You can only lend what you own: Inferring daily institutional trading from security lending supply*

March 8, 2025

Abstract

Institutions make their equity holdings lendable, allowing us to use the daily change in lendable shares to proxy daily net institutional trading in each stock. Our proxy better tracks quarterly changes in institutional ownership than existing alternatives, or even a subset of actual institutional trades, especially if we allow the corresponding elasticity to vary across stocks. Using this proxy, we document (1) price momentum anomaly obtains only if institutional trading and intraday returns oppose during the portfolio formation period, consistent with under-reaction; (2) negative short-term return predictability, consistent with transitory institutional price impacts; (3) institutions unwind holdings before earnings announcements and re-establish them afterwards, suggesting the earnings announcement premium is trade-driven; and (4) institutions provide liquidity to retail investors, e.g., around stock splits.

JEL Classification Codes: G14

Keywords: Institutional Trading, Security Lending, Lendable Equity, Proxy, Machine Learning

1 Introduction

In the US, almost 80% of total shares outstanding are held by institutions (Blume and Keim (2012)), raising many important questions regarding the role of institutional investors in financial markets. However, researchers face a challenge to perform analyses that require high frequency data on institutional trading activity. This is because institutions are only required to disclose their equity positions quarterly (via 13-F filings) and may seek to obscure their holdings to minimize transaction costs and maximize the value of their information. One possibility is to examine actual institutional trades from databases like ANcerno. However, these data account for a modest fraction of overall institutional trading and are only available until 2015 (Hu, Jo, Wang, and Xie (2018)). The second alternative is to resort to proxies of such activity. For example, some infer institutional buy and sell trades using transaction sizes and trade directions inferred by Lee and Ready (1991)'s algorithm (e.g., Lee and Radhakrishna (2000) and Campbell, Ramadorai, and Schwartz (2009)). However, in today's modern equity markets, the accuracy of such trade and quote based algorithms diminishes as institutions increasingly rely on sophisticated dynamic order splitting strategies, which render identifying their trades via algorithms difficult (O'Hara, 2015). We propose a new proxy for daily net institutional trading that addresses these concerns.

Our proxy is based on the simple premise that changes in the total amount of equity holdings that institutions make available for lending proxies changes in institutional ownership (IO). Institutions routinely make some of their holdings available for lending to potential borrowers of security loans in order to earn loan fees.[1] S&P Global Insights (formerly Markit) estimates the total number of lendable shares (lendable quantity) for each stock on a daily basis—see Section 2 for details—and makes them commercially available. According to IHS Markit's Quant Summary, lendable equity "*measures the supply/lendable quantity of the stock to be borrowed. It can be used as a high-frequency proxy for institutional ownership.*"

[1] According to an Office of Financial Research Survey, the majority of these lending assets are provided by investment firms, pension funds, and endowment funds.

Thus, our proxy is not subject to measurement complications that reflect trade execution strategies or the limited availability of data on actual institutional trades.[2]

Lendable quantity underestimates IO because each institution may not lend more than one-third of its total holdings.[3] Empirically, lendable quantity, on average, accounts for 35% of institutionally held shares, and this ratio, denoted *Lratio*, varies across stocks.[4] For example, the ratio tends to be lower for growth and volatile stocks, and stocks with concentrated institutional ownership. Importantly, at the stock level, the institutional lending propensity is highly persistent. For example, the average quarterly autocorrelation in the ratio is 86%, consistent with Dong and Zhu (2024)'s finding that lending supply is inelastic to price changes. These observations lead us to use the daily change in the lendable quantity (*dLend*), divided by *Lratio* at the end of the previous quarter, to proxy daily changes in IO, i.e., net institutional trading activity. Indeed, we confirm that the ability *dLend*/*Lratio* to track institutional trading is similar across stocks with different *Lratio*s.

