP to obtain daily trading information. My final ETF sample contains 365 distinct ETFs for 2010 to 2018. For ETF holdings data, I use the Thomson-Reuters Mutual Fund Ownership database by following procedures. First, I identify the *fundno* of the ETFs by matching the historical ticker-date with the *s12type8* dataset in Thomson-Reuters. Then, I retrieve the last available detailed quarterly holding information using table *s12type3* dataset.

At the end of each day, for each stock i in the CRSP stock file universe, I construct the daily ETF flow as

$$ETF\_flow_{i,t} = \frac{\sum_{j=1}^{J} w_{i,j,t} * abs(\$Flows_{j,t})}{\$Volumes_{i,t}} \tag{1}$$

where j is the set of ETFs that hold stock i; $w_{i,j,t}$ is computed as the most recently available fund portfolio percentage weight of ETF j in stock i (i.e., the dollar ownership by the fund portfolio in stock, divided by total equity assets in the fund portfolio, based on the most recent quarterly report); $abs(\$Flows)_{j,t}$ is the absolute value for the daily dollar flow for each fund j at time t, where the daily dollar flow is measured based on the change of total shares outstanding at day-end from Bloomberg for each ETF; $\$Volumes_{i,t}$ is the dollar trading volume over the same period t. The motivation to use the sum of the absolute value of flows is based on the premise that different funds cannot trade directly to cancel out the opposite trading need.

For market-on-close data, I use TAQ second (pre-2014) and millisecond (post-2014) data to identify each stock's market-on-close volume based on its primary listing exchange for each day with the sale condition of "6" (closing print) mainly for NYSE-listed securities. For Nasdaq-listed securities, trades from the closing auction can also be identified by the sale condition of "M" (market center close price). However, in most

[19] AUM from leveraged and inverse-leveraged ETFs account for 2% of the whole sector according to BlackRock (December 2014).

cases, these trades have identical price and volume information with the sale conditions of "6" and will be discarded to avoid double accounting.

My sample period for closing auction volume spans March 2010 to June 2018 because the NYSE (also NYSE Arca) did not report the closing auction until July 2008 and because the regulation on the dissemination of the closing order imbalance changed in July 2008 and in March 2010 (as described in section 2). I cross-match with CRSP, and I retain only common stocks (*shrcd* in 10 or 11), which means that I exclude securities such as warrants, preferred shares, American Depositary Receipts, closed-end funds, and REITs.[20] I also exclude stocks with zero daily trading volume or zero market-on-close volume. Stocks with prices of less than $5 are also excluded to ensure market micro-structure issues do not affect the analysis. I define the *market-on-close volume* measure as the total shares of market-on-close volume divided by the total trading volumes on each date. For control variable order imbalance, I use the modified Lee and Ready (1991) algorithm to identify buy or sell orders following Blume, Easley, and O'Hara (1994) and Holden and Jacobsen (2014). For the portion of my analysis that uses intraday prices or returns, I use the last available valid trade price from the TAQ data as the end-of-period stock price, and these returns are winsorized at the 1st and 99th percentile to avoid extreme outliers. Additionally, I use the NYSE close imbalance feeds data from July 2008 to June 2018 to identify the direction of the market-on-close order imbalances. The closing auction imbalance feeds data provide information for each stock on imbalance quantity, paired quantity, imbalance side, and indicated match prices every five seconds from 15:45 until 16:00.

For the analysis using the Russell reconstitution, I first collect the constituents of the Russell indices (RTY index, RIY index, and RAY index) from the last trading day in May of each year before reconstitution and the last trading day in June each year after reconstitution. Because my sample mainly starts from 2010, after Russell Inc. implemented a new banding policy around the 1000 cutoff to mitigate index turnover,[21] I follow the procedure from Coles, Heath, and Ringgenberg (2017) to construct new upper and lower bounds in order to construct new thresholds for identifying potential index switchers.

Under the new policy, after the initial market capitalization breakpoints are determined, existing members

[20] As suggested by Boehmer, Jones, and Zhang (2008), ticker symbols are sometimes reused, and ticker symbols in CRSP do not always match the ticker symbols in the NYSE data, especially for firms with multiple share classes. Tickers and CUSIPs are used to ensure accurate matching.

[21] Russell reconstitution guide from Russell Inc.:https://www.ftse.com/products/downloads/Russell-US-indexes.pdf.

are reviewed to determine whether they fall within an accumulative 5% of the market cap range around the new breakpoints. This creates an upper band and an lower band around the Russell 1000/2000 index threshold. The width of each band equals 2.5% of the total market cap of the Russell 3000E (which consists of 4,000 stocks or the whole stock universe if the total number of eligible stocks is less than 4,000). Stocks within the bands do not switch their index assignment from previous year.

To construct the post-banding sample, I perform the following procedure. First, all candidate stocks from the CRSP are ranked based on their end-of-May market cap, and the threshold for rank 1,000 is then imputed. Then, I calculate the cumulative market cap from stocks ranked 1 to 1000 (CMC_{1000}) and the cumulative market cap from stocks ranked 1 to 4000 (CMC_{all}). The upper band is then determined based on the largest n, such that the cumulative market cap from stocks ranked 1 to n (CMC_n) is smaller than $CMC_{1000} - 0.025 \times CMC_{all}$. The lower band is determined in a similar way.

Finally, information regarding earnings announcements and analyst forecasts is extracted from the Institutional Brokers' Estimate System (I/B/E/S).

[Table I here]

Table I presents summary statistics for the whole sample. The sample consists of 6,663,021 stock-day observations over the whole period from July 1, 2008, to June 30, 2018. Over this period, the average volume for the market-on-close orders across all stocks and all days is about 62,000 shares, which on average accounts for 6.4% of total daily volume. There is $1.57 million in absolute flows per stock-day (i.e., 14.1% of daily volume). Variable definitions are available in the Appendix A.I.

4 Results

4.1 ETF flows and market-on-close volume

My first hypothesis is to test whether ETFs influence the underlying assets through the closing auction at the end of the trading day. In this section, I begin the analysis by first showing that ETF flows are correlated with market-on-close volume at the individual stock level. APs are motivated to create or redeem either because they are actively involved in arbitrage activities or because they are trasmitting demand from the tracking indices of institutional or retail investors. In both cases, since the orders are based on the closing prices of the underlying securities and are cleared only once per day at 16:00, APs have the incentive to place a market-on-close order. To formally test this hypothesis, I estimate a panel regression of daily market-on-close volume on ETF flows as follows:

$$MOC\ Volume_{i,t} = \alpha + \beta ETF\ flow_{i,t} + \gamma X_{i,t} + Fixed\ Effect + \varepsilon_{i,t} \tag{2}$$

where $MOC\ Volume_{i,t}$ is the market-on-close trading volume for stock i from the TAQ data, scaled by total trading volume in day t; the key variable of interest is $ETF\ flow_{i,t}$, the aggregated absolute flows at time t for stock i from ETFs holding the stock, as defined in Equation (1); and $X_{i,t}$ denotes a vector of time-varying controls for stock characteristics.

To address potential omitted variables issue, I include stock and day fixed effects to account for both time-invariant stock effects and time-varying effects. I also control for a number of stock characteristics that might be related to closing auction trading. The variable *log(market cap)* is the natural log of market capitalization of the stock. *Amihud ratio* measures illiquidity and is computed as the absolute daily return divided by the total daily volume in millions of dollars, following Amihud (2002). $Return_t$ and $Return_{t-1}$ represent stock returns over the same day and over the previous day, respectively. Finally, I include *Turnover* and *Order imbalance* as explanatory variables to capture the aggregate-level trading, since APs could use alternative trading strategies, such as submitting a market order during the day or spreading orders throughout the normal trading session. *Order imbalance* is calculated using the modified Lee and Ready (1991) algorithm, following Blume, Easley, and O'Hara (1994) and Holden and Jacobsen (2014). It is defined as the total dollar purchase minus total dollar sell, normalized by the total dollar trading volume. *Turnover* is the total daily trading volume divided by market capitalization. To ease interpretation, for each day, I standardize both the dependent variable and the main explanatory variables to have a mean of zero and a unit of standard deviation. Standard errors are clustered at the stock and daily levels to account for error correlation within stocks and days as suggested in Petersen (2009).

The results are presented in Table II. Column 1 reports the estimate of a simple benchmark regression of market-on-close trading volume on stock-level ETF flows after including both stock and day fixed effects. Consistent with the hypothesis, ETF flows are positively associated with market-on-close trading volume. In particular, a one standard deviation increase in ETF flow is associated with a 27% of one standard deviation increase in market-on-close trading volume. The coefficient is statistically significant at the 1% level. Columns 2 to 4 add stock-characteristics controls and various fixed effects. Column 2 adds only firm fixed effects in addition to all stock characteristics. The estimated coefficient shrinks to 25.5% but retains its statistical significance. The estimate of β from Column 3 is still significant after controlling for the day fixed effect. The larger coefficient of 30.2% suggests that certain time-invariant characteristics are important in determining the market-on-close trading volume. The coefficient in Column 4 shrinks to 25.4%, but remains significant. The estimated impact is economically sizable: A one-standard-deviation increase in ETF flows leads to a increase of 2.01% (0.255 $\times$ 7.9%) of market-on-close trading volume, which is a 31.4% increase relative to the mean market-on-close volume of 6.4%.

[Table II here]

The independent variable *ETF flow*$_{i,t}$ used in Panel A is measured by the sum of absolute stock-level flows.[22] The assumption behind using this measure is that APs, who are responsible for creation and redemption from various ETFs, will not cross-trade stocks that are commonly held by multiple ETFs.[23] For example, APs, who are responsible for both the creation of the iShares S&P 500 ETF and the redemption of the State Street's S&P 500 ETF will submit separate orders through a secondary market or through market makers. However, this assumption might not always be true because APs could submit netted orders. In addition, the clearing process for creation or redemption requires that deposit securities must be delivered to a Depository Trust Company (DTC) through the National Securities Clearing Corporation ("NSCC"). Therefore, the total trading for each stock could be netted at a clearing center. To take cross-trading or netting into consideration, an alternative measure of ETF flows, which nets the flows at the stock level, is defined as

$$ETF\_flow_{i,t} = \frac{abs(\sum_{j=1}^{J} w_{i,j,t} * \$Flows_{j,t})}{\$Volumes_{i,t}} \tag{3}$$

[22] Ben-David, Franzoni, and Moussawi (2018) argues that this component of flows as the sum of absolute stock-level flows captures part of non-fundamental explanations since investors might reshuffle money across the ETFs holding a given stock.

[23] For mutual funds, within the same fund family, different funds may cross-trade or do inter-fund lending.

Panel B of II presents the results that repeat the same analysis as in Panel A but uses the alternative definition of ETF flows described above. Across all four specifications, the coefficients of the estimate are still significant at the 1% level, suggesting market-on-close trading volume is driven mainly by net inflows of money into the ETFs. For instance, the specification in Column 4 that adds all stock controls and both stock and day fixed effects shows that a one standard deviation increase in ETF flow is associated with a 22.5% of a standard deviation increase in market-on-close trading volume. The small difference between the magnitude of the coefficient in Panel A and the coefficient in Panel B suggests that this new ETF flow measure is noisier and that netting or cross-trading might be a less important factor when determining the market-on-close trading volume.

Table II, Panel C shows that the effect has increased over time and is robust across different exchanges. The fact that the effect intensifies is consistent with the significant increase in passive funds over the last 10 years. For the sake of brevity, in each of the subsample analyses, I only show the coefficient on $\textit{ETF flow}_{i,t}$. When using all controls, the coefficient on $\textit{ETF flow}_{i,t}$ changes from 0.248 for the sample period of July 2008 to June 2014 to 0.250 for the sample period of July 2014 to June 2018. The analysis using stocks from different exchanges shows that this effect is pronounced across all stocks. The effect is stronger for Nasdaq stocks (0.276) compared to NYSE stocks (0.185). To summarize, the main conclusion from this analysis is that ETF flows predict statistically and economically significant variation in market-on-close trading volume.

4.2 A quasi-natural experiment: Russell reconstitution

The OLS results discussed in the previous section suggest that an increase in market-on-close trading volume is strongly correlated with passive fund flows. Even though I include fixed effects and control for a number of time-varying stock characteristics, the causal inference is still questionable if certain time-varying variables that co-determine market-on-close trading volume and ETF flows are omitted. In this section, I conduct analyses using the annual Russell index reconstitution experiment as a quasi-natural experiment, which allows me to exploit mechanical changes in passive flows.[24]

The Russell 1000 and 2000 indices are constructed based on the end-of-May market capitalization ranks

[24] See Appel, Gormley, and Keim (2016); Chang, Hong, and Liskovich (2015); Ben-David, Franzoni, and Moussawi (2018); and Coles, Heath, and Ringgenberg (2017).

each year. Russell Inc. reconstitutes the indices on the last Friday in June of each every year based on end-of-May stock capitalization; the composition remains constant for the rest of the year. Before 2007, index assignment followed a simple threshold: Stocks ranked from 1 to 1,000 were assigned to the Russell 1000, while stocks ranked from 1,001 to 3,000 were assigned to the Russell 2000. Such a simple threshold creates a quasi-natural experiment to exploit mechanical changes in passive index.

Starting in 2007, Russell implemented a new assignment banding procedure that replaced the original rule. After the initial market capitalization breakpoints are determined, existing members are reviewed to determine whether they fall within in accumulative 5% market cap range around the new breakpoints. This creates an upper band and a lower band around the Russell 1000/Russell 2000 index threshold. The width of each band equals 2.5% of the total May cap of the Russell 3000E (which is 4000 stocks or the whole stock universe if the total number of eligible stocks is less than 4000). Stocks within the bands do not switch their index assignment from the previous year. After the 2007 banding, an existing member will not be replaced if its market cap is within 5% of the cutoff range. This significantly decreases the addition and deletion of stocks due to Russell reconstitution.

However, even under the new regime, firms still cannot control their rankings precisely, thus being assigned to the left or right of the cutoff (i.e., either upper band or lower band) is quasi-random. This upper band/lower band division actually creates the new thresholds for index switching, and the assignment is essentially quasi-random around the band threshold (Coles, Heath, and Ringgenberg (2017)).

Because the Russell indices are value-weighted, this random assignment leads to significant differences in the portfolio weights and further differences in passive ownership around the threshold. The weights of the top stocks in the Russell 2000 are about 10 times larger than those of the bottom stocks in the Russell 1000 (Chang, Hong, and Liskovich (2015)). Consequently, a significantly larger amount of passive money tracks the top Russell 2000 stocks than the bottom Russell 1000 stocks. Stocks that were at the top of the Russell 2000 but then switched to the Russell 1000 would receive less institutional ownership and would expect lower passive fund flows.

I conduct a two-stage least squares estimation. For my main analysis, the sample is composed of stocks that in May, before index reconstitution, are potential switchers in the Russell 2000 around the upper band threshold. The sample composition remains constant for all the months between June, the first month after index reconstitution, and April of the following year. In the first stage, I conduct a regression of ETF flows on an indicator variable of whether the stock switches index membership in June:

$$ETF\ flows_{i,t} = \alpha + \beta\ \mathbf{1}(Switcher_{i,t}) + \gamma X_{i,t} + Fixed\ Effect + \varepsilon_{i,t} \tag{4}$$

In the second stage, I regress market-on-close trading volume on the fitted value of ETF flows from the first stage. The regression is

$$MOC\ Volume_{i,t} = \alpha + \beta \widehat{ETF\ flow}_{i,t} + \gamma X_{i,t} + Fixed\ Effect + \varepsilon_{i,t} \tag{5}$$

In addition to the control for the rank of market capitalization, which is the assignment variable, I include the same set of controls as in the OLS regressions. I include only time fixed effects since the identification strategy relies on the cross-sectional variation by comparing *switchers* versus *nonswitchers*. Standard errors are double clustered at the stock and day level. Similar to the OLS regression, all variables, except return variables, are standardized cross-sectionally to have a mean of zero and a unit of standard deviation for ease of interpretation. I use first-degree linear polynomial specification for the ranking variable.

Table III, Panel A shows the first-stage regressions. Columns 1 to 2 report the baseline regression with different bandwidths ($\pm$100 and $\pm$200). The results of this test show that switching indices has a strong and significant impact on the ETF flows. The coefficient on the switch indicator in Column 1 suggests that the average ETF flow in the 12 months after reconstitution decreases by about 17.2% of a standard deviation for stocks that switched to the Russell 1000. The magnitude of the estimate is similar, at -0.155 in Column 2, where the bandwidth increases to 200. To mitigate concerns that there might be some anticipated flow-driven trading during the migration period of June, the sample in Columns 3 and 4 omits the day observations from the month of June and the analysis is repeated. The conclusion remains unchanged.