We compare the ability of *dLend*/*Lratio* as a proxy of net change in IO to those of three alternative proxies of institutional trading. Since true institutional holdings are only observable quarterly, we use quarterly changes in IO as the benchmark. We aggregate *dLend*/*Lratio* and three alternative proxies to the quarterly level before examining their associations with quarterly changes in IO. In the spirit of Lee and Radhakrishna (2000) and Campbell et al. (2009), the first alternative is the net amount of signed large trades whose values exceed $50,000.[5] Another proxy is the imbalance in retail trading volume identified using Boehmer, Jones, Zhang, and Zhang (2021)'s (BJZZ's) algorithm multiplied by -1, with the premise that most non-retail trades capture institutional trades.[6] Finally, we look

[2]A small fraction of lendable equity reflects equity holdings of individual investors made available for lending by retail brokers as custodians. Even though quantifying this fraction is difficult, it introduces noise to our proxy of net daily IO change. Hence, it's effects attenuate our findings, rather than drive them.

[3]Investment companies typically do not have more than one-third of the value of their portfolio on loan at any given point in time. This limitation stems from the asset coverage requirements in section 18 of the Investment Company Act.

[4]Aggarwal, Saffi, and Sturgess (2015) report that average share of lendable quantity to market-cap is about 28%. Dividing 28% by 35% implies an approximate 80% institutional ownership as expected.

[5]We find similar results using other cutoff points such as $20,000.

[6]This is consistent with the negative association between BJZZ retail imbalances and institutional trade

at a subset of actual institutional trades available from ANcerno, which are limited to a shorter sample period from 2010 through 2014.

We first evaluate the in-sample ability of the four proxies in tracking the actual quarterly change in IO. Specifically, we examine the slope coefficients from panel regressions of each of the four daily proxies (aggregated to quarterly) on quarterly changes in institutional trading, with or without quarter and stock fixed effects. We find that the change in lendable quantity has the strongest association with the actual change in IO. Concretely, one standard deviation increase in $dLend/Lratio$ is associated with 0.34–0.39 units increase in the standardized actual change in IO, depending on the set of fixed effects used. ANcerno trades yield the second best fit, with analogous estimates between 0.18 to 0.20. The proxies based on large trades and BJZZ trades perform poorly, with slope coefficients under 0.02. Similar results hold when all four proxies enter multivariate regressions.

More striking evidence of our proxy's superior performance obtains in out-of-sample analyses during 2013-2021. Each quarter, we use data from the prior 20 quarters in simple OLS regressions to predict next quarter's change in IO using each of the four proxies, i.e., we skip one quarter between estimation and prediction periods. To examine predictive power, we then run cross-sectional regressions of the actual quarterly IO change on the predicted IO change in each quarter, averaging the resulting R-squareds across quarters. We find an average R-squared of 13.8% using our proxy of net institutional trading. This average R-squared remarkably exceeds the analogues obtained using the other three proxies, i.e., only 0.34% for BJZZ trades, 0.29% for large transactions, and 5.80% for ANcerno trades.

Our baseline analysis assumes that the elasticity of IO with respect to lendable equity is constant. However, the in-sample ability of the scaled lendable share changes ($dLend/Lratio$) to track IO changes varies across stocks. The association between these two quarterly vari-

imbalances documented by Barardehi, Bernhardt, Da, and Warachka (2024) using ANcerno data from 2010-2014. Alternatively, Battalio, Jennings, Salgam, and Wu (2024) report a positive correlation between retail imbalances estimated by BJZZ's algorithm and a subset of institutional trades in S&P500 stocks from Jan, 2010 through Mar, 2011. We primarily focus on the explanatory power of BJZZ imbalances for institutional trading, rather than the direction of the correlation. We find similar results using the improvements that Barber, Huang, Jorion, Odean, and Schwarz (2023) propose on BJZZ's algorithm.

ables is stronger among stocks with higher lending activity, as reflected by higher utilization rates or lower average loan tenure. This association is also stronger among stocks with more dispersed institutional ownership, large stocks, growth stocks, volatile stocks, and recent winners. Again, these are stocks that are likely to be associated with more lending activity. The slope coefficient reflecting this association varies between 0.31 and 0.46 conditional on these characteristics. These findings suggest that the elasticity between lendable equity and IO varies across stocks, leading us to relax the constant-elasticity assumption for our out-of-sample analysis. In fact, allowing the association between changes in IO and changes in lendable equity to vary with stock characteristics improves the accuracy of our out-of-sample predictions, elevating the average R-squared from the 13.8% baseline to 17.7%.

Observing that the elasticity of IO with respect to lendable equity can be a complex function of stock characteristics, we also employ several machine learning methods, including Random Forest, Gradient Boosting, and various Ensamble methods, to predict quarterly changes in IO using our proxy. However, we find that these machine learning methods underperform the simple OLS approach when evaluated based on the out-of-sample R-squareds. This likely reflects machine learning algorithms' tendency to overfit outliers. Consistent with this conjecture, the average R-squared from machine learning methods improve to be slightly above 17.7%, when we trim the most extreme 10% of $dLend$ observations. Given this negligible improvement despite trimming a significant fraction of data when employing these non-OLS alternatives, we rely on the parsimonious OLS approach whenever we employ predicted daily institutional trading in several applications.

We use our proxy in several applications, providing novel evidence on the ability of net daily changes in institutional ownership to predict stock returns as well as the dynamics of net institutional trading around key corporate/stock events. We first use $dLend/Lratio$ to provide new evidence consistent with a link between the momentum anomaly and price under-reaction, building on the premise that institutional investors' trading incorporate long-lived information in prices (Sias, Starks, and Titman, 2006). Barardehi, Bogousslavsky, and

Muravyev (2024) attribute the momentum anomaly to under-reaction of investors to price signals that are associated with trading by showing that momentum strategies only work when portfolios are formed using past intraday returns, but not past overnight returns. We refine this result by showing that it holds when portfolios are formed on past intraday returns that *oppose* the directions the corresponding daily changes in IO. The economic link between intraday returns and daily changes in IO reflects the fact that institutional investors primarily trade intraday (Lou, Polk, and Skouras, 2019). Our findings suggest that when institutions trade in the same direction as the information, reflected by intraday price changes, then prices adjust correctly and there is no under-reaction. If, however, the sign of net institutional trading opposes those of intraday returns, e.g., reflecting liquidity timing by institutions, then momentum strategies are profitable suggesting price under-reaction.

Second, we show that $dLend/Lratio$'s short-term return predictability aligns with price dynamics associated with institutional liquidity consumption that exerts price pressure and is followed by reversals (Campbell, Grossman, and Wang, 1993; Hendershott and Menkveld, 2014). Daily long-short strategies that buy stocks in $dLend/Lartio$'s top decile, i.e., reflecting institutional buying pressure, and sell stocks in $dLend/Lartio$'s bottom decile, i.e., reflecting institutional selling pressure, are associated with negative future returns. The equally-weighted average 10-day raw or risk-adjusted returns to these strategies are over 31bps, while the value-weighted counterparts are over 18bps. We find similar results when we, instead, use out-of-sample predictions of daily institutional trading. Specifically, the long-short strategy yields equally-weighted average 10-day raw or risk-adjusted returns of over 29bps and value-weighted returns of over 20bps. These analyses further validate our proxy of directional institutional trading.

Third, we examine net institutional trading around earnings announcements, documenting that institutions unwind their holdings before earnings announcements and re-establish them afterwards. Our findings match those of Di Magio, Franzoni, Kogan, and Xing (2023) based on actual institutional trades from ANcerno data. Johnson and So (2018) report a similar tendency among financial intermediaries that leads to increased trading costs for sellers prior to earnings announcements. That is, institutional investors tend to unwind positions prior to announcements, when market makers also limit liquidity provision to sellers. Together, these two effects may exacerbate negative transitory institutional price impacts prior to announcements that are then followed by positive price impact as institutions re-establish holdings. Thus, our findings complement Di Magio et al. (2023) by offering suggestive evidence of trade-driven increased stock-return commonalities around earnings announcements that reflect the earnings announcement premium, well documented by Beaver (1968) and Frazzini and Lamont (2007), among many others.[7]