[Table III here]

Table III, Panel B reports the second-stage coefficient estimates of the effect of ETF flows on market-on-close trading volume. The effect of ETF flows on market-on-close trading volume is significant across all samples and bandwidths. The coefficients range from 0.341 to 0.526 and are statistically significant at the 1% level. The economic magnitude from the IV estimates is larger than the OLS estimates. This is not surprising because, first of all, the OLS sample includes all common stocks that have low ETF ownership, while the IV sample focuses on the stocks included in the Russell indices. Another reason is that IV estimates reflect the local average treatment effect, which is the effect of the treatment group compared to the control group around the Russell threshold. The drastic status change for index membership results in a greater impact than for average stock.

Overall, the results using Russell reconstitution provide support for the causal interpretation of the positive relation between ETF flows and market-on-close trading volume.

4.3 Closing auction and cross-sectional return predictability

4.3.1 Price impact during auction period

Given the evidence that ETF flows lead to more market-on-close trading volume, in this section I investigate whether the process of the closing auction affects asset price dynamics. The trading activity in a security provides valuable data about the information structure and subsequent price moves in the security.[25] There is strong empirical evidence that demand and supply shocks can affect individual stock prices. For instance, event studies that focus on compositional changes in the S&P 500 index show that additions increase share prices while delisting decreases prices.[26] Also, the literature on block trades generally finds evidence of temporary price pressure on individual securities.[27] Greenwood (2008) examines transitory price effects upon a weighting change to the Nikkei 225. Furthermore, fire sales from distressed mutual funds also distort stock prices over a long horizon (Coval and Stafford (2007)).

In a closing auction, an auction order can be submitted during normal trading hours until 15:45, but the aggregated market-on-close order imbalance information will not be disseminated for NYSE-listed stocks until 15:45 to members of the public who subscribe to the feed service. Similar to the price pressure hypothesis of

[25]See Blume, Easley, and O'Hara (1994), Campbell, Grossman, and Wang (1993), Conrad, Hameed, and Niden (1994).

[26] See Goetzmann and Massa (2003); Harris (1986); Shleifer (1986); Beneish and Whaley (1996); Lynch and Mendenhall (1997).

[27] Kraus and Stoll (1972), Lakonishok, Shleifer, and Vishny (1992), Chan and Lakonishok (1993), Chan and Lakonishok (1995).

Scholes (1972), where stock prices can diverge from their information-efficient values because of uninformed shocks to excess demand to compensate those who provide liquidity, the price impact should be the strongest for stocks that have a high market-on-close trading volume imbalance when the information disseminates. Furthermore, since this excess demand is not driven by fundamental information about firms, the price impact for stocks that have high market-on-close orders should be symmetrical in terms of buying and selling. To verify this conjecture, I rely on the NYSE proprietary closing auction order imbalance feed data, which provide the direction of the order imbalance, allowing for a comparison of the price impact on both buying and selling orders. The sample in this analysis consists of all NYSE stocks that have at least a $5 share price from March 1, 2010 (when the NYSE changed its dissemination time rule), to June 30, 2018. The price impact is computed as the stock return for the period starting from 15:45 to 16:00, based on the trading prices from the TAQ data.[28] I then run a Fama-Macbeth regression of this return measure on the measure for market-on-close trading volume cross-sectionally. Standard errors are corrected for heteroskedasticity and autocorrelation up to five lags.[29]

[Table IV here]

Table IV shows the Fama-Macbeth regression results, regressing the last 15-minute return on *highmoc_buy* and *highmoc_sell* as the main independent variables, where *highmoc_buy* equals to 1 if the stock is in the top quintile of market-on-close trading volume across all stocks listed on the NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'buy'; *highmoc_sell* is defined in a similar way. Control variables include the turnover ratio and order imbalance.

The results show that, on average, the 15-minute return of stocks with high buying market-on-close volume is significantly higher (10.1 bps) than that of stocks with low market-on-close volume. In contrast, the 15-minute return for stocks with high selling market-on-close volume is significantly lower (-10.8 bps) than that of stocks with low market-on-close volume. These results suggest a price impact during the last 15 minutes for stocks with high market-on-close order imbalances, and the price movement is economically large, regardless of whether the imbalance side is buy or sell.

[28] The price at 15:45 is based on the first observed transaction price that occurred in the period starting at 15:45:00, after intertwining with quota data to exclude outliers, as in Holden and Jacobsen (2014) and Bogousslavsky (2018).

[29] Results are virtually the same when I use various lags.

To address the concern that the price movement might be caused by information other than the closing auction feeds, Panel B of Table IV shows the results of tests that use the 15-minute return for the period from 15:15 to 15:30 and the return for the period from 15:30 to 15:45 as dependent variables. Interestingly, the analysis shows that the price moves in a statistically significant way even before order imbalance disseminates, suggesting a possible information leak.[30] However, from the economic magnitude perspective, the price impacts during both periods are small: the price impact during 15:15 to 15:30 is about 0.5 bps (-0.2 bps) for stocks with high buying (selling) market-on-close order imbalances, and the price impact during 15:30 to 15:45 is about 1.2 bps (-1.4 bps) for stocks with high buying (selling) market-on-close order imbalances.

As a robustness check, I also use the market-on-close trading volume from the TAQ database to repeat the analysis. I use the absolute value of the stock return from 15:45 to 16:00 to capture the price movement because the TAQ database does not reveal the market-on-close order imbalance direction (i.e., the market-on-close order can either buy or sell).

Table A.II, Panel A reports the Fama-Macbeth regression coefficients. The result in Column 1 confirms that the price impact is larger when the market-on-close volume is high, with market-on-close volume again defined as market-on-close trading volume divided by the total trading volume for the day. A 1% increase in market-on-close trading volume leads to 0.28 bps change in price movement for the 15-minute period, and the coefficient is significant at 1%.

To address the concern that the market-on-close trading volume might be skewed, I use the measure, *highmoc*, which is an indicator of whether the stock is in the top quintile of the market-on-close volume for the day. Column 2 reports the coefficient estimates using this measure. Again, the result is consistent with my hypothesis. Cross-sectionally, stocks in the high quintile of market-on-close volume move 2.17 bps more than stocks that are not in the top quintile. Panel B repeats the same placebo tests that use absolute returns for the period from 15:15 to 15:30 and the absolute return for the period from 15:30 to 15:45 as dependent variables. Contrary to the results in Table IV, the coefficient estimates on the *highmoc* from both columns are insignificant at the 5% level. This supports the argument that the price impact during the last 15 minutes is due to the closing auction dissemination.

[30] The U.S. Securities Exchange Commission fined the NYSE for improper distribution of market data (including closing auction imbalance data) in 2012 and 2014. Mayhew, McCormick, and Spatt (2009) show that specialists could front-run the closing auction before such information is disseminated to floor traders.

The analysis so far shows the absolute level of the price impact due to the high market-on-close order imbalance. For investors who submit market-on-close orders, the alternative trading strategy could be simply placing market orders during normal trading hours. Given the large quantity of trading needs, the price impact of such block trades could also be large. To quantify the economic magnitude of the transaction costs from using market-on-close orders relative to counterfactual situations, I conduct a market impact cost analysis following Korajczyk and Sadka (2004).

Specifically, I consider four alternative price impact measures: the Breen, Hodrick, and Korajczyk (2002) (henceforth BHK) measure; the Glosten and Harris (1988) (henceforth GH) measure; and the effective and quoted spread. Korajczyk and Sadka (2004) use intraday transaction data from January 1993 to May 1997 to estimate market impact costs. These costs are estimated each month for each stock using the TAQ data. Using cross-sectional relationships between these market impact costs and firm characteristics, the out-of-sample price impact costs can then be estimated.[31]

[Table V here]

Table V presents the counterfactual price impact estimates together with the actual price impact for stocks with high market-on-close order imbalance. The results show that the price impact would ranges from 0.12 bps to 3.39 bps among all alternative price impacts estimates.[32] The estimates for a high buying market-on-close and a high selling market-on-close are similar. However, all the estimates from these four specifications are less than the actual price impact (11.24 bps for buying and 11.46 bps for selling).

Taken together, these results suggest that the price impact due to the high market-on-close order imbalance volume is large during the last 15 minutes of the trading day and it is symmetrical in terms of buying and selling orders.

[31] See the Appendix for a detailed estimation of each measure.

[32] The estimate is based on the same market-on-close trading volume and the time interval is based on 30 minutes, following Korajczyk and Sadka (2004). The effective spread and the quoted spread are proportional to the trading volume, while the BHK measure and the GH measure are nonproportional to the trading volume.

4.3.2 Closing auction and price reversal

Having provided evidence consistent with a price impact for stocks that have high market-on-close trading volume during the last 15 minutes of the trading day, I now test whether the closing auction process impacts the price movement beyond this 15-minute interval. Since the demand shocks originate from nonfundamental shocks, such a price impact should be transitory and should lead to price reversal. To test this hypothesis, I use Fama-Macbeth regressions at a daily frequency to analyze the cross-sectional relation between return reversals and market-on-close trading volume as follows.

$$Return_{i,t+1} = \alpha + \beta_1 \ Return_{i,t} + \beta_2 \ Return_{i,t} \times highmoc_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \tag{6}$$

The sample consists of all common NYSE stocks from March 2010 to June 2018 with market-on-close trading volume determined from the TAQ data, as in the previous section. The dependent variable in our baseline regression is the stock's daily return (close-to-close return). To capture the price reversal, I interact the indicator variable *highmoc* with the previous day's return, so that a negative and significant coefficient β_2 would suggest that, cross-sectionally, a stock with *highmoc* exhibits a negative serial correlation. I control for a number of stock characteristics that might relate to the prediction of short-term returns: *log(market cap)*, *Amihud ratio*, *order imbalance*, and *turnover*. Standard errors are corrected for heteroskedasticity and autocorrelation up to five lags (Newey and West (1987)).

[Table VI here]

Table VI reports the coefficient estimates from the Fama-MacBeth regression. The results in Column 1, Panel A show two findings. First, there is weak evidence of negative first-order serial correlation at a daily frequency for stocks that do not have a high market-on-close order imbalance in our sample period. The coefficient on $Return_t$ is negative (−0.004) but insignificant.[33] Second, the β_2 coefficient implies that stocks with high market-on-close trading volume exhibit strong negative autocorrelation, which is consistent with my hypothesis that the price pressure is temporary.[34] The estimate of the coefficient is −0.022 with the t-statistic

[33] Chordia et al. (2005, Table 1) find that these autocorrelations are smaller in more recent subperiods.

[34] A large literature studying stock return serial correlation shows that at a monthly or annual frequency, there is significant negative

of −7.06. The magnitude of the reversal for stocks that have high market-on-close trading volume is more than five times larger than the reversals for stocks that have lower trading volume. Column 2 shows this result after controlling for stock characteristics that relate to short-term return predictability. The β_2 coefficient is −0.019 and remains significant at the 1% level.

The evidence so far suggests that stocks in the top quintile of high market-on-close trading volume have a stronger reversal using *close-to-close* returns. To closely investigate when such a reversal occurs and whether it is persistent, I decompose the close-to-close returns into overnight returns and intraday returns, following Lou, Polk, and Skouras (2018) and Bogousslavsky (2018).[35] Specifically,

$$r^i_{intraday,s} = \frac{P^i_{close,s}}{P^i_{open,s}} - 1, \tag{7}$$

$$r^i_{overnight,s} = \frac{1 + r^i_{close-to-close,s}}{1 + r^i_{intraday,s}} - 1.$$

Columns 3 and 4 of Table VI, Panel A report the cross-sectional regression results, using these two returns as dependent variables, respectively. The results show that the reversal is quite persistent and significant both during the next open-to-close period (−0.011) and during the overnight period (−0.007). Interestingly, β_1,the coefficient on $Return_{i,t}$ is also significant and negative when the dependent variable is the overnight return. This result comports with the finding in Lou, Polk, and Skouras (2018) that stocks with relatively high overnight returns over the previous month have, on average, relatively low intraday returns in the subsequent month.[36]

The evidence using the TAQ data thus far suggests that stock returns for high market-on-close trading volume will reverse regardless of whether the market-on-close order imbalance is high *buying* or high *selling*. As a robustness check of whether the reversal pattern might differ for buying versus selling orders, I conducted an additional subsample analysis using NYSE proprietary closing auction order imbalance data focusing on

serial correlation (Jegadeesh (1990)); however, at a daily frequency, the results are weak (Fama and French (1988)). Collin-Dufresne and Daniel (2016) show that the reversal effect for residual return is remarkably strong in even the largest 100 most liquid stocks.

[35]The decomposition assumes that dividend adjustments, share splits, and other corporate events that could mechanically move prices take place overnight, and this assumption is reasonable, as argued in Lou, Polk, and Skouras (2018).

[36]See also Hendershott, Livdan, and Rösch (2018) and Bogousslavsky (2018) about the return difference over the trading day and overnight.

NYSE-listed stocks. Specifically, I run the similar Fama-MacBeth regression as follows.

$$\begin{aligned} Return_{i,t+1} = \alpha + \beta_1\ Return_{i,t} + \beta_2\ highmoc\_buy_{i,t} \times Return_{i,t} + \\ \beta_3\ highmoc\_sell_{i,t} \times Return_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{8}$$

The independent variable and dependent variable are the daily returns, and the dummy variable *highmoc_buy* is an indicator of whether (a) the stock is in the top quintile of market-on-close trading volume across all stocks listed on the NYSE and (b) the imbalance side of the closing auction at the beginning of dissemination is buy. The indicator *highmoc_sell* is defined similarly. The main variables of interest are β_2 and β_3, which capture the return reversal for high buying market-on-close orders and high selling market-on-close orders, respectively.

[Table VII here]

Table VII shows the Fama-Macbeth coefficient results. Consistent with the previous finding, the coefficients β_2 and β_3 load significantly negatively across all regressions. This suggests a strong return reversal for both high buying market-on-close orders and high selling market-on-close orders. Furthermore, for stocks that have high buying market-on-close orders, this reversal is slightly stronger than the reversal pattern for stocks with high selling market-on-close orders. Columns 3 and 4 show the results using subsequent intraday returns and subsequent overnight returns as their respective dependent variables. The reversal pattern is persistent and stronger during the overnight period.

To address the concern the my return analysis used is based on the close-to-close return, which might be affected by the price pressure from previous day's closing auction, Table A.III, Panel A shows the repeated analysis using the same day open-to-close return as the independent variable and also in the interaction term. The coefficient of interest β_2 is -0.015 (t-stat=-5.09) when the dependent variable is the close-to-close return and -0.008 (t-stat=-2.81) when the dependent variable is the subsequent intraday return. As an additional robustness check, Table A.III, Panel B uses the return from 15:45 to 16:00 as the independent variable and in the interaction term. The conclusions remain unchanged.

Studies show that market microstructure bias and small cap stocks could contribute to the short-term reversal. To alleviate the concern that the short-term reversal documented here is not driven by these biases, I conduct two additional tests, and results are presented in the Appendix.

The first test controls for size explicitly by dividing stocks into different groups. In particular, at the end of each month, stocks are sorted into five groups based on their market capitalization. Then, I run Fama-Macbeth regressions for those stocks classified into different groups. The results in Table A.IV show that, across all size groups, the coefficient β_2 is negative and significant. (Even for stocks from the large size group, the estimate is -0.0197 with t-stat of -3.64.)

The second test reported in A.V uses the average of the bid and ask quotes at closing to calculate the close-to-close returns, and the results suggest that the price reversal cannot be explained by bid-ask bounce bias. Nonsynchronous trading bias is another market microstructure bias that potentially leads to negative serial correlation. However, since all stocks in the sample involve the closing auction, which occurs at same time across all stocks at 16:00, it is unlikely that the reversal pattern can be explained by nonsynchronous trading.

Taken together, stocks with high market-on-close trading volume experience a larger price impact during the last 15 minutes of trading. Since flow-driven trading is due to nonfundamental shocks, the returns reverse overnight and in the following day. Combining the evidence that ETF flow leads more market-on-close orders, the results lend supportive evidence to the recent finding that ETF ownership leads to increasing deviation from a random walk for underlying assets (Ben-David, Franzoni, and Moussawi (2018); Coles, Heath, and Ringgenberg (2017)).

4.3.3 Long–short trading strategies

Given the significant serial correlation documented here for stocks that have high market-on-close trading volume, one might wonder whether a closing auction imbalance can be used as a signal to form profitable trading strategies. In this section, I consider two trading strategies. The first is to form daily equal-weighted, long–short portfolios based on the direction of the closing auction imbalance side on the previous day and the closing auction total imbalance quantity from the NYSE proprietary data. Specifically, at the end of day t, stocks that are (a) in the highest quintile of the closing auction imbalance quantity and (b) on the imbalance side of selling based on the first disseminated closing auction information at 15:45 will be bought. Meanwhile, the stocks will be sold short if they are (a) in the highest quintile of the closing auction total imbalance quantity but (b) on the imbalance side of buying. The sample consists of all nonzero closing auction volume stocks available from the NYSE order imbalance feed data from July 2008 to June 2018.