Finally, we analyze institutional trading around stock splits. We find that institutions tend to become net sellers on the day of splits. This is consistent with institutions, who intend to unwind positions, timing their liquidity consumption to trade against retail investors, as suggested by Kaniel, Saar, and Titman (2008), who enter the market as the per-share price drops due to a split (Easley, O'Hara, and Saar, 2001). Moreover, institutions become net buyers in several days following the split. We attribute this to an expansion of institutional holdings due to reduced institutional trading costs following a stock split (O'Hara, Saar, and Zhong, 2019; Chung, Lee, and Rösch, 2020).

We contribute to the literature by proposing an effective proxy of *daily* changes in institutional ownership that addresses limited availability of high-frequency data on actual institutional trading activity. We motivate our proxy based on an extensive discussion of the current institutional details of security lending markets that informs researchers interested in studying these markets. Our proxy is simple, intuitive, and unaffected by methodological

[7]We do not investigate the economics behind these patterns. The literature offers different economic explanations for this premium. For example, while Patton and Verardo (2012) and Savor and Wilson (2016) offer explanations based on information spillovers and the resulting changes in systematic risk; Barber, De George, Lehavy, and Trueman (2013) and Yang, Zhang, and Zhang (2020) provide evidence of increased idiosyncratic risk around announcements. However, Di Magio et al. (2023) attribute institutional trading behavior around earnings announcements to post-announcement fund flow sensitivity, using data from 331 mutual funds that they match between ANcerno and CRSP. They argue that such flow sensitivity discourages institutions from taking advantage of earnings announcement premium and therefore constitutes a source of limits to arbitrage for institutional investors.

challenges that render existing proxies based on transaction sizes and inferred trade directions inaccurate in today's electronic order-driven markets. Finally, reflecting its accessibility, our proxy comes from a commercially available database that covers non-U.S. securities.

2 Institutional Details

2.1 U.S. Securities Lending Markets

A securities loan is a transaction where the owner of a security temporarily transfers legal ownership of a security to a borrower in an over-collateralized transaction.[8] The compensation that the lender receives depends on the type of collateral used to secure the loan. For cash collateralized loans, the most common form of collateral for U.S. equity loans, the lender re-invests the cash and earns interest. The lender rebates a pre-determined fixed rate back to the borrower and earns the difference between the interest earned on the securities and rebate rate as their fee.[9] For non-cash collateralized loans, the borrower must pay a cash fee that is generally a fraction of the loan value.

The securities lending market is divided into two segments sometimes referred to as the wholesale and retail segment of the market. The retail segment of the market refers to loans from broker dealers to their customers to facilitate specific short selling transactions. The terms of these loans from broker-dealers to their customers are often spelled out in the prime

[8]This transfer includes voting rights and the rights to dividends. See Aggarwal et al. (2015) for additional discussion of the role of securities lending on corporate voting actions. Securities lending agreements generally require that lenders be reimbursed for any dividends received while the stock is on loan by receiving a substitute dividend. See Dixon, Fox, and Kelley (2021) and Blocher, Reed, and Van Wesep (2013) for additional discussion of securities lending and dividends.

[9]If the security is in high demand, the rebate rate may be negative implying that the lender keeps all of the re-invested interest plus the borrower must provide additional compensation to the lender equal to the rate of the negative rebate. Borrowing costs for cash collateralized loans are often converted from rebate rates to lending fees, which can be more easily compared to non-cash collateralized loans. This is done by subtracting the rebate rate from the federal funds rate or the overnight bank funding rate (OBFR). It is also possible for the lender to lose money on the loan if their investment returns do not cover the rebate rate. This reality played a significant role in downfall of AIG during the 2008 financial crisis when AIG reinvested cash collateral from securities loans in to risky assets which ultimately did not pay off leaving AIG responsible to return the cash from securities loans plus the agreed upon rebate rate. See Peirce (2014).

brokerage agreement.[10] In the retail segment of the market no securities actually exchange hands. This is because broker-dealers generally facilitate clearing and settlement for their customers. Consequently a sale of any kind, short or otherwise, by one customer simply creates an obligation for the broker-dealer to deliver shares on the settlement date. This obligation is not account by account, but is netted across all the broker-dealer's accounts creating a net obligation for the broker dealer to deliver shares on the settlement day.

For broker dealers, the profit associated with lending to their customers and facilitating short sales is the difference between what they charge their customers for 'loans' and what it costs them to deliver their net share obligation at clearing and settlement. Broker dealers will typically source shares in the following order. First they will use their own inventory or from customer margin accounts, because these are the least expensive source of shares since there is no fee involved to acquire the shares. If they do not have sufficient shares to meet their clearing and settlement requirements from these sources they will then look to their own customers with fully paid lending agreements, which allow the broker-dealer to lend a customer's shares. If they still cannot source sufficient shares they will turn to the wholesale market to borrow shares.

The wholesale market comprises all "non-retail" loans. The primary purpose for loans in this market is to facilitate the net clearing and settlement obligations of various market participants—mostly broker dealers.[11] A market participant, usually a broker-dealer, who needs to borrow shares in the wholesale market will maintain relationships with one or more lending programs and will negotiate bilaterally with the lending program for the loan of the shares. Transactions in the wholesale market are made bilaterally, and often with a phone call, although electronic negotiations are increasingly common. High search costs

[10] Retail loans often have a pre-determined fixed rate associated with borrowing shares that are easy to borrow and cost-plus model to price loans for securities that are harder to borrow. For harder to borrow loans, the cost to borrow is benchmarked off of a reference rate, which is frequently the prevailing wholesale market rate plus a markup.

[11] An OFR Pilot Survey indicated that approximately 85% of all wholesale loans went to broker dealers. The remainder generally went to large entities like exceptionally large hedge funds or pension and sovereign wealth funds that are large enough to bypass broker-dealers in the borrowing process and maintain their own relationships with lending programs and facilitate clearing and settlement internally.

characterize this market (Kolasinski, Reed, and Ringgenberg (2013), D'avolio (2002), Duffie, Garleanu, and Pedersen (2002)). Lending rates for wholesale loans are negotiated bilaterally, and while the forces of supply and demand play a key role in determining lending rates, other factors, combined with high search costs, can be significant and thus rates can vary significantly, even for similar loans on the same day.

The key feature of the wholesaler market from the perspective of our study is that the primary suppliers of shares in this market are institutional investors such as investment firms, pension and endowment funds, banks, insurance companies, and government entities.[12] Most of these entities do not supply more than one-third of their holdings' value according to Section 18(f)1 of the the Investment Company Act. Institutional investors make their shares available to loan by either offering the shares to a lending agent who runs a lending program, or if they are large enough, by running their own lending program. By far the largest lending programs are the major custodian banks who typically offer a reduction in their custodian fees per share of the lending revenue to customers who allow their shares to be lent by the custodian bank. Shares can be made available for lending on the day that the investor takes custody of the shares, i.e., the settlement date.[13]

2.2 Security Lending Data Sources

The securities lending market is opaque. There is limited transparency in the retail segment of the market.[14] In the wholesale market, data primarily comes from three main data providers (S&P Global Insights (formerly Markit), FIS, and Datalend). These companies obtain data via a give-to-get model, whereby participants in the wholesale securities lending

[12]Shares from non-institutional traders play a reduced role in the wholesale lending market because retail traders are less likely to make their shares available for lending and when they do, their shares are often used to facilitate the net clearing and settlement obligations of their own broker-dealer rather than the wholesale market in general. That said, broker-dealers of non-institutional traders will sometimes lend out the shares of their customers with fully paid lending agreements in the wholesale market.