[Table VIII here]

Table VIII, Panel A reports the alphas and factor loadings for such a long–short portfolio. The results show that the alpha for the long–short portfolio is statistically and economically large: 13.29 bps for the CAPM model (with a t-stat of 10.11) and 13.22 bps for the Fama–French three-factor model (with a *t*-stat of 9.81) and 13.20 bps for a four-factor model (with a t-stat of 9.91). This is equivalent to a 33.26% risk-adjusted annualized return. The factor loadings on SMB, HML, and UMD are insignificant, suggesting the return from this long–short trading strategy is not driven by these common factors. The return decomposition from the long and short legs of the portfolio shows that both legs yield significant alphas (6.07 bps for the long leg and −7.23 for the short leg).

The second trading strategy is to form daily equal-weighted, long–short portfolios based on the direction of returns in the previous day and the market-on-close trading volume, using only the TAQ data.This trading strategy does not rely on the proprietary real time closing auction order imbalance data. Nagel (2012) also uses the lagged return as a noisy proxy for unobserved market-maker inventory imbalance when investigating the short-term reversal strategies. Specifically, at the end of day t, stocks that are (a) in the highest quintile of market-on-close trading volume and (b) also have a negative return will be bought, and stocks that are (a) in the highest quintile of market-on-close trading volume but (b) have a positive return will be sold short. The sample consists of all common stocks with nonzero market-on-close volume from July 2008 to June 2018, and the holding horizon starts from the close of day *t* to the close of the following day.

Table VIII, Panel B reports the alphas and factor loadings from this long–short portfolio. The zero–investment portfolio has, on average, a market-adjusted return of 7.44 bps per day and it is statistically significant. After adjusting for the Fama–French three factors and Fama–French plus momentum, the alphas maintain almost the same magnitude (7.36 bps and 7.37 bps, respectively) and are still significant at 1%. Similar to the results from the first trading strategy, both the long and short legs yield significant alphas (2.69

bps for the long leg and -4.70 for the short leg).

A caveat with this long/short strategy is that it assumes the return direction (positive or negative) is the same as the order imbalance direction, which might not be true if the price movement during the day were information-driven. Another assumption is that stocks with a high market-on-close order imbalance can be quickly identified, using the record of the TAQ data, which may not be available until after the market closes. This strategy assumes that orders can be placed at the closing price, which may not be realistic. Given the fact that the price reversal is persistent during the overnight period and also the subsequent intra day, modified trading strategies that hold stock during the overnight period and the next intraday period are also considered.

Table VIII, Panel C reports the different holding period returns from these two trading strategies. The holding horizon starts either from the same day's close to the next morning's open (i.e., the overnight return in Columns 1 and 3) or from next morning's open to the next day's close (i.e., the open-to-close return in Columns 2 and 4). The results show that the return for the long–short portfolio comes not only from the overnight period (5.55 bps with a *t*-stat of 13.83) but also from the open-to-close period (7.84 bps with a t-stat of 7.38).

As a robustness check, Table A.VI shows the results of the subperiod analyses. The results are robust across different sample periods. The fact that the alphas from the second period are lower than the first period is consistent with the conjecture that the market becomes more efficient regarding liquidity provision during the closing auction. Value-weighted portfolios are also formed under both trading strategies, and the results presented in Table A.VII show consistent but weaker abnormal returns, suggesting some of the abnormal returns can be attributed to the liquidity constraints from the small cap stocks.

While the size and significance of such a long–short strategy indicates the return predictability based on closing auction order imbalances, transaction costs for individual investors could erode these profits. Nevertheless, it is possible for professional investors (e.g., high-frequency traders) with low trading costs to exploit such a trading strategy. Another concern for such a trading strategy is information availability, especially the closing auction volume. These trading strategies suggest that providing liquidity to passive funds by assessing the closing auction order imbalances is more profitable than using the public TAQ data.[37]

[37]Major exchanges provide the closing auction feeds that provide information about closing auction volume at high frequency during the closing auction. The subscription fee to the stock-exchange market-data feed is not cheap, and it has increased recently,

Overall, the takeaway from this section is that the price reversal, due to the closing auction order imbalance, caused by ETF flows, can be exploited. And, by providing liquidity to passive funds, the trading strategies result in significant abnormal returns.

4.4 Closing auction, ETF ownership, and volatility

Ben-David, Franzoni, and Moussawi (2018) find that stocks with higher ETF ownership display significantly higher volatility because the liquidity shocks propagate to the underlying securities through the arbitrage channel. Besides the continuous arbitrage activities throughout the day, closing auction orders placed by passive funds and authorized participants could cause additional volatility to the underlying securities due to the price impact and the price reversals documented here. To verify this conjecture, I first replicate the main analysis in Ben-David, Franzoni, and Moussawi (2018) using S&P 500 stocks from 2010 to 2018 by regressing volatility on ETF ownership and controls. The ETF ownership is defined as follows.

$$ETF\ ownership_{i,t} = \frac{\sum_{j=1}^{J} w_{i,j,t} AUM_{j,t}}{MktCap_{i,t}} \tag{9}$$

Following their procedure, the analysis is conducted at the monthly frequency, where the daily volatility is computed using all daily returns within the month. All controls are from the end of the prior month. Standard errors are double clustered at the stock and month level. Volatility and ETF ownership are standardized by subtracting the sample mean and then dividing by the sample standard deviation in the entire sample. To purge the possible effect of the closing auction, I calculate the daily return using the price from opening and the transaction price at 15:45. Then the monthly alternative volatility is computed based on this return measure. The correlation between these two volatility measures is around 0.77.

[Table IX here]

Table IX presents the results. Columns 1 and 2 show that, consistent with the results from Ben-David, Franzoni, and Moussawi (2014), stocks with higher ETF ownership have higher volatility, which is computed based on the close-to-close return. The magnitude of the coefficient resembles the one they document: a

though the SEC recently overturned the approval of increasing higher fees for certain feed data. See https://www.wsj.com/articles/sec-to-rule-nyse-nasdaq-didnt-justify-market-data-fee-increases-1539721232?mod=mktw.

one-standard-deviation increase in ETF ownership is associated with a 10.5% of a standard deviation change in daily volatility. Columns 3 and 4 show the same panel regression results with the volatility based on an alternative return measure. The magnitude of the relationship between ETF ownership and volatility decreases by 42.6% (60.4%) to 0.060 (0.026) in Column 3 (Column 4). These results provide suggestive evidence that the closing auction might contribute to the higher volatility of stocks with higher ETF ownership.

4.5 Do informed traders use market-on-close orders?

The results so far have focused on the impact of passive investing on stock price dynamics through the channel of market-on-close orders, based on the premise that passive funds and authorized participants are the main users of market-on-close orders. However, it remains unclear whether market-on-close orders could be used by other market participants. Models, such as that of Admati and Pfleiderer (1988), shows that the increasing level of liquidity trading induces more informed trading, consistent with the empirical finding that the average volume of shares traded is U-shaped and clustered within one day. Given the increasing intensity of market-on-close volume over the past decade, are informed traders induced to place market-on-close orders to camouflage their trading intentions? In this section, I provide some suggestive evidence supporting this argument.

To begin with, I construct the following measure to quantify market-on-close trades that do not originate with passive funds:

$$Volume\ residual_{i,t} = \frac{\$Volume_{moc,i,t} - \$Volume_{ETF\ flows,i,t}}{\$Volume_{total\ tradingi,t}} \tag{10}$$

where $\$Volume_{moc,i,t}$ represents the total dollar trading volume of the market-on-close orders, and $\$Volume_{ETF\ flows,i,t}$ represents the total aggregate ETF flows in dollar volume. The difference between these two is then normalized by the total daily trading volume.

[Figure 2 here]

Figure 2 provides the time-series trend of this measure, averaged cross-sectionally among S&P 500 stocks and all common stocks. The pattern shows that over time, the market-on-close trading volume that does not originate from the ETF flows have increased. This evidence is consistent with the hypothesis that other market participants might also use market-on-close orders. However, caution is called for regarding this noisy proxy: First of all, due to the limits of data availability, I can capture only the daily flows from ETFs, not the daily flows from index mutual funds.[38] The increasing trend could be due to the unmeasured flows from index mutual funds and other index-linked trading strategies. Second, the measure assumes ETF flows are being executed 100% by the closing auction, which might not be true. Therefore, I conducted two additional tests.

4.5.1 Closing auction, informed trading, and price reversals

The main analysis of price reversals shows that stocks with high closing auction order imbalances will reverse because their prices are driven by passive flows that are not related to fundamental. However, if the price impact from the high closing auction is due to private information, the price impact will be permanent and will not induce negative serial correlation (Glosten and Milgrom (1985) and Nagel (2012)). Among stocks in the high quintile of closing auction volume residual, defined in Equation (10) cross-sectionally, market-on-close orders with a direction of sell (buy) but with positive (negative) aggregate ETF inflows are more likely to be information-driven orders. In other words, a stock that is sold (bought) by aggregate ETFs (inferred based on the direction of flows) but has an aggregate buying (selling) market-on-close order imbalance might be a stock traded by informed traders. To formally test this, I repeat the same Fama-Macbeth regression based on the specification in Equation (8) across different subsamples.

[Table X here]

Table X presents the results. The first column is the same as in Table VII, including all NYSE stocks with standard filters. Both the coefficients on the interaction terms are significant, and the estimate in Column 1 are slightly different from those in Table VII because the sample analysis here requires that the aggregated ETF flows at the stock level cannot be missing. The second column repeats the analysis for those stocks that are less likely to have information-driven closing trades, and the last column shows the coefficient estimates for those

[38] A number of studies that focus on daily mutual fund flows, such as those by Chalmers, Edelen, and Kadlec (2001); Edelen and Warner (2001); Greene and Hodges (2002); Rakowski (2010); Kaniel and Parham (2017); and Agarwal, Jiang, and Wen (2018) use TrimTab data, which rely on voluntary disclosure from fund managers and therefore have limited coverage of the funds. For instance, TrimTab data do not cover passive funds from Vanguard and Blackrock, which account for more than 50% of the market share of passive funds, according to Morningstar.

stocks that are more likely to have information-driven closing trades. Consistent with the conjecture, the price reversal for stocks that are more likely to have information-driven closing trades is weaker and insignificant.

4.5.2 Closing auction and informed trading around earnings announcement

In this subsection, I utilize earnings announcements as a laboratory setting to provide further evidence that market-on-close orders could be used by informed traders. The goal of earnings announcements, as major company information events, is to release information to the market, and informed investors should be especially active at these times.

For instance, Jegadeesh and Titman (1993) estimate that the three-day returns around earnings announcements represent approximately 25% of momentum profits. La Porta, Lakonishok, Shleifer, and Vishny (1997) report that about 25% of the returns to various value strategies considered by Lakonishok, Shleifer, and Vishny (1994) are concentrated on the three days around earnings announcements. More recently, Engelberg, Mclean, and Pontiff (2018) investigate 97 stock return anomalies and find that anomaly returns are six times higher on earnings announcement days.

This evidence suggests that the short window around earnings announcements is a period in which informed trades occur. Market-on-close orders could be used by informed traders as a way to enter or exit the positions since trading along with passive funds using market-on-close trades can hide their intentions and increase the profits they gain from private information.(e.g., Admati and Pfleiderer (1988)).

To begin with, I follow the literature (Dellavigna and Pollet (2009); Kaniel, Liu, Saar, and Titman (2012); and Hirshleifer, Lim, and Teoh (2009)) by constructing the normalized earnings surprise and abnormal returns during the earnings announcement as proxies for the surprise. Specifically, I define the standardized unexpected earnings (SUE) as actual earnings minus the earnings forecast then divided by the price on the forecast date. The earnings forecast is the mean analyst forecast one month before the earnings announcement. Another earnings surprise proxy is the abnormal return at the time of the earnings announcement, which is defined as CAR[-1,1], the cumulative buy-and-hold return from one day before the earnings announcement to one day after, adjusted for the characteristics benchmarks described in Daniel, Grinblatt, Titman, and Wermers (1997) (i.e., size, book-to-market, and momentum).

To measure the abnormal market-on-close trading volume that does not derive from passive ETF flows, I follow Dellavigna and Pollet (2009) and define

$$\Delta V_{t,i}^{(h,H)} = \sum_{u=\tau+h}^{\tau+H} \log\left(Volume\ residual_{t,i}^{u}\right)/(H-h+1) - \sum_{u=\tau-14}^{\tau-5} \log\left(Volume\ residual_{t,i}^{u}\right)/10 \tag{11}$$

where $Volume\ residual_{t,i}^{u}$ is calculated in Equation (10) as the market-on-close trading residual on day u and τ is the date of the earnings announcement in quarter t for company i. The measure $\Delta V_{t,k}^{(h,H)}$ is the percentage increase in volume around the announcement date at horizon (h,H). In particular, I investigate three abnormal volume during the three horizons: $\Delta V_{t,k}^{(-4,-2)}$, $\Delta V_{t,k}^{(-1,1)}$, and $\Delta V_{t,k}^{(2,4)}$.

In each calendar quarter from 2010 to 2018, quarterly earnings announcement observations are sorted in that quarter based on the absolute value of the earnings surprise (CAR[-1,1]). I then calculate the mean abnormal market-on-close residual trading volumes before the announcement, around the announcement, and after the announcement. The abnormal trading volume difference between earnings surprise quintiles 5 and 1 captures the trading volume difference that responds to different magnitudes of earnings surprises.

[Table XI here]

Table XI, Panel A reports the results. During the announcement period [-1, 1] and the post-announcement period [2, 4], there are significant increases in abnormal trades that use the market-on-close orders. Furthermore, stocks in the top quintile of earning surprises experience significantly higher abnormal market-on-close trading volume than stocks in the bottom quintile of earnings surprises across three periods. There is little evidence of increasing abnormal market-on-close trades in the pre-announcement period [-4, -2] for stocks in the bottom quintile of earnings surprises. An analysis using the cumulative abnormal return as a proxy for earnings surprise shows a similar pattern.

Table XI, Panel B takes a step further to analyze the trading imbalance pattern during the earnings announcement periods. By leveraging the direction of both ETF flows and order imbalances, I construct the following signed measure of market-on-close trades:

$$Trade\ Imb\ residual_{i,t} = \frac{Signed\ \$Volume_{closing\ auction,i,t} - Signed\ \$Volume_{ETF\ flows,i,t}}{\$Volume_{total\ tradingi,t}} \tag{12}$$

As mentioned earlier, this measure has two limitations that can undermine the statistical power of empirical tests. First, it captures only the flows from ETFs, not from index mutual funds. Second, it assumes that all ETF flows are executed entirely through the closing auction. Nevertheless, I attempt to provide the best possible evidence using this measure. The abnormal trading imbalances are constructed in a way similar to abnormal trading volume in Equation (11), with a negative sign meaning abnormal selling and a positive sign meaning abnormal buying.

One distinct pattern in Panel B is the substantial buying for stocks with negative earnings surprises and selling for stocks with positive earnings surprises in the post-earnings announcement period [2, 4]. This suggests that investors use market-on-close orders to exit the position they accumulated during or before the earnings announcement. This result is consistent with the pattern of "buying the rumor and selling the news" in the context of individual investors (Kaniel, Liu, Saar, and Titman (2012)) as well as institutional investors (Kadan, Michaely, and Moulton (2018)).

Taken together, the results in this section provide several lines of supportive evidence that the market-on-close orders can be used by informed traders. This is consistent with the theoretical implication that high liquidity trades induce more informed trades.

5 Conclusion

This paper investigates the popularity of increasing closing auction trades. First, I show that due to the increase in passive investing, market-on-close trading volumes are increasing dramatically. Stocks with more passive flows from ETFs have higher market-on-close trading volume than otherwise similar securities. Using a quasi-natural experiment based on the reconstitution of the Russell indices, I provide a causal interpretation of this result.

Next, I investigate the impact of increasing market-on-close trading volume on stock price dynamics. I show that stocks with high market-on-close trading volume have larger price movement during the end of day auction period. Furthermore, since the demand shock from ETFs is nonfundamental, stocks that have high market-on-close trading volume cross-sectionally show a strong return reversal pattern that is consistent with the temporary price pressure hypothesis, and investors are compensated by providing liquidity to passive investors. A long–short trading strategy that exploits this temporary price pressure results in a significant risk-adjusted 13.2 bps per day, or 33.26% annualized return.

Besides passive funds, market-on-close orders might also be used by informed traders, consistent with the theoretical implications of the clustering of informed trades and liquidity trades.

Overall, these findings provide direct evidence that the closing auction is an important channel through which demand shocks from ETFs can propagate into underlying securities. The significant price impact for stocks that have high market-on-close trading volume suggests that the closing price for some stocks may be distorted due to the increasing volume of market-on-close orders, which originate from aggregate passive flows. My findings have important implications for investor welfare. If the popularity of market-on-close orders causes price to deviate from their fundamental value, then some passive funds may benefit from executing their orders at other than closing prices at the expense of larger tracking errors.

References

Admati, Anat R., and Paul Pfleiderer, 1988, A Theory of Intraday Patterns: Volume and Price Variability, *The Review of Financial Studies* 1, 3–40.

Admati, Anat R., and Paul Pfleiderer, 1991, Sunshine Trading and Financial Market Equilibrium, *The Review of Financial Studies* 4, 443–481.

Agarwal, Vikas, Paul Hanouna, Rabih Moussawi, and Christof Stahel, 2019, Do ETFs Increase the Commonality in Liquidity of Underlying Stocks?, *Working Paper* .

Agarwal, Vikas, Lei Jiang, and Quan Wen, 2018, Why Do Mutual Funds Hold Lottery Stocks?, *Working paper* .

Akbas, Ferhat, Will J. Armstrong, Sorin Sorescu, and Avanidhar Subrahmanyam, 2015, Smart money, dumb money, and capital market anomalies, *Journal of Financial Economics* 118, 355–382.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2012, Performance of institutional trading desks: An analysis of persistence in trading costs, *Review of Financial Studies* 25, 557–598.

Antón, Miguel, and Christopher Polk, 2014, Connected Stocks, *Journal of Finance* 69, 1099–1127.

Appel, Ian R., Todd A. Gormley, and Donald B. Keim, 2016, Passive investors, not passive owners, *Journal of Financial Economics* 121, 111–141.

Appel, Ian R, Todd A Gormley, and Donald B Keim, 2018, Standing on the shoulders of giants: The effect of passive investors on activism, *The Review of Financial Studies* 32, 2720–2774.

Ben-David, ITZHAK, Francesco Franzoni, Augustin Landier, and Rabih Moussawi, 2013, Do hedge funds manipulate stock prices?, *The Journal of Finance* 68, 2383–2434.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2014, Do ETFs Increase Volatility?, Technical Report w20071, National Bureau of Economic Research, Cambridge, MA.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl, 2011, The Price Pressure of Aggregate Mutual Fund Flows, *The Journal of Financial and Quantitative Analysis* 46, 585–603.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do ETFs Increase Volatility?, *Journal of Finance, forthcoming* .

Beneish, Messod D., and Robert E. Whaley, 1996, An Anatomy of the "S & P Game": The Effects of Changing the Rules, *Journal of Finance* 51, 1909–1930.

Blume, Lawrence, David Easley, and Maureen O'Hara, 1994, Market Statistics and Technical Analysis: The Role of Volume, *Journal of Finance* 49, 153–181.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2008, Which Shorts Are Informed?, *Journal of Finance* 63, 491–527.

Bogousslavsky, Vincent, 2018, The Cross-Section of Intraday and Overnight Returns, *Working Paper* .

Box, Travis, Ryan Davis, Richard Evans, and Andrew Lynch, 2019, Intraday Arbitrage Between ETFs and Their Underlying Portfolios, *Working paper* 73.

Breen, William J, Laurie Simon Hodrick, and Robert A Korajczyk, 2002, Predicting Equity Liquidity, *Management Science* 48, 15.

Broman, Markus S, and Pauline Shum, 2018, Relative liquidity, fund flows and short-term demand: Evidence from exchange-traded funds, *Financial Review* 53, 87–115.

Brown, David C., Shaun Davies, and Matthew Ringgenberg, 2018, ETF Arbitrage and Return Predictability, SSRN Scholarly Paper ID 2872414, Social Science Research Network, Rochester, NY.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *The Quarterly Journal of Economics* 108, 905–939.

Chalmers, John M. R., Roger M. Edelen, and Gregory B. Kadlec, 2001, On the Perils of Financial Intermediaries Setting Security Prices: The Mutual Fund Wild Card Option, *Journal of Finance* 56, 2209–2236.

Chan, Louis K. C., and Josef Lakonishok, 1993, Institutional trades and intraday stock price behavior, *Journal of Financial Economics* 33, 173–199.

Chan, Louis K. C., and Josef Lakonishok, 1995, The Behavior of Stock Prices Around Institutional Trades, *Journal of Finance* 50, 1147–1174.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression Discontinuity and the Price Effects of Stock Market Indexing, *The Review of Financial Studies* 28, 212–246.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485–518.

Coles, Jeffrey L., Davidson Heath, and Matthew Ringgenberg, 2017, On Index Investing, *Working paper* .

Collin-Dufresne, Pierre, and Kent Daniel, 2016, Liquidity and Return Reversals, *Working paper* 42.

Comerton-Forde, Carole, and Tālis J Putniņš, 2011, Measuring closing price manipulation, *Journal of Financial Intermediation* 20, 135–158.

Conrad, Jennifer S., Allaudeen Hameed, and Cathy Niden, 1994, Volume and Autocovariances in Short-Horizon Individual Security Returns, *Journal of Finance* 49, 1305–1329.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Cushing, David, and Ananth Madhavan, 2000, Stock returns and trading at the close, *Journal of Financial Markets* 3, 45–67.

Da, Zhi, and Sophie Shive, 2018, Exchange traded funds and asset return correlations, *European Financial Management* 24, 136–168.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *Journal of Finance* 52, 1035–1058.

Dellavigna, Stefano, and Joshua M. Pollet, 2009, Investor Inattention and Friday Earnings Announcements, *Journal of Finance* 64, 709–749.

Duffie, Darrell, 2010, Presidential address: Asset price dynamics with slow-moving capital, *The Journal of finance* 65, 1237–1267.

Edelen, Roger M, and Jerold B Warner, 2001, Aggregate price effects of institutional trading: a study of mutual fund flow and market returns, *Journal of Financial Economics* 59, 195–220.

Engelberg, Joseph, R. David Mclean, and Jeffrey Pontiff, 2018, Anomalies and News, *Journal of Finance* 73, 1971–2001.

Etula, Erkko, Kalle Rinne, Matti Suominen, and Lauri Vaittinen, 2019, Dash for Cash: Monthly Market Impact of Institutional Liquidity Needs, *The Review of Financial Studies* hhz054.

Fama, Eugene F., and Kenneth R. French, 1988, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.

Frazzini, Andrea, and Owen A. Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

Gao, Lei, Yufeng Han, Sophia Zhengzi Li, and Guofu Zhou, 2018, Market intraday momentum, *Journal of Financial Economics* 129, 394–414.

Glosten, Lawrence R, and Lawrence E Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123–142.

Glosten, Lawrence R, and Paul R Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Glosten, Lawrence R, Suresh Nallareddy, and Yuan Zou, 2016, Etf activity and informational efficiency of underlying securities, *Columbia Business School Research Paper* .

Goetzmann, William N., and Massimo Massa, 2003, Index Funds and Stock Market Growth, *Journal of Business* 76, 1–28.

Greene, Jason T, and Charles W Hodges, 2002, The dilution impact of daily fund flows on open-end mutual funds, *Journal of Financial Economics* 65, 131–158.

Greenwood, Robin, 2008, Excess Comovement of Stock Returns: Evidence from Cross-Sectional Variation in Nikkei 225 Weights, *Review of Financial Studies* 21, 1153–1186.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Harris, Lawrence, 1986, A transaction data study of weekly and intradaily patterns in stock returns, *Journal of Financial Economics* 16, 99–117.

Hendershott, Terrence, Dmitry Livdan, and Dominik Rösch, 2018, Asset Pricing: A Tale of Night and Day, *Working paper* .

Heston, Steven L, Robert A Korajczyk, and Ronnie Sadka, 2010, Intraday patterns in the cross-section of stock returns, *Journal of Finance* 65, 1369–1407.

Hillion, Pierre, and Matti Suominen, 2004, The manipulation of closing prices, *Journal of Financial Markets* 7, 351–375.

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

Holden, Craig W., and Stacey Jacobsen, 2014, Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions, *Journal of Finance* 69, 1747–1785.

Israeli, Doron, Charles M. C. Lee, and Suhas A. Sridharan, 2017, Is there a dark side to exchange traded funds? An information perspective, *Review of Accounting Studies* 22, 1048–1083.

Jegadeesh, Narasimhan, 1990, Evidence of Predictable Behavior of Security Returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* 48, 65–91.

Kadan, Ohad, Roni Michaely, and Pamela C. Moulton, 2018, Trading in the Presence of Short-Lived Private Information: Evidence from Analyst Recommendation Changes, *Journal of Financial and Quantitative Analysis* 53, 1509–1546.

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual Investor Trading and Return Patterns around Earnings Announcements, *Journal of Finance* 67, 639–680.

Kaniel, Ron, and Robert Parham, 2017, WSJ Category Kings – The impact of media attention on consumer and mutual fund investment decisions, *Journal of Financial Economics* 123, 337–356.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are Momentum Profits Robust to Trading Costs?, *Journal of Finance* 59, 1039–1082.

Kraus, Alan, and Hans R. Stoll, 1972, Price Impacts of Block Trading on the New York Stock Exchange, *Journal of Finance* 27, 569–588.

La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny, 1997, Good News for Value Stocks: Further Evidence on Market Efficiency, *Journal of Finance* 52, 859–874.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, *Journal of Finance* 49, 1541–1578.

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

Lou, Dong, 2012, A Flow-Based Explanation for Return Predictability, *Review of Financial Studies* 25, 3457–3489.

Lou, Dong, Christopher Polk, and Spyros Skouras, 2018, A Tug of War: Overnight Versus Intraday Expected Returns, *Journal of Financial Economics, forthcoming* 51.

Lynch, Anthony W., and Richard R. Mendenhall, 1997, New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index, *The Journal of Business* 70, 351–383.

Mayhew, Stewart, Tim McCormick, and Chester S. Spatt, 2009, The Information Content of Market-on-Close Imbalances, the Specialist and NYSE Equity Prices, *Working paper* .

Nagel, Stefan, 2012, Evaporating Liquidity, *The Review of Financial Studies* 25, 2005–2039.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Pan, Kevin, and Yao Zeng, 2019, Etf arbitrage under liquidity mismatch, *Working Paper* .

Petersen, Mitchell A., 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* 22, 435–480.

Rakowski, David, 2010, Fund Flow Volatility and Performance, *The Journal of Financial and Quantitative Analysis* 45, 223–237.

Schmidt, Cornelius, and Rüdiger Fahlenbrach, 2017, Do exogenous changes in passive institutional ownership affect corporate governance and firm value?, *Journal of Financial Economics* 124, 285–306.

Scholes, Myron, and Joseph Williams, 1977, Estimating betas from nonsynchronous data, *Journal of Financial Economics* 5, 309–327.

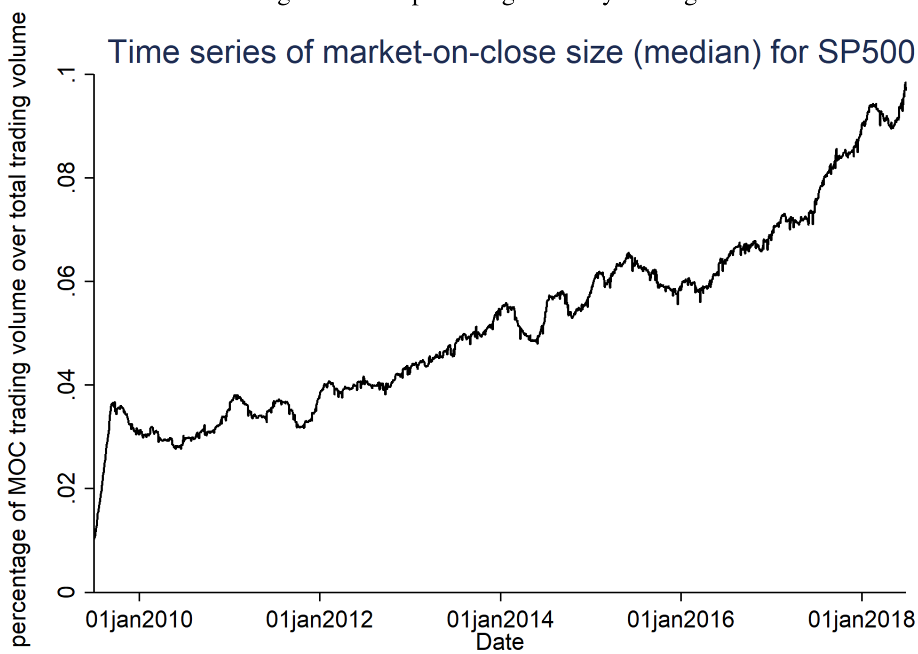
Scholes, Myron S., 1972, The Market for Securities: Substitution Versus Price Pressure and the Effects of Information on Share Prices, *Journal of Business* 45, 179–211.

Shleifer, Andrei, 1986, Do Demand Curves for Stocks Slope Down?, *Journal of Finance* 41, 579–590.

Figure 1: Times Series Trend of Market-on-close Trading Volumes

This figure shows the time series trend of the market-on-close trading volume from July 2008 to June 2018 for S&P 500 stocks (Panel A) and for all NYSE/NASDAQ/Amex stocks (Panel B). The market-on-close trading volume is computed as the trading volume in TAQ dataset with the sale condition of '6' (or 'M'), scaled by the total daily trading volume.

Panel A: Market-on-close trading volume as percentage of daily trading volume for stocks in SP500



Panel B: Market-on-close trading volume as percentage of daily trading volume for all stocks

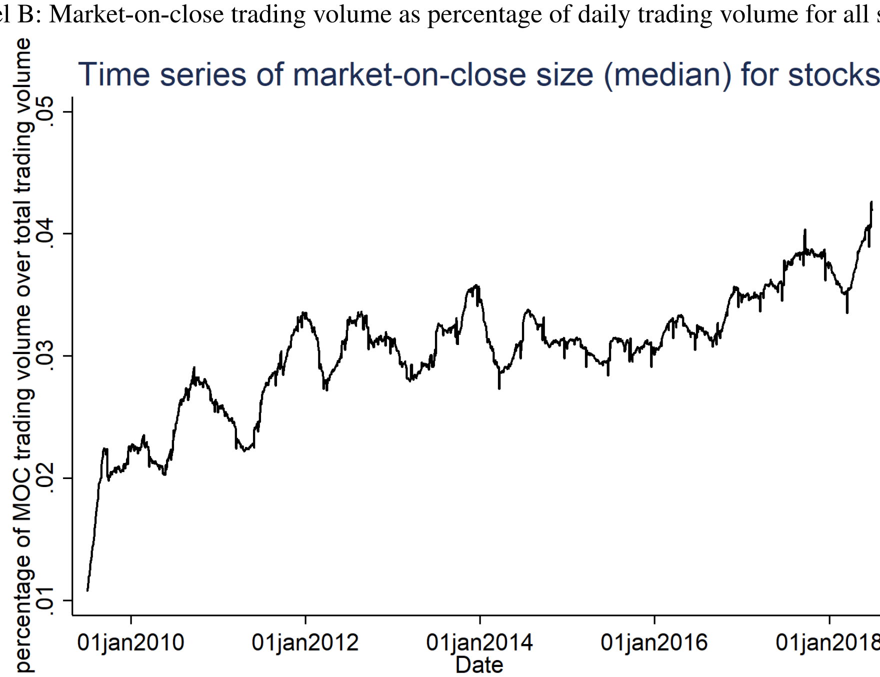


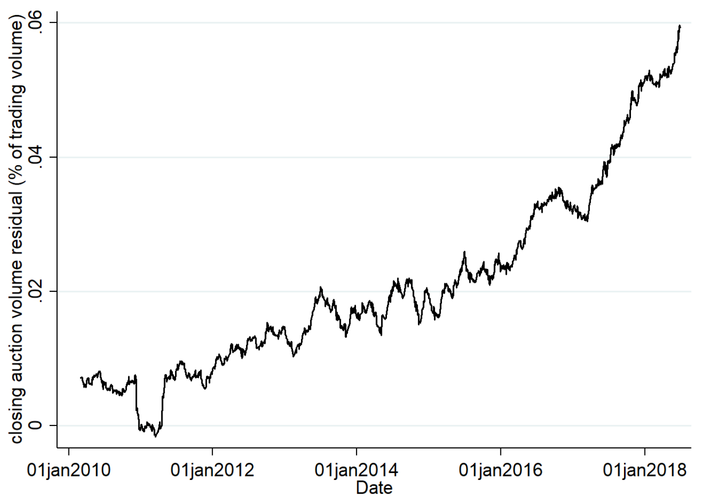
Figure 2: Times Series Trend of Market-on-close Trading Volumes not Originated from ETF Flows

This figure shows the time series trend of the residual of market-on-close trading volume that is not originated from ETF flows from January 2010 to June 2018 for S&P 500 stocks (Panel A) and for all NYSE/NASDAQ/Amex stocks (Panel B). The residuals of market-on-close trading volume is defined as follows.

$$Volume\ residual_{i,t} = \frac{\$Volume_{moc,i,t} - \$Volume_{ETF\ flows,i,t}}{\$Volume_{total\ tradingi,t}}$$

where $\$Volume_{moc,i,t}$ represents the total dollar trading volume of the market-on-close orders, and $\$Volume_{ETF\ flows,i,t}$ represents the total aggregate ETF flows in dollar volume. The difference between these two is then normalized by the total daily trading volume.