[13]For additional institutional details regarding the structure of the securities lending market see the Economic Baseline section of recently adopted SEC Rule 10c-1a.

[14]There are some data providers that survey asset managers in the retail segment of the market about their lending experiences in order to gain insight into the retail segment of the market, but the coverage of these datasets is relatively small

market are required to give data to the vendor in exchange for the right to buy the aggregated data from the vendor, and usually only those with data to contribute can purchase concurrent access to the data.[15] Relevant for our study, participants often report the quantity of shares they have on loan along with the associated utilization rate, measuring the on-loan fraction of all shares a participant would make available as landable quantity—however, participants may or may not directly report the lendable quantity.

Additional data aggregation details highlight the challenging nature of inferring lendable quantity from the aggregate quantities of shares on loan and utilization rates. Each data provider has its own proprietary process for collecting, cleaning, and aggregating the data it receives. Key variables offered by the major wholesale market data providers include information about the distribution of loan fees and the quantity of shares on loan, e.g., average and standard deviation of these variables on loan across participants, at the *stock-day* level. Major data providers often do not provide direct estimates of the lendable quantity, but instead provide estimates of the utilization rate. This variable is computed by surveying multiple lending programs about their own utilization rates and then using a proprietary process to compute an *average* utilization rate. However, dividing average shares on loan by average utilization rate produces a highly noisy estimate of lendable quantity at the daily frequency for several reasons: (1) the data received from participants are aggregated using proprietary processes, which may weight observations based on undisclosed factors; (2) the estimate reflecting the ratio of two averages will be biased reflecting the likely non-zero cross-participant correlations between shares on loan and utilization rate; and (3) the lending programs providing utilization rate information to the data providers are not necessarily the same as those providing shares on loan information.[16]

[15]The quality and comprehensiveness of the data provided by these three companies is comparable. The give-to-get model limits access to the data and is designed to maximize participation since many participants would be unwilling to contribute data if they knew that it was being offered to, e.g. hedge funds and HFTs, that could potentially use the data form trading strategies that could harm them. Some providers make exceptions and allow academics and regulators to purchase the data.

[16]Observing that due to some reported utilization rates being extremely close to zero can result in absurd values. Consequently, some researchers estimate shares available using the formula SharesAvaliable = min(IO,SharesOnLoan/Utilization) were IO is the most recent institutional ownership based on 13F filings

S&P Global Insights (Markit), stands out among peer data providers as it provides users with direct measures of lendable quantity. We rely on these measures to develop our estimates of directional institutional trading. Plausibly, these lendable quantity measures are based on a proprietary aggregation process that is consistent with those that Markit employs to construct their reported metrics of shares on loan and utilization rates. Moreover, Markit's lendable quantity estimates can benefit from the aggregator's access to the distributional properties of shares on loan and utilization rates across contributing participants. Our empirical findings supportive of these conjectures: for example, quarterly changes in lendable quantity reported by Markit are strongly correlated with changes in quarterly changes in 13F institutional ownership measures; whereas, a weak analogous association obtains when we back out lendable quantity as the ratio of shares on loan to utilization rates reported by FIS.

2.3 Security Lending vs. Equity Trade Settlement Gap

The securities lending market has same day settlement while the equities market does not. Consequently, the loan of a security does not happen on the day that the equity market transaction occurred, but rather on the settlement day for that transaction.[17] Prior to September 5, 2017 the United States operated on a t+3 settlement cycle, meaning that shares for equity transactions were actually delivered three trading days after the transaction occurred. On September 5, 2017 the United States moved to t+2 settlement. On May 28, 2024, the United States moved to t+1 settlement for the equities market. To accurately capture the timing of net changes in institutional ownership, we account for the gap between security lending versus equity trade settlement periods. That is, we shift the date for a given

(Dixon et al. (2021)). We cannot use this approach since we aim to estimate daily IO using lendable quantity.