Panel A: market-on-close trading volume as percentage of daily trading volume for stocks in SP500



Panel B: market-on-close trading volume as percentage of daily trading volume for all stocks

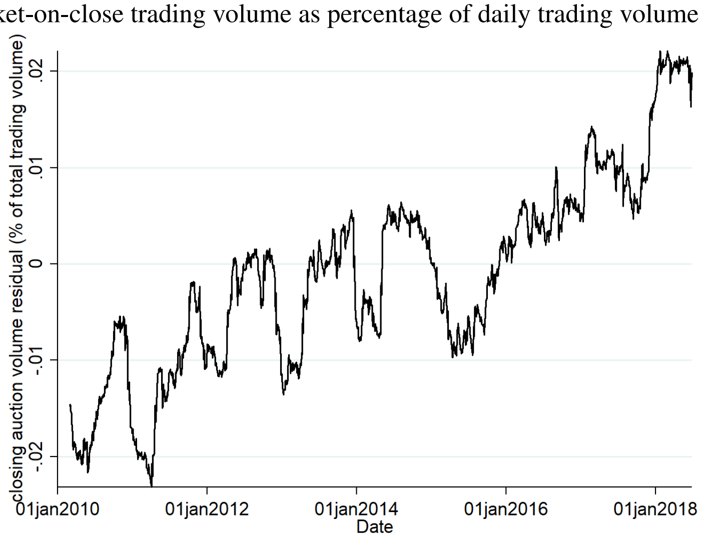


Table I: Summary Statistics

The table reports the summary statistics for the variables used in the study. The samples range between July 2008 and June 2018. Detailed definition of variables are available in appendix A.I.

Variables:	N	Mean	Std Dev	p10	Median	p90
Close-to-close return	6,663,021	0.001	0.026	-0.026	0.000	0.028
Open-to-close return	6,663,019	0.001	0.026	-0.024	0.000	0.025
Overnight return	6,621,943	0.000	0.011	-0.011	0.000	0.011
Last15ret	6,234,519	0.000	0.006	-0.005	0.000	0.005
Close-to-close return (quote)	6,662,918	0.001	0.052	-0.026	0.000	0.027
Daily Trading Volume (shares million)	6,663,021	1.442	5.986	0.015	0.292	3.090
Market-on-Close Volume (shares million)	6,663,021	0.062	0.329	0.000	0.010	0.118
Market-on-Close Volume (%)	6,663,021	6.40	7.90	0.80	4.10	13.7
Turnover	6,663,021	8.804	9.988	1.225	5.852	18.95
Order Imbalance	6,015,804	-0.001	0.019	-0.017	0.000	0.015
Amihud	6,662,618	0.598	12.168	0.000	0.002	0.088
1/Price	6,662,859	0.058	0.049	0.014	0.043	0.13
Log (Market Cap)	6,663,021	13.937	1.755	11.777	13.818	16.305
Flow-related variables:						
Total Flows ($ million)	6,663,021	0.146	4.564	-0.644	0.003	0.980
Total abs flows	6,663,021	1.566	6.027	0.004	0.258	3.141
Total flows/Daily $ Volume	6,663,021	0.045	93.285	-0.065	0.001	0.082
Total abs flows/ Daily $ Volume	6,663,021	0.141	132.495	0.003	0.035	0.181

Table II: ETF flows and Market-on-close Volume

Panel A of this table reports estimates from panel regressions of daily market-on-close trading volume on aggregated ETF flows and controls. The sample consists of all common stocks (*shrcd* in 10 or 11, excluding those are priced less than $5) from July 2008 to June 2018. In Panel B, the independent variable is based on the alternative definition of aggregated *ETF flows*. Panel C presents the same regressions but with the sample split by different periods and different stock exchanges. The frequency of the observations is daily. The dependent variable, market-on-close trading volume, is computed as the trading volume in the TAQ dataset with the sale condition of '6'(or 'M), scaled by the total daily trading volume. The aggregated ETF flows used in Panel A is defined in the Equation 1 while the alternative ETF flows used in Panel B is defined as in Equation 3. The controls in all panels include order imbalance, turnover, daily return, the lagged return, the lagged order imbalance, the lagged turnover, the lagged logged market capitalization, and the lagged Amihud (2002) ratio. All variables except return variables have been standardized each day by subtracting the mean and dividing by the standard deviation. Standard errors are double clustered at the stock and day level. t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A

	Dep.Variable: market-on-close trading volume			
	(1)	(2)	(3)	(4)
ETF flows: aggregate of absolute $_t$	0.270***	0.255***	0.302***	0.254***
	(37.38)	(35.56)	(42.85)	(35.23)
Order imbalance$_t$		0.00717***	0.00711***	0.00780***
		(4.84)	(4.48)	(5.29)
Turnover$_t$		-0.0745***	-0.0999***	-0.0749***
		(-20.67)	(-27.04)	(-20.78)
Return$_t$		-0.119***	-0.135***	-0.227***
		(-3.63)	(-2.76)	(-5.35)
Return$_{t-1}$		0.00207	0.0867*	0.000727
		(0.07)	(1.94)	(0.02)
Order imbalance$_{t-1}$		-0.00337***	-0.00337***	-0.00307***
		(-3.48)	(-2.95)	(-3.21)
Turnover$_{t-1}$		-0.0480***	-0.0805***	-0.0482***
		(-21.79)	(-28.90)	(-21.94)
Log (Market Cap)$_{t-1}$		0.0167***	-0.00564*	-0.0256***
		(3.34)	(-1.86)	(-4.64)
Amihud$_{t-1}$		0.00356*	0.00838*	0.00354*
		(1.82)	(1.87)	(1.82)
Firm Fixed Effect	Yes	Yes	No	Yes
Day Fixed Effect	Yes	No	Yes	Yes
N	6,662,992	5,969,882	5,969,894	5,969,881
Adj-R^2	0.295	0.269	0.178	0.271

Panel B: Alternative Definition of ETF Flows

	Dep.Variable: market-on-close trading volume			
	(1)	(2)	(3)	(4)
ETF flows:absolute of aggregate $_t$	0.245***	0.227***	0.276***	0.225***
	(37.66)	(34.80)	(40.43)	(34.52)
Order imbalance$_t$		0.00711***	0.00779***	0.00782***
		(4.70)	(4.77)	(5.21)
Turnover$_t$		-0.0863***	-0.114***	-0.0865***
		(-23.87)	(-30.66)	(-23.97)
Return$_t$		-0.150***	-0.189***	-0.274***
		(-4.54)	(-3.79)	(-6.40)
Return$_{t-1}$		-0.0225	0.0446	-0.0316
		(-0.74)	(0.99)	(-0.84)
Order imbalance$_{t-1}$		-0.00341***	-0.00267**	-0.00306***
		(-3.50)	(-2.31)	(-3.18)
Turnover$_{t-1}$		-0.0489***	-0.0827***	-0.0491***
		(-22.11)	(-29.31)	(-22.28)
Log (Market Cap)$_{t-1}$		0.0195***	-0.00386	-0.0265***
		(3.86)	(-1.27)	(-4.72)
Amihud$_{t-1}$		0.00349*	0.00799*	0.00347*
		(1.81)	(1.86)	(1.81)
Firm Fixed Effect	Yes	Yes	No	Yes
Day Fixed Effect	Yes	No	Yes	Yes
N	6,662,992	5,969,882	5,969,894	5,969,881
Adj-R^2	0.287	0.259	0.162	0.261

Panel C: Subsamples

	Dep. Variable: market-on-close trading volume				
	July/14-June/18	July/08-June/14	NYSE	NASDAQ	American
ETF flows: aggregate of absolute $_t$	0.248***	0.250***	0.185***	0.276***	0.263***
	(25.07)	(30.05)	(16.18)	(30.24)	(9.71)
Control Variables	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes	Yes
Day Fixed Effect	Yes	Yes	Yes	Yes	Yes
N	2,785,245	3,184,631	2,659,709	3,191,285	118,248
Adj-R^2	0.273	0.308	0.231	0.304	0.309

Table III: ETF flows and Market-on-close Trading Volume: Instrumental Variable Regressions

This table reports estimates from a quasi-natural experiment relying on the reconstitution of the Russell 1000 and Russell 2000 indexes. The sample is at the stock-day stock level. Panel A presents the first-stage regression where the dependent variable is ETF flows. The explanatory variables (instruments) are a dummy for inclusion in the Russell 1000, for stocks that are in the Russell 2000 and potential switchers around the upper bound threshold before index reconstitution. To determine the upper bound threshold, stocks are ranked by market cap by the end of May each year and the upper bound threshold is calculated as the market cap of stock ranked in 1000 plus 0.25 times the cumulative market cap of Russell 3000E stocks. Column 1 and 2 present bandwidths ranging from 100 to 200 stocks. The stocks enter the sample in the June after index reconstitution and remain in the sample until the following May (excluded). The controls include order imbalance, turnover, daily return, the lagged return, the lagged order imbalance, the lagged turnover, the lagged logged market capitalization, and the lagged Amihud (2002) ratio. Samples for Column 3 and 4 exclude the daily observations in the months of June. In Panel B, the dependent variable is market-on-close trading volume, defined as the trading volume in the TAQ dataset with sale condition of '6'(or 'M), scaled by the total daily trading volume. The main explanatory variable is instrumented ETF flow with switching indicators as instruments. All variables except return variables have been standardized each day by subtracting the mean and dividing by the standard deviation. Standard errors are double clustered at the stock and day level. t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: First Stage Regression

	Dep.Variable: ETF flows			
	(1)	(2)	(3)	(4)
	±100	±200	±100 exclude June	±200 exclude June
Switch Indicator	-0.172***	-0.155***	-0.186***	-0.169***
	(-3.23)	(-4.39)	(-3.42)	(-4.68)
Order imbalance$_{t-1}$	0.013	0.012	0.014	0.014*
	(1.28)	(1.62)	(1.27)	(1.71)
Turnover$_{t-1}$	-0.121***	-0.135***	-0.120***	-0.134***
	(-10.38)	(-12.61)	(-10.10)	(-12.47)
Log (Market Cap)$_{t-1}$	0.155**	0.088	0.152**	0.089
	(2.13)	(1.35)	(2.11)	(1.36)
1/Price$_{t-1}$	4.891**	1.394	4.784**	1.398
	(2.03)	(0.80)	(2.03)	(0.81)
Return$_{t-1}$	-0.435**	-0.433***	-0.471***	-0.439***
	(-2.31)	(-2.43)	(-2.69)	(-3.76)
Return$_t$	-0.415***	-0.443***	-0.398***	-0.439***
	(-2.62)	(-3.64)	(-2.69)	(-3.76)
Amihud$_t$	68.143***	78.272***	65.722***	76.573***
	(5.58)	(7.96)	(5.61)	(8.01)
Day Fixed Effect	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes
N	55,337	110,991	49,926	100,314
Adj-R^2	0.279	0.248	0.282	0.249
F-stat	367.140	206.249	379.414	220.635

Panel B: Second Stage Regression

	Dep.Variable: market-on-close trading volume			
	(1)	(2)	(3)	(4)
	±100	±200	±100 exclude June	±200 exclude June
ETF flows(instrumented)	0.486**	0.341**	0.526***	0.389***
	(2.16)	(2.04)	(2.42)	(2.47)
Order imbalance$_{t-1}$	-0.007	-0.011	-0.001	-0.011
	(-0.64)	(-1.33)	(-0.12)	(-1.30)
Turnover$_{t-1}$	-0.080***	-0.102***	-0.075***	-0.095***
	(-3.14)	(-4.59)	(-3.12)	(-4.59)
Log(Market Cap)$_{t-1}$	0.039	0.049**	0.031	0.041*
	(0.95)	(2.07)	(0.76)	(1.78)
1/Price$_{t-1}$	-0.949	-0.712	-1.220	-0.894*
	(-0.83)	(-1.42)	(-1.08)	(-1.86)
Return$_{t-1}$	0.524**	0.427**	0.650***	0.535***
	(2.22)	(2.05)	(2.58)	(3.07)
Return$_t$	0.093	-0.040	0.156	-0.028
	(0.55)	(-0.30)	(0.97)	(-0.21)
Amihud $_t$	12.605	17.713	10.488	14.458
	(0.82)	(1.27)	(0.73)	(1.14)
Daily time Fixed Effect	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes
N	55,337	110,991	49,926	100,314
Adj-R^2	0.178	0.216	0.160	0.207

Table IV: Price Movement and Closing Auction Imbalance: Evidence from NYSE Closing Auction Feed Data

Panel A of this table reports Fama-Macbeth regression of 15-minute return (in percentage) from 15:45 to 16:00 on market-on-close trading volume measure and controls, using NYSE closing auction feed data from March 2010 to June 2018. *highmoc_buy* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Buy', where *moc_size_nyse* is the total imbalance quantity from the first disseminated feed from the NYSE closing auction data scaled by the total trading volume per stock. *highmoc_sell* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Sell'. Column 1 reports baseline regression result, Column 2 reports regression result with all controls for whole sample period. Column 3 and 4 report subsample periods with March 2010 to May 2014 for Column 3 and June 2014 to June 2018 for Column 4. Panel B reports the Fama-MacBeth regression of 15-minute return from 15:15 to 15:30 (Column 1), or 15:30 to 15:45 (Column 2) on the closing auction measure *highmoc_buy* and *highmoc_sell*, as placebo tests. The controls in regressions include order imbalance and turnover, defined in the appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Using All NYSE Feed Data

Dependent variable:	Return from 15:45 to 16:00			
	(1)	(2)	(3)	(4)
Highmoc_buy	0.101***	0.090***	0.121***	0.060***
	(27.93)	(27.77)	(25.81)	(19.12)
Highmoc_sell	-0.108***	-0.100***	-0.115***	-0.087***
	(-9.16)	(-9.73)	(-26.53)	(-4.48)
Turnover		-0.001	-0.001	-0.002
		(-1.43)	(-1.46)	(-1.29)
Order imbalance		5.965***	4.058***	7.734***
		(4.52)	(27.39)	(3.05)
Intercept	0.006	0.012	0.000	0.023
	(0.54)	(0.91)	(0.07)	(0.92)
Average Adj-R^2	0.033	0.068	0.078	0.059

Panel B: Placebo Test Using Returns from Other Time Periods

Dependent variable:	*Ret*_15 : 15 to 15 : 30	Ret_15 : 30 to 15 : 45
	(1)	(2)
Highmoc_buy	0.005***	0.012***
	(5.03)	(10.67)
Highmoc_sell	-0.002**	-0.014***
	(-2.27)	(-10.83)
Turnover	0.0001	0.000*
	(1.14)	(2.15)
Order imbalance	1.024***	1.705***
	(12.48)	(17.83)
Intercept	0.003	0.013***
	(1.04)	(4.09)
Average Adj-R^2	0.013	0.015

Table V: The Economic Magnitude of Price Impact: Counterfactual Analysis

This table contrasts the actual estimates and four alternative estimates of price impact during 15 minute interval from 15:45 to 16:00 for stocks listed in NYSE from March 2010 to June 2018. The four alternative measures of trading costs are: the Breen, Hodrick, and Korajczyk (2002) (BHK) specification, the Glosten and Harris (1988) (GH) specification, the effective spread, and the quoted spread. The first two measures are nonproportional to trade sizes while the later two are independent of the trading size. The costs are estimated out of samples per stock per month following the approach in Korajczyk and Sadka (2004): Korajczyk and Sadka (2004) first use intraday transactions data over the period January 1993 to May 1997 to estimate market impact costs. Then based on the in-sample cross-sectional relationships between these market impact costs and firm characteristics, the costs can then be estimated out of samples. Appendix A provides the detailed procedure for each estimate. Stocks with high buying (selling) market-on-close order imbalance are stocks that are in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the market-on-close orders at the beginning of dissemination is 'Buy'('Sell'), where *moc_size_nyse* is the total imbalance quantity from the first disseminated feed from the NYSE closing auction data scaled by the total trading volume per stock. Standard errors are presented in parentheses.

	Actual estimate	BHK estimate	GH estimate	Effective spread	Quoted spread	Diff between 1 and 2	Diff between 1 and 3
A. Stocks with high buying	11.243	0.059	2.753	0.325	0.992	11.013	8.319
closing auction order imbalance	(0.030)	(0.000)	(0.072)	(0.000)	(0.000)	(0.030)	(0.050)
B. Stocks with high selling	11.457	0.062	2.666	0.317	0.971	11.257	8.654
closing auction order imbalance	(0.045)	(0.000)	(0.087)	(0.000)	(0.001)	(0.047)	(0.066)

Table VI: Closing Auction and Price Reversal

This table reports estimates from Fama-Macbeth regression of following model:

$$Return_{i,t+1} = \alpha + \beta_1\ Return_{i,t} + \beta_2\ Return_{i,t} \times highmoc_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t}$$

The sample consists of all common stocks (*shrcd* in 10 or 11, excluding those are priced less than $5) with non-zero closing auction volume in NYSE stock exchange from July 2008 to June 2018. Return variable used in the independent variable and the interaction term is the close to close daily return. The dependent variables include close-to-close return (columns 1 and 2), intraday return (column 3), and overnight return (column 4), defined in Equation 7. *highmoc* is an indicator as to whether the stock is in the top quintile of *Market-on-close Trading Volume* for the day, where *Market-on-close Trading Volume* is defined as the total market-on-close volume in the TAQ dataset scaled by the total trading volume per stock. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Return in Independent Variables: $Ret_{close\_to\_close,t}$

Dependent variable:	$Ret_{close\_to\_close,t+1}$		$Ret_{Intraday,t+1}$	$Ret_{overnight,t+1}$
	(1)	(2)	(3)	(4)
Return_t	-0.004 (-1.19)	-0.004 (-1.43)	-0.002 (-0.74)	-0.003*** (-2.69)
Highmoc # Return$_t$	-0.022*** (-7.06)	-0.019*** (-6.31)	-0.011*** (-4.10)	-0.007*** (-7.64)
Highmoc_t	-0.008 (-1.14)	-0.003 (-0.50)	0.003 (0.68)	-0.006*** (-3.56)
Turnover_t		0.000 (0.62)	-0.001*** (-2.61)	0.001*** (7.02)
Order imbalance_t		-1.68*** (-9.72)	-0.535*** (-3.37)	-1.19*** (-20.35)
Amihud_t		-0.103* (-1.70)	0.013 (0.21)	-0.060** (-2.44)
log(Market $\text{Cap})_t$		-0.003 (-0.93)	-0.002 (-0.86)	0.000 (-0.17)
$1/\text{Price}_t$		-0.163* (-1.65)	-0.306*** (-3.48)	0.098*** (3.27)
Intercept	0.069*** (2.96)	0.109** (1.96)	0.091* (1.77)	0.018 (0.88)
Average Adj-R^2	0.02	0.056	0.055	0.055

Table VII: Closing Auction and Price Reversal: Evidence from NYSE Closing Auction Feed Data

This table reports estimates from Fama-Macbeth regression of the following model:

$$Return_{i,t+1} = \alpha + \beta_1\ Return_{i,t} + \beta_2\ highmoc\_buy_{i,t} \times Return_{i,t} + \beta_3\ highmoc\_sell_{i,t} \times Return_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t}$$

The sample consists of all common stocks (*shrcd* in 10 or 11, excluding those are priced less than \$5) from NYSE Closing auction feed data for the period July 2008 to June 2018. Return variable used in the independent variable and the interaction term is the close to close daily return. The dependent variables include close-to-close return (columns 1 and 2), intraday return (column 3), and overnight return (column 4), defined in Equation 7. *highmoc_buy* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Buy', where *moc_size_nyse* is the total imbalance quantity from the first disseminated feed from the NYSE closing auction data scaled by the total trading volume per stock. *highmoc_sell* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Sell'. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Dependent variable:	$Ret_{close\_to\_close,t+1}$		$Ret_{Intraday,t+1}$	$Ret_{overnight,t+1}$
	(1)	(2)	(3)	(4)
Return	0.000	0.000	0.000	-0.001
	(0.13)	(-0.08)	(0.09)	(-0.34)
Highmoc_buy # Return	-0.033***	-0.026***	-0.018***	-0.008***
	(-7.05)	(-5.54)	(-4.29)	(-3.87)
Highmoc_sell # Return	-0.024***	-0.019**	-0.009	-0.009***
	(-3.23)	(-2.52)	(-1.28)	(-3.45)
Highmoc_buy	-0.059***	-0.033**	-0.016	-0.017***
	(-3.88)	(-2.27)	(-1.35)	(-3.38)
Highmoc_sell	0.032**	0.041***	0.025*	0.017***
	(2.21)	(3.08)	(1.82)	(4.11)
Turnover		0.000	-0.001***	0.001***
		(-0.10)	(-2.83)	(5.01)
Order Imbalance		-1.31***	-0.439**	-0.824***
		(-6.31)	(-2.21)	(-8.02)
Amihud		-0.566	0.405	-0.989
		(-1.09)	(1.14)	(-1.07)
log(Market Cap)		-0.006*	-0.005	-0.001
		(-1.86)	(-1.59)	(-0.88)
1/Price$_t$		-0.073	-0.282*	0.203***
		(-0.52)	(-2.08)	(3.95)
Intercept	0.059**	0.145**	0.125**	0.020
	(2.10)	(2.30)	(2.14)	(0.80)
Average Adj-R^2	0.020	0.052	0.053	0.052

Table VIII: Long/Short Trading Strategies

This table reports raw returns and risk-adjusted returns for daily equal weighted long/short trading strategies formed based on signal from closing auction volume. The sample consists of all common stocks with nonzero closing auction volume from July 2008 to June 2018. The trading strategy for Panel A is to form portfolio at the end of each day by buying (selling) stocks that are within highest quintile of closing auction total imbalance quantiy and with imbalance side of 'Sell' ('Buy') based on the NYSE imbalance feed information at 15:45. The holding period starts from the end of day until the end of the next day. The trading strategy for Panel B is to form portfolio at the end of each day by buying (selling) stocks that are within the highest quintile of closing auction volume and also have a negative (positive) daily return. The holding period starts from the end of the day until the end of the next day. Daily Fama-French risk factors are used to calculate risk-adjusted alphas. Panel C reports the average of raw returns from these two trading strategies for different holding horizons. The holding horizon starts from the same day's close to the next morning's open (i.e., *Overnight*) in columns 1 and 3 and starts from next morning's open to the next day's close (i.e., *Open-to-Close*) in columns 2 and 4. All standard errors are adjusted for autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Long-Short Strategy Using NYSE Feed Data

	Alpha(bp)	MktRf	SMB	HML	UMD
Long	6.063***	1.106***			
	(4.78)	(49.46)			
	6.109***	1.022***	0.531***	0.192***	
	(5.58)	(36.29)	(8.31)	(4.75)	
	6.071***	1.011***	0.533***	0.135***	-0.076***
	(6.23)	(37.29)	(7.33)	(5.86)	(-5.59)
Short	-7.365***	1.086***			
	(-6.16)	(38.36)			
	-7.214***	0.979***	0.607***	0.266***	
	(-6.96)	(47.53)	(8.56)	(5.62)	
	-7.233***	0.973***	0.609***	0.237***	-0.039**
	(-6.96)	(46.13)	(8.57)	(4.91)	(-2.20)
Long-Short	13.291***	0.020			
	(10.11)	(0.87)			
	13.218***	0.044**	-0.074**	-0.077	
	(9.81)	(2.10)	(-1.98)	(-1.29)	
	13.198***	0.038*	-0.073**	-0.106*	-0.038
	(9.91)	(1.85)	(-1.97)	(-1.80)	(-1.23)

Panel B: Long-Short Strategy Using TAQ data

	Alpha(bp)	MktRf	SMB	HML	UMD
Long	2.652**	1.139***			
	(2.04)	(45.23)			
	2.709**	1.038***	0.680***	0.218***	
	(2.42)	(46.17)	(13.64)	(6.75)	
	2.686**	1.031***	0.681***	0.183***	-0.047**
	(2.42)	(47.32)	(13.73)	(5.50)	(-2.40)
Short	-4.794***	1.048***			
	(-3.71)	(39.70)			
	-4.655	0.929***	0.688***	0.292***	
	(-1.47)	(44.78)	(14.22)	(9.94)	
	-4.697***	0.919***	0.691***	0.239***	-0.072***
	(-4.40)	(45.51)	(14.31)	(7.34)	(-3.57)
Long-Short	7.442***	0.090***			
	(5.17)	(5.02)			
	7.356***	0.108***	-0.009	-0.073**	
	(5.12)	(5.74)	(-0.24)	(-2.29)	
	7.370***	0.111***	-0.010	-0.056	0.023
	(5.13)	(5.86)	(-0.26)	(-1.63)	(1.04)

Panel C: Long-Short Strategy: Different holding horizons

	TAQ data		NYSE Imbalance Feed Data	
Holding Period:	Overnight	Open-to-Close	Overnight	Open-to-Close
Time series average return (bp)	3.77***	4.02***	5.55***	7.84***
	(3.28)	(6.97)	(13.83)	(7.38)

Table IX: Closing Auction, ETF Ownership, and Volatility of Underlying Securities

This table reports estimates from ordinary least squares (OLS) regressions of different measures of daily volatility on ETF ownership and controls. The sample consists of S&P 500 stocks and ranges from June 2010 to May 2018. The frequency of the observations is monthly. Volatility in columns 1 and 2 is computed using all daily returns within the month, while volatility in columns 3 and 4 is computed using all daily alternative return, computed based on open price and 15:45 price. The controls include lagged book-to-market ratio, lagged Amihud (2002) ratio, lagged inverse share price, and logged marekt capitalization. The dependent variable and the ownership variable have been standardized by subtracting the mean and dividing by the standard deviation. Standard errors are double clustered at the stock and month levels. t-statistics are presented in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Volatility (close to close)		Volatility (open to 15:45)	
	(1)	(2)	(3)	(4)
ETF Ownership	0.105***	0.0662**	0.0603***	0.0262*
	(2.63)	(2.26)	(2.68)	(1.82)
1/Price		4.072***		1.484**
		(6.11)		(2.42)
Amihud		0.196***		-0.0829***
		(5.92)		(-2.91)
B/M		158.6**		104.6**
		(2.09)		(2.27)
log(market cap)		-0.238***		-0.306***
		(-4.21)		(-5.70)
Constant	0.00864***	3.754***	0.00648***	4.985***
	(319.11)	(3.92)	(88.70)	(5.45)
Firm Fixed Effect	Yes	Yes	Yes	Yes
Month Fixed Effect	Yes	Yes	Yes	Yes
N	56,815	52,509	56,815	52,509
Adj-R^2	0.652	0.698	0.748	0.775

Table X: Do Informed Traders Use Closing Auction Orders? Evidence from Price Reversal

This table reports estimates from Fama-Macbeth regression of the following model:

$$Return_{i,t+1} = \alpha + \beta_1\ Return_{i,t} + \beta_2\ highmoc\_buy_{i,t} \times Return_{i,t} + \beta_3\ highmoc\_sell_{i,t} \times Return_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t}$$

across different subsamples. The sample consists of all common stocks (*shrcd* in 10 or 11, excluding those are priced less than $5) from NYSE Closing auction feed data for the period July 2008 to June 2018. Return variable used in the independent variable, dependent variable, and the interaction term is the close to close daily return. *highmoc_buy* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Buy', where *moc_size_nyse* is the total imbalance quantity from the first disseminated feed from the NYSE closing auction data scaled by the total trading volume per stock. *highmoc_sell* is an indicator as to whether the stock is in the top quintile of daily *moc_size_nyse* cross-sectionally across all stocks listed in NYSE and the imbalance side of the closing auction at the beginning of dissemination is 'Sell'. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. The sample from column 1 consists of all stocks while the sample from column 2 (3) consists of stocks which the closing auction orders are less (more) likely to be placed by informed traders. A stock within high quintile of closing auction volume residual, defined in Equation 10 and with a direction of "Sell" ("Buy") of closing auction but with a positive (negative) aggregate ETF inflows is classified as more likely to be stocks traded by informed traders using the closing auction orders. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
Return	0.001	0.002	-0.005
	(0.57)	(0.63)	(-0.86)
Highmoc_buy # Return	-0.029***	-0.034***	-0.026
	(-6.92)	(-7.37)	(-1.47)
Highmoc_sell # Return	-0.021***	-0.022***	-0.012
	(-4.57)	(-3.93)	(-1.28)
Highmoc_buy	-0.052***	-0.055***	-0.031
	(-4.27)	(-4.75)	(-1.47)
Highmoc_sell	0.059***	0.063***	0.027*
	(6.11)	(6.65)	(1.84)
Turnover	0.000	0.000	0.003***
	(-0.11)	(-0.39)	(2.67)
Order Imbalance	-1.31***	-1.24***	-1.53**
	(-7.25)	(-6.68)	(-2.45)
Amihud	0.386	-0.074	-0.778
	(0.56)	(-0.25)	(-0.31)
log(Market Cap)	-0.007**	-0.008***	0.001
	(-2.45)	(-2.83)	(0.27)
1/Price$_t$	-0.147	-0.166	0.240
	(-1.17)	(-1.34)	(0.97)
Intercept	0.152***	0.174***	0.020
	(2.84)	(3.20)	(0.28)
Average Adj-R^2	0.043	0.043	0.107

Table XI: Do Informed traders Use Closing Auction Orders? Evidence from Earning Announcement

This table reports non-ETF-flow-driven abnormal market-on-close trading volumes (Panel A) and abnormal market-on-close trading imbalances (Panel B) around the quarterly earnings announcement dates. Stocks in CRSP are matched to quarterly earnings announcements in I/B/E/S from January 2010 to June 2018. In event time, day 0 is the day of the announcement. The earnings surprise (SUE) for an announcement is the difference between actual earnings for the quarter recorded by I/B/E/S and the mean analyst forecast one month before the quarterly earnings scaled by the stock price on the forecast date. The cumulative abnormal return (*CAR[-1,1]*) for each stock is the buy-and-hold return from one day before earnings announcement to one day afterward, adjusted for the Daniel, Grinblatt, Titman, and Wermers (1997) characteristics benchmark. Abnormal market-on-close trading volume that is not from passive ETF flows is the average log *volume residual* on the day of and the day after the announcement, divided by the average log *volume residual* for the period -14 to -5 in event time (10 trading days), where *volume residual* is defined in Equation 10. Abnormal market-on-close trading imbalance that is not from passive ETF flows is defined as the average log *trade imbalance residual* on the day of and the day after the announcement, divided by the average log *trade imbalance residual* for the period -14 to -5 in event time (10 trading days), where *trade imbalance residual* is defined in Equation 12. In each calendar quarter, quarterly earnings announcements during that quarter are sorted into quintiles based on absolute value of SUE (Panel A, columns 3-5), absolute value of CAR (Panel A, columns 8-10), relative value of SUE (Panel B, columns 3-5), relative value of CAR (Panel B, columns 8-10). The sample average of abnormal trading volumes and abnormal trading imbalances are reported with t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Abnormal Trading Volumes

		Abnormal volume			Obs		Abnormal volume			Obs
Horizon:		[-4, -2]	[-1, 1]	[2, 4]			[-4, -2]	[-1, 1]	[2, 4]	
$\|SUE\|$	1	0.0004	0.0043	0.0085	17124	$\|CAR\|$	0.0016	0.0070	0.0113	11590
		(1.01)	(12.54)	(16.54)			(2.80)	(13.86)	(14.87)	
	2	0.0003	0.0042	0.0089	17093		0.0025	0.0087	0.0122	11590
		(0.74)	(11.50)	(17.01)			(4.67)	(17.86)	(16.99)	
	3	0.0014	0.0066	0.0105	16880		0.0027	0.0081	0.0102	11621
		(3.28)	(17.27)	(19.54)			(5.27)	(17.65)	(14.66)	
	4	0.0017	0.0090	0.0132	16551		0.0030	0.0100	0.0125	11585
		(3.90)	(22.96)	(23.89)			(6.25)	(23.51)	(20.86)	
	5	0.0032	0.0095	0.0114	15715		0.0032	0.0111	0.0147	11490
		(7.25)	(24.46)	(20.07)			(6.92)	(28.41)	(29.08)	
	Diff 5-1	0.0028	0.0052	0.0029			0.0034	0.0041	0.0016	
		(4.84)	(10.09)	(3.77)			(3.73)	(6.39)	(2.19)	

Panel B: Abnormal Trading Imbalances

		Abnormal trading imbalances			Obs		Abnormal trading imbalances			Obs
		[-4, -2]	[-1, 1]	[2, 4]			[-4, -2]	[-1, 1]	[2, 4]	
SUE	1	0.0043	0.0042	0.0047	8252	CAR	0.0023	0.0079	0.0141	5519
		(4.63)	(5.73)	(4.20)			(2.21)	(10.14)	(13.15)	
	2	-0.0001	0.0032	0.0030	8940		0.0023	0.0077	0.0071	5605
		(-0.08)	(4.50)	(2.79)			(2.12)	(8.50)	(5.02)	
	3	0.0006	0.0025	0.0020	8876		0.0015	0.0044	0.0058	5596
		(0.71)	(3.95)	(2.06)			(1.33)	(4.69)	(3.46)	
	4	0.0005	0.0035	0.0026	8866		0.0001	0.0009	-0.0016	5620
		(0.64)	(5.47)	(2.67)			(0.06)	(1.03)	(-1.28)	
	5	0.0010	0.0034	0.0013	8449		0.0023	0.0019	-0.0043	5551
		(1.21)	(4.94)	(1.34)			(2.26)	(2.46)	(-3.56)	
	Diff 5-1	-0.0032	-0.0008	-0.0034			0.0000	-0.0060	-0.0184	
		(-2.59)	(-0.78)	(-2.31)			(0.00)	(-5.43)	(-11.41)	

Appendix

Figure A.1 shows one example of Authorized Participant handbook and agreement.