[17]Rule 203(b)(1) of Reg SHO requires that broker-dealers have reasonable grounds to believe that a stock is available for borrowing when settlement is due known as the "locate" requirement, which is intended to help ensure that they will be able to deliver the shares on the settlement date. In order to facilitate their own and their customer's short sales, a broker dealer obtains the 'locate' from a lending program on the day of the transaction. A 'locate' is an assurance from a lending program that shares will be available to borrow on the settlement date. Lending programs frequently offer locates for free for easy to borrow stocks by posting a list of easy to borrow stocks. For stocks that are harder to borrow, a lending program may charge a fee, in addition to whatever lending fee is charged, to provide a locate.

change in lendable shares backward by three business days before September 5, 2017, and by two business days after that date as our sample period ends in 2021.

Figure 1 provides an illustrative example, where we rely on non-informational institutional trading triggered in common stocks by Russell 1000/2000 index reconstitutions from 2010 through 2016. We show that one must account for settlement misalighments to accurately proxy the net change in institutional ownership using changes in lendable shares. Our example compares three outcomes across index-switcher stocks and the otherwise similar stocks in the indexes: (1) absolute changes in lendable equity, itself; (2) absolute estimated changes in institutional ownership; and (3) the true institutional trading volume obtained from ANcerno. We only shift dates associated with quantities of (1) and (2) to account for settlement differences, since (3) is a direct measure of institutional trading activity.

Each year, index-switching stocks between Russell-1000 and Russell-1000 indexes on the last Friday of June are selected as "treatment" stocks. For each index-switching (treated) stock, the two stocks whose Russell-1000/2000 rankings in the preceding May fall immediately above and below the treated stock are used as control stocks. Panel A plots the medians of $|dLend_{jt}|$ for treated and control firms in 30-day event windows around reconstitution dates. Panel B plots the medians of absolute estimated changes in institutional ownership, reflecting $dLend_{j,t+3}$ divided by the ratio of $Lend$ three days after the previous-quarter's end in institutionally held shares, IO, at the end of the previous quarter. Panel C plots the median share of actual institutional trading volume, observed in ANcerno data, in total number of shares outstanding with *no adjustments for settlements*. The alignments of the spikes in three panels support our approach to account for the settlement gaps.

3 Data and Sample Construction

3.1 Sample construction

Our sample includes all NMS-listed common shares between January 2007 through December 2021, merging data from 13F, Markit, Daily TAQ, CRSP, and I/B/E/S. From 13F, we collect quarterly information on institutional ownership. We obtain estimates of lendable shares and other security loan characteristics, including security loans tenure and utilization rates, are obtained from Markit. From WRDS Intraday Indicators, we obtain the volumes of buyer- and seller-initiated trades (identified by the Lee and Ready (1991)'s algorithm) whose transaction values exceed $50,000[18] as well as the volumes of buyer- and seller- initiated "retail" trades identified by the BJZZ algorithm. We obtain daily and monthly trading and price information, as well as risk-factor returns, from CRSP. Earning announcement dates and surprise scores come from I/B/E/S. Finally, for the period 2007 through 2014, we construct trading volumes of actual institutional buy and sell trades at the stock-day level using ANcerno.

We then apply the following filters to the data: First, we exclude observations where institutional holdings and lendable shares are either missing, exceed the total shares outstanding, or where lendable shares surpass institutional holdings. Such data points represent 2.1% of the initial sample. Second, we exclude observations with missing firm characteristics such as size, book-to-market value, Amihud illiquidity, volatility, turnover ratio, average return over the past year, institutional holdings, and idiosyncratic volatility. These observations account for 3.2% of the initial sample. Third, we require institutional holdings and lendable shares over consecutive quarters, in order to compute quarterly changes. This requirement excludes 11.4% of the initial sample. Fourth, we trim the data by removing observations with quarterly turnover ratios in the lowest 1% (0.7%) of the remaining (initial) sample. Stocks with exceptionally low turnover ratios are unlikely to experience substantial changes in in-

[18] We use a $20,000 cutoff to examine robustness