Figure A.2 illustrates two scenarios of creation and redemption.

Table A.I defines variables used in this paper.

Appendix A describes the procedure to estimate alternative price impacts.

Table A.II shows the results for the price impact of high market-on-close order imbalances using TAQ data.

Table A.III shows the return reversal pattern using alternative measures of returns.

Table A.IV shows the return reversal controlling for market cap.

Table A.V shows the return reversal using average quoted price for return calculation.

Table A.VI shows subperiod alphas from the long/short trading strategies.

Table A.VII shows alphas from the value-weighted scheme for forming the long/short trading strategies.

Figure A.1: Authorized Participant Handbook and Agreement: An Example

The figure provides one example of authorized participant handbook and agreement. It specifies the procedure how authorized participants can purchase/redeem shares. The outlines in this document requires that the Authorized Person of the authorized participant must make telephone call not later than the closing time of the regular trading session on the listing exchange.

FOR ALPS ETF TRUST

<u>TO PLACE A PURCHASE ORDER FOR</u>
<u>CREATION UNIT(S) OF SHARES OF ONE OR MORE FUNDS OF</u>
ALPS ETF TRUST

1. PLACING A PURCHASE ORDER.

Purchase Orders for Creation Units of Shares of ALPS ETF Trust may be initiated only on days that the Listing Exchange is open for trading ("Business Days"). Purchase Orders may only be made in whole Creation Units of Shares of each Fund.

To begin a Purchase Order, the Authorized Participant ("AP"") must telephone the BNY ETF Administrator at (718) 315-4512 or such other number as the Distributor designates in writing to the AP. This telephone call must be made by an Authorized Person of the AP not later than the closing time of the regular trading session on the Listing Exchange which is ordinarily 4:00 p.m. Eastern Time ("Listing Exchange Closing Time"). Upon verifying the authenticity of the AP (as determined by the use of the appropriate PIN Number), BNY ETF Administrator will request that the AP place the Purchase Order. To do so, the AP must provide the appropriate ticker symbols when referring to each Fund . After the AP has placed the Purchase Order, BNY ETF Administrator will read the Purchase Order back to the AP. The AP then must affirm that the Purchase Order has been taken correctly by BNY ETF Administrator. If the AP affirms that Purchase Order has been taken correctly, BNY ETF Administrator will issue a Confirmation Number to the AP.

PLEASE NOTE: A PURCHASE ORDER REQUEST IS NOT COMPLETE UNTIL THE CONFIRMATION NUMBER IS ISSUED BY BNY ETF ADMINISTRATOR. AN ORDER MAY NOT BE CANCELED BY THE AP AFTER THE CONFIRMATION NUMBER IS ISSUED. INCOMING TELEPHONE CALLS ARE QUEUED AND WILL BE HANDLED IN THE SEQUENCE RECEIVED. CALLS PLACED BEFORE THE LISTING EXCHANGE CLOSING TIME WILL BY PROCESSED EVEN IF THE CALL IS ANSWERED BY BNY ETF ADMINISTRATOR AFTER THE LISTING EXCHANGE CLOSING TIME. ACCORDINGLY, THE AP SHOULD NOT HANG UP AND REDIAL. INCOMING CALLS THAT ARE RECEIVED AFTER THE LISTING EXCHANGE CLOSING TIME WILL NOT BE ANSWERED BY BNY ETF ADMINISTRATOR. ALL TELEPHONE CALLS WILL BE RECORDED.

2. RECEIPT OF TRADE CONFIRMATION.

Subject to the conditions that a properly completed telephone Purchase Order has been placed by the AP (either on its own or its customer's behalf) not later than the Listing Exchange Closing Time, the Transfer Agent will accept the Purchase Order on behalf of Trust

18

and Distributor and will confirm in writing to the AP that its Purchase Order has been accepted by 4:45 p.m. Eastern Standard Time on the Business Day that the Purchase Order is received.

Figure A.2: Examples of creation and redemption process

The figure provides an example of how arbitrage activities from creation and redemption are related to closing auction. The top panel provides an illustrative example when the price of a ETF is below the implied price from basket while the bottom panel provides the opposite scenario when the ETF price is above the implied price from basket. In the first case, ETF authorized participants may buy ETF shares and sell short underlying basket. To clear the transaction, at the end of day, authorized participants deliver ETF shares they purchased from market and receive underlying basket to cover the short positions. The bottom panel shows the example when the ETF price is above the implied NAV.(source: BlackRock)

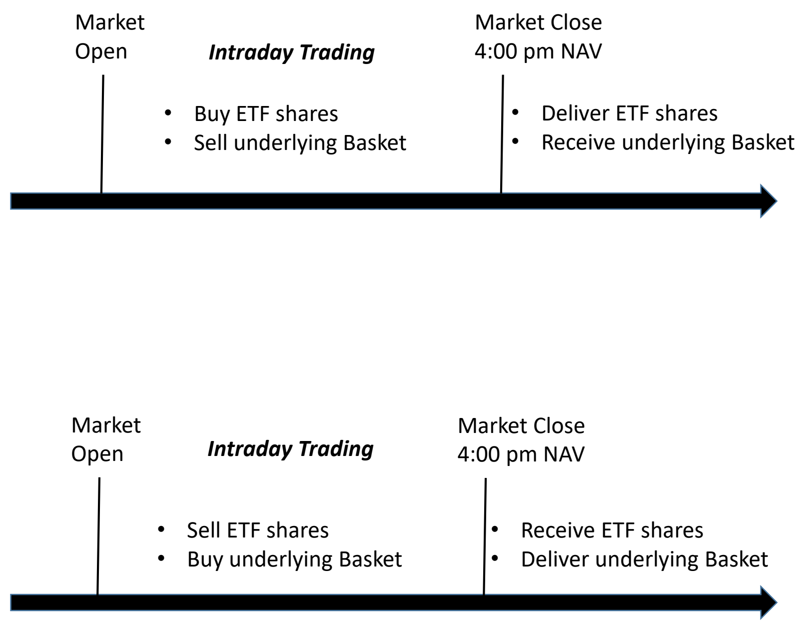


Table A.I: Variable Definition

Variable Names	Variable Definition
MOCVolume	The trade volume from TAQ data with sale condition of "6" or "M", scaled by total daily trading volume in the whole day.
highmoc	A dummy variable equal to one if the stock is in the top quintile of market-on-close volume *MOCVolume* cross-sectionally.
moc_size_nyse	The indicated market-on-close trading volume from NYSE closing auction feed data, scaled by total daily trading volume in the whole day.
highmoc_buy	A dummy variable equal to one if the stock is in the top quintile of indicated market-on-close trading volume *moc_size_nyse* from NYSE imbalance feed data at 15:45:00 and the imbalance side of the closing auction is "Buy"
highmoc_sell	A dummy variable equal to one if the stock is in the top quintile of indicated market-on-close trading volume *moc_size_nyse* from NYSE imbalance feed data at 15:45:00 and the imbalance side of the closing auction is "Sell"
ClosetoCloseReturn	The daily stock return that calculated based on the closing price at $t-1$ and t from CRSP.
IntradayReturn	The return calculated based on the opening price and closing price on day t from CRSP, following Lou, Polk, and Skouras (2018).
OvernightReturn	The return calculated based on equation(7) from CRSP, following Lou, Polk, and Skouras (2018).
$Return_{last15}$	The return calculated based on stock price at 15:45 and closing price, in which stock price at 15:45 is based on the first available valid transaction price that occurred in the time period starting at 15:45:00, after intertwining with quota data to exclude trades outliers.
$Return_{15to30}$	The return calculated based on stock price at 15:15 and price at 15:30, in which both prices are based on the first available valid transaction price that occurred in the time period starting at 15:15:00 and 15:30:00, respectively, after intertwining with quota data to exclude trades outliers.
$Return_{30to45}$	The return calculated based on stock price at 15:30 and price at 15:45, in which both prices are based on the first available valid transaction price that occurred in the time period starting at 15:30:00 and 15:45:00, respectively, after intertwining with quota data to exclude trades outliers.

Variable Names	Variable Definition
Turnover	Share volume divided by total shares outstanding from CRSP
Order imbalance	Total dollar buys minus total dollar sells, then scaled by the daily total share volumes between 9:30 and 16:00 from TAQ data. Trades are classified to buy and sell using a modified algorithm following Holden and Jacobsen (2014)
Log(Market Cap)	The natural log of market capitalization (in millions), in which it is the product of stock price and number of shares outstanding
Amihud Ratio	The monthly average of the absolute daily return divided by the total dollar daily volume in millions, following Amihud(2002)
ETF flows (fund level)	The dollar change in shares outstanding based on the end day NAV at day t and $t-1$ from Bloomberg
ETF Flows (sum of absolute)	Stock-day-level measure that is weighted average of absolute value of dollar *ETF Flows (fund level)* from ETFs holding the stock, scaled by total dollar value of trading volume on day t. The weight is the ETF ownership of the stock per ETF.
ETF Flows (sum of net)	Stock-day-level measure that is absolute value of weighted average of dollar *ETF Flows (fund level)* from ETFs holding the stock, scaled by total dollar value of trading volume on day t. The weight is the ETF ownership of the stock per ETF.

Appendix A Price impact estimates

This section describes the procedure on how to estimate the counterfactual price impact for market-on-close trades, following the approach in Korajczyk and Sadka (2004). Specifically, four alternative measures of trading costs are considered: the Breen, Hodrick, and Korajczyk (2002) specification, the Glosten and Harris (1988) specification, effective spread, and quoted spread. The first two measures are nonproportional to trade sizes while the later two are independent of the trading size. Korajczyk and Sadka (2004) first use intraday transactions data over the period January 1993 to May 1997 to estimate market impact costs. Then based on the in-sample cross-sectional relationships between these market impact costs and firm characteristics, the costs can then be estimated out of samples per stock per month.

Using the results in Table II of Korajczyk and Sadka (2004), I obtain the estimate of buying and selling trading costs:

$$
\begin{aligned}
\lambda_{bhk,i,t} = {} & 2.49 + 0.25 * \text{relative\_mktcap} - 0.61 * \text{volpct} - 0.63 * \text{ret6m} + 0.34 * \text{abs\_ret6m} \\
& - 1.68 * \text{sp500dum} + 4.15 * \text{div\_prc} - 3.42 * \text{r\_squared} - 0.26 + 3.23 * \text{prc\_inverse})/10^5;
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{gh,i,t} = {} & (prc\_inverse) * (0.454 + 0.005 * \text{relative\_mktcap} - 0.069 * \text{volpct} - 0.06 * \text{ret6m} \\
& + 0.25 * \text{abs\_ret6m} - 2.01 * \text{sp500dum} - 0.796 * \text{div\_prc} \\
& - 0.540 * \text{r\_squared} + 0.566 * \text{prc\_inverse})/10^6;
\end{aligned}
$$

$$
\begin{aligned}
\phi_{gh,i,t} = {} & (prc\_inverse) * (26.60 - 0.14 * \text{relative\_mktcap} - 0.71 * \text{volpct} - 54.13 * \text{ret6m} \\
& + 59.83 * \text{abs\_ret6m} - 14.13 * \text{sp500dum} + 19.13 * \text{div\_prc} \\
& - 27.91 * \text{r\_squared} + 121.92 * \text{prc\_inverse})/10^4;
\end{aligned}
$$

$$
\begin{aligned}
effective\_pread_{i,t} = {} & (35.47 - 0.03 * \text{relative\_mktcap} - 1.93 * \text{volpct} - 57.00 * \text{ret6m} + 70.74 * \text{abs\_ret6m} \\
& - 18.44 * \text{sp500dum} - 51.56 * \text{div\_prc} - 36.08 * \text{r\_squared} + 168.64 * \text{prc\_inverse})/10^4;
\end{aligned}
$$

$$
\begin{aligned}
quoted\_spread_{i,t} = {} & (108.29 - 0.07 * \text{relative\_mktcap} - 6.01 * \text{volpct} - 158.17 * \text{ret6m} + 192.43 * \text{abs\_ret6m} \\
& - 48.09 * \text{sp500dum} - 315.34 * \text{div\_prc} - 86.14 * \text{r\_squared} + 363.96 * \text{prc\_inverse})/10^4;
\end{aligned}
$$

where $relative\_mktcap$ is the market cap at the end of last month divided by the average market cap of CRSP, minus one, $volpct$ is the total volume during the last three months divided by the average firm volume on NYSE, minus one; $ret6m$ is stock price at the end of last month divided by the price six month prior, minus one; abs_ret6m is the absolute value of $ret6m$; $sp500dum$ is the dummy variable that equals to 1 if the firm is included in the S&P 500 index; div_prc is the dividend yield and $r\_squared$ is R2 of monthly returns regressed on NYSE index over the last 36 months (required at least 24 observation); $prc\_inverse$ is the inverse of the stock price in the previous month. Since only consider NYSE stocks, NYSE dummy variable are set to 1.

The total cost of a purchasing of q units of the asset can then be expressed as following:

$$pq + \int_0^q f(p,q)\mathrm{d}q = x$$

where f(p,q) is the price impact cost function: for BHK, $f(p,q) = p(e^{(\lambda_{bhk,i,t}/shrout)\times q} - 1)$; for GH, $f(p,q) = \lambda_{gh,i,t} \times q + \phi_{gh,i,t} \times p$; for effective and quoted spread, $f(p,q) = p * effectivespread_{i,t}$ or $= p * effspread_{i,t}$.

Therefore, given a market-on-close volume, the final price impact normalized by initial investment of $p \times q$ can then be expressed as:

$$\begin{aligned}
price_impact_{bhk,i,t} &= \frac{e^{(\lambda_{bhk,i,t}/shrout)\times q} - 1}{\lambda_{bhk,i,t}/shrout \times q} - 1; \\
price_impact_{gh,i,t} &= \frac{(\lambda_{gh,i,t} \times q)}{2 \times p} + \phi_{gh,i,t}; \\
price_impact_{effect_spread} &= effective_spread_{i,t}; \\
price_impact_{quoted_spread} &= quoted_spread_{i,t};
\end{aligned}$$

Since these price impact estimates are based on the intraday transaction data from 1993 to 1997, this likely results in our overestimating the costs for the current sample period.

Table A.II: Price Movement and Closing Auction Imbalance

Panel A of this table reports Fama-Macbeth regression of price movement starting from 15:45 to 16:00 on the market-on-close trading volume measures and controls. The sample consists of all common stocks(share code of 10 or 11, excluding those are priced less than $5)) with non-zero market-on-close trading volume in NYSE stock exchange from July 2008 to June 2018. Price movement is defined as the absolute value (in percentage) of the return during the period and the measures of market-on-close trading volume include *moc volume*, which is defined as the total market-on-close trading volume in the TAQ dataset scaled by the total trading volume per stock, *highmoc*, which is an indicator as to whether the stock is in the top quintile of *moc volume* for the day. Panel B reports the Fama-Macbeth regression of 15 minutes price movements from 15:15 to 15:30 (Column 1), or 15:30 to 15:45 (Column 2) on the market-on-close trading measure *highmoc* as placebo tests. The controls in regression include order imbalance, turnover, defined in the appendix. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: All NYSE Stocks from March 2010 to June 2018

Dependent variable:	Price movement from 15 : 45 to 16 : 00		
	(1)	(2)	(3)
MOC volume	0.2817*** (4.51)		
Highmoc		0.0217** (2.12)	0.0438*** (7.26)
Turnover			0.0021*** (5.29)
Order imbalance			0.6047 (0.88)
Intercept	0.232*** (28.99)	0.2415*** (22.93)	0.2154*** (18.62)
Adj-R^2	1.45%	1.06%	3.41%

Panel B: Placebo Test Using Returns from Other Time Periods

Dependent variable:	Price movement from 15 : 15 to 15 : 30	Price movement from 15 : 30 to 15 : 45
	(1)	(2)
Highmoc	-0.003* (-1.91)	0.001 (0.36)
Turnover	0.003*** (48.14)	0.003*** (36.86)
Order imbalance	-0.124 (-2.14)	0.087 (0.88)
Intercept	0.168 (50.74)	3.722 (2.13)
Adj-R^2	3.34%	3.21%

Table A.III: Closing Auction and Price Reversal: Alternative Measure of Returns

This table reports estimates from Fama-Macbeth regression of following model:

$$Return_{i,t+1} = \alpha + \beta_1\ Return_{i,t} + \beta_2\ Return_{i,t} \times highmoc_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t}$$

The sample consists of all common stocks (*shrcd* in 10 or 11, excluding those are priced less than \$5) with non-zero market-on-close trading volume in NYSE stock exchange from July 2008 to June 2018. Return variable used in the independent variable and the interaction term is open to close return (Panel A) and is return from 15:45 to 16:00 (Panel B). The dependent variables include close-to-close return (columns 1), intraday return (column 2), and overnight return (column 3), defined in Equation 7. *highmoc* is an indicator as to whether the stock is in the top quintile of *MOC Volume* for the day, where *MOC Volume* is defined as the total market-on-close trading volume in the TAQ dataset scaled by the total trading volume per stock. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Return in Independent Variables is $Ret_{open\_to\_close,t}$

Dependent variable:	$Ret_{close\_to\_close,t+1}$	$Ret_{Intraday,t+1}$	$Ret_{overnight,t+1}$
	(1)	(2)	(3)
Return_t	-0.005**	0.002	-0.008***
	(-2.05)	(0.81)	(-8.84)
Highmoc # Return$_t$	-0.015***	-0.008***	-0.007***
	(-5.09)	(-2.81)	(-6.49)
Highmoc_t	-0.003	0.004	-0.007***
	(-0.66)	(0.78)	(-4.15)
Turnover_t	0.000	-0.001***	0.001***
	(0.39)	(-2.84)	(6.86)
Order imbalance_t	-1.64***	-0.658***	-1.04***
	(-9.41)	(-4.15)	(-17.23)
Amihud_t	-0.098	0.016	-0.055**
	(-1.60)	(0.25)	(-2.26)
log(Market Cap) $_t$	-0.002	-0.002	0.000
	(-0.74)	(-0.73)	(0.03)
$1/\text{Price}_t$	-0.164*	-0.312***	0.104***
	(-1.66)	(-3.54)	(3.45)
Intercept	0.095*	0.080	0.014
	(1.69)	(1.55)	(0.70)
Average Adj-R^2	0.054	0.053	0.053

Panel B: Return in Independent Variables is $Ret_{15:45\_to\_16:00,t}$

Dependent variable:	$Ret_{close\_to\_close,t+1}$	$Ret_{Intraday,t+1}$	$Ret_{overnight,t+1}$
	(1)	(2)	(3)
Return_t	-0.006	0.081***	-0.091***
	(-0.54)	(8.28)	(-20.09)
$\textit{Highmoc \# Return}_t$	-0.072***	-0.054***	-0.020**
	(-3.50)	(-3.45)	(-2.36)
Highmoc_t	0.007	0.010**	-0.004***
	(1.15)	(2.04)	(-2.58)
Turnover_t	0.001	-0.001**	0.002***
	(0.90)	(-2.43)	(5.57)
Order imbalance_t	-1.99	-1.09	-0.902
	(-6.38)	(-4.63)	(-8.30)
Amihud_t	-1.96	0.308	-1.65
	(-0.23)	(0.04)	(-1.00)
$\log(\text{Market Cap})_t$	-0.002	-0.001	0.000
	(-0.53)	(-0.50)	(0.07)
$1/\text{Price}_t$	-0.221**	-0.347***	0.087***
	(-2.01)	(-3.59)	(2.61)
Intercept	0.081	0.065	0.013
	(1.32)	(1.19)	(0.60)
Average Adj-R^2	0.051	0.051	0.050

Table A.IV: Closing Auction and Price Reversal: Controlling for Size

This table reports the coefficient estimates from Fama-Macbeth regression of the same regression in Table VI for different size quintiles. Return variable used in the independent variable, dependent variable, and the interaction term is the close to close daily return. At the end of each month, all common stocks (share code of 11, 12 , excluding those are priced less than $5, with non-zero market-on-close trading volume) are sorted into five quintile groups based on the market capitalization. Among each size group, Fama-Macbeth regression is estimated. *highmoc* is an indicator as to whether the stock is in the top quintile of *MOC Volume* for the day, where *MOC Volume* is defined as the total market-on-close trading volume in the TAQ dataset scaled by the total trading volume per stock. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Dependent variable:	$Ret_{close\_to\_close,t+1}$				
	Large Cap				Small Cap
	(1)	(2)	(3)	(4)	(5)
Return_t	-0.0059	-0.0060*	0.0019	0.0012	-0.0074***
	(-1.43)	(-1.65)	(0.52)	(0.37)	(-2.45)
Highmoc # Return$_t$	-0.0197***	-0.0132***	-0.0106**	-0.0152***	-0.0281***
	(-3.64)	(-2.46)	(-1.99)	(-2.84)	(-5.57)
Highmoc_t	0.0197***	0.0111	0.0116	0.0006	-0.0288***
	(2.82)	(1.49)	(1.32)	(0.07)	(-2.84)
Turnover_t	0.0005	0.0003	-0.0002	0.0000	0.0013*
	(0.63)	(0.49)	(-0.36)	(0.07)	(1.69)
Order imbalance_t	-2.4642***	-2.2847***	-2.0819***	-2.4593***	-1.1588***
	(-3.87)	(-4.63)	(-5.36)	(-8.65)	(-5.79)
Amihud_t	-38.3644*	-10.5531*	-0.6714	-1.3776***	0.0206
	(-1.90)	(-1.65)	(-0.64)	(-2.70)	(0.41)
log(Market Cap)$_t$	-0.0061**	-0.0068	-0.0201	-0.0068	0.0131**
	(-2.09)	(-0.88)	(-1.61)	(-0.57)	(1.97)
Intercept	0.1597***	0.1719	0.3590*	0.1634	-0.1166
	(2.83)	(1.33)	(1.84)	(0.91)	(-1.33)
Average Adj-R^2	0.0570	0.0482	0.0436	0.0375	0.0342

Table A.V: Closing auction and price reversal: Return as bid ask spread

This table reports the coefficient estimates from Fama-Macbeth regression of the same regression in Table VI using the average of bid-ask as the closing price to calculate the return. Return variable used in the independent variable, dependent variable, and the interaction term is the close to close daily return. The sample consists of all common stocks (share code of 10 or 11 , excluding those are priced less than $5) with non-zero market-on-close trading volume from July 2008 to June 2018. *highmoc* is an indicator as to whether the stock is in the top quintile of *MOC Volume* for the day, where *MOC Volume* is defined as the total market-on-close trading volume in the TAQ dataset scaled by the total trading volume per stock. The controls include order imbalance, turnover, logged market capitalization, lagged Amihud (2002) ratio, and inverse of stock price, defined in appendix A.I. All standard errors are adjusted for heteroskedasticity and autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Return variable is average of bid ask spread

Dependent variable:	$Ret_{close\_to\_close,t+1}$	
	(1)	(2)
Return_t	0.0004	-0.0005
	(0.13)	(-0.18)
Highmoc # Return$_t$	-0.0151***	-0.0157***
	(-5.08)	(-5.50)
Highmoc_t	-0.0029	0.0016
	(-0.41)	(0.29)
Turnover_t		0.0004
		(0.81)
Order imbalance_t		-0.1396
		(-0.94)
Amihud_t		-0.0712
		(-1.62)
log(Market $\text{Cap})_t$		-0.0018
		(-0.58)
Intercept	0.0553***	0.0761
	(2.42)	(1.27)
Average Adj-R^2	0.0194	0.0454

Table A.VI: Long/Short Trading Strategies: Subperiods

This table reports raw returns and risk-adjusted returns for daily equal weighted long/short trading strategies formed based on signal from market-on-close trading volume for different subperiods. The trading strategy for Panel A is to form portfolio at the end of each day by buying (selling) stocks that are within highest quintile of market-on-close total imbalance quantity and with imbalance side of 'Sell' ('Buy') based on the NYSE imbalance feed information at 15:45. The holding period starts from the end of day until the end of the next day. The trading strategy for Panel B is to form portfolio at the end of each day by buying (selling) stocks that are within the highest quintile of market-on-close trading volume and also have a negative (positive) daily return. The holding period starts from the end of the day until the end of the next day. Daily Fama-French risk factors are used to calculate risk-adjusted alphas. Columns 2 to 6 represent the risk-adjusted returns and factor loadings for the sample period from July 2008 to June 2013 and columns 7 to 11 represent the risk-adjusted returns and factor loadings for the sample period from July 2013 to June 2018. All standard errors are adjusted for autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Long-Short Strategy Using NYSE feed data

Subperiod:	07/2008 to 6/2013					7/2013 to 06/2018				
	Alpha(bp)	MktRf	SMB	HML	UMD	Alpha(bp)	MktRf	SMB	HML	UMD
Long	11.11***	1.13***				1.40	1.02***			
	(4.89)	(41.51)				(1.29)	(45.84)			
	10.72***	1.05***	0.56***	0.13**		1.79**	0.97***	0.49***	0.30***	
	(5.33)	(29.97)	(5.74)	(2.31)		(2.43)	(89.95)	(24.96)	(15.54)	
	10.42***	1.03***	0.58***	0.04	-0.11***	1.81**	0.97***	0.48***	0.29***	-0.02
	(5.48)	(28.38)	(5.93)	(0.86)	(-3.46)	(2.46)	(91.45)	(24.85)	(14.30)	(-1.26)
Short	-12.45***	1.10 ***				-2.05*	1.02***			
	(-5.79)	(31.30)				(-1.92)	(43.51)			
	-12.75***	0.98***	0.66***	0.26***		-1.68**	0.97***	0.52***	0.27***	
	(-6.76)	(36.87)	(6.06)	(4.18)		(-2.35)	(86.21)	(21.45)	(14.14)	
	-12.91***	0.97***	0.67***	0.22***	-0.06**	-1.66**	0.97***	0.51***	0.26***	-0.02
	(-6.78)	(33.00)	(6.02)	(3.17)	(-2.08)	(-2.32)	(88.30)	(21.56)	(13.65)	(-1.17)
Long-Short	23.3***	0.03				3.44***	-0.01			
	(10.01)	(0.94)				(4.16)	(-0.63)			
	23.29***	0.07***	-0.10*	-0.14*		3.47***	-0.01	-0.03	0.03	
	(9.78)	(3.04)	(-1.87)	(-1.73)		(4.20)	(-0.34)	(-1.00)	(1.61)	
	23.15***	0.07**	-0.09*	-0.18**	-0.05	3.47***	-0.003	-0.03	0.03	-0.001
	(10.03)	(2.48)	(-1.74)	(-2.26)	(-1.12)	(4.20)	(-0.34)	(-1.00)	(1.55)	(-0.02)

Panel B: Long-Short Strategy Using TAQ data

Subperiod:	07/2008 to 6/2013					7/2013 to 06/2018				
	Alpha(bp)	MktRf	SMB	HML	UMD	Alpha(bp)	MktRf	SMB	HML	UMD
Long	3.89*	1.17***				1.97	1.00***			
	(1.69)	(38.57)				(1.59)	(36.5)			
	3.42*	1.07***	0.77***	0.178***		2.32**	0.95***	0.51***	0.24***	
	(1.73)	(39.48)	(10.55)	(4.41)		(2.32)	(48.42)	(21.70)	(8.16)	
	3.19*	1.05***	0.78***	0.12***	-0.08***	2.31**	0.95***	0.52***	0.24***	0.01
	(1.65)	(38.65)	(10.72)	(2.66)	(-3.03)	(2.3)	(48.31)	(21.3)	(8.62)	(0.41)
Short	-7.59***	1.06***				-1.70	0.98***			
	(-3.42)	(32.29)				(-1.29)	(40.63)			
	8.04***	0.93***	0.78***	0.28***		-1.30	0.93***	0.53***	0.31***	
	(-4.30)	(33.06)	(11.23)	(6.85)		(-1.32)	(58.18)	(17.96)	(12.37)	
	-8.39***	0.91***	0.80***	0.19***	-0.12***	-1.30	0.93***	0.52***	0.30***	-0.01
	(-4.52)	(32.63)	(11.57)	(4.11)	(-4.16)	(-1.31)	(58.82)	(18.34)	(11.72)	(-0.83)
Long-Short	11.49***	0.11***				3.70***	0.02			
	(4.57)	(5.17)				(2.69)	(0.95)			
	11.46***	0.14***	-0.01	-0.10**		3.62***	0.02	-0.01	-0.07*	
	(4.59)	(5.95)	(-0.15)	(-2.42)		(2.65)	(0.94)	(-0.35)	(-1.92)	
	11.57***	-0.01	-0.07	0.04	0.00***	3.60***	0.02	-0.01	-0.06	0.02
	(4.63)	(-0.25)	(-1.56)	(1.15)	(4.63)	(2.63)	(0.93)	(-0.24)	(-1.64)	(1.04)

Table A.VII: Long/Short Trading Strategies: Value weighted

This table reports raw returns and risk-adjusted returns for daily value-weighted long/short trading strategies formed based on signal from market-on-close trading volume. The weights are based on the market cap, calculated as prior end-of-month stock price times shares outstanding. The sample consists of all common stocks with nonzero market-on-close trading volume from July 2008 to June 2018. The trading strategy for Panel A is to form portfolio at the end of each day by buying (selling) stocks that are within highest quintile of market-on-close trading total imbalance quantity and with imbalance side of 'Sell' ('Buy') based on the NYSE imbalance feed information at 15:45. The holding period starts from the end of day until the end of the next day. The trading strategy for Panel B is to form portfolio at the end of each day by buying (selling) stocks that are within the highest quintile of market-on-close trading volume and also have a negative (positive) daily return. The holding period starts from the end of the day until the end of the next day. Daily Fama-French risk factors are used to calculate risk-adjusted alphas. All standard errors are adjusted for autocorrelation with a lag length of 5 days (Newey and West (1987)). t-statistics are presented in parentheses. *, **, and *** indicate statistical significance for it at the 10%, 5%, and 1% levels, respectively.

Panel A: Long-Short Strategy Using NYSE feed data

	Alpha(bp)	MktRf	SMB	HML	UMD
Long	2.90***	0.93***			
	(2.74)	(40.18)			
	2.87***	0.94***	-0.04	-0.03	
	(2.65)	(34.56)	(-0.60)	(-0.75)	
	2.85***	0.94***	-0.04	-0.06	-0.04
	(2.68)	(34.93)	(-0.58)	(-1.41)	(-1.49)
Short	-1.44	0.85***			
	(-1.40)	(34.30)			
	-1.44	0.85***	0.02	0.01	
	(-1.38)	(25.61)	(0.27)	(0.08)	
	-1.41	0.86***	0.02	0.04	0.04**
	(-1.37)	(26.16)	(0.25)	(0.70)	(2.57)
Long-Short	4.26***	0.08***			
	(2.96)	(2.81)			
	4.24***	0.09**	-0.06	-0.04	
	(2.91)	(2.25)	(-0.82)	(-0.52)	
	4.19***	0.08**	-0.06	-0.1	-0.09**
	(2.94)	(2.04)	(-0.79)	(-1.37)	(-2.29)

Panel B: Long-Short Strategy Using TAQ data

	Alpha(bp)	MktRf	SMB	HML	UMD
Long	1.96	1.001***			
	(1.37)	(32.13)			
	2.13	0.955***	0.116	0.165***	
	(1.48)	(36.69)	(1.50)	(3.27)	
	2.11	0.949***	0.117	0.135***	-0.039
	(1.48)	(37.19)	(1.52)	(2.81)	(-1.56)
Short	-1.33***	0.903***			
	(-3.71)	(41.94)			
	-1.22	0.862***	0.150***	0.133***	
	(-0.95)	(38.78)	(2.45)	(4.01)	
	-1.25	0.854***	0.152**	0.091**	-0.056***
	(-0.98)	(39.86)	(2.49)	(2.38)	(-2.54)
Long-Short	3.33*	0.098***			
	(1.93)	(3.36)			
	3.38*	0.094***	-0.034	0.032	
	(1.95)	(3.70)	(-0.62)	(0.54)	
	3.39*	0.096***	-0.035	0.045	0.017
	(1.95)	(3.85)	(-0.62)	(-0.78)	(0.63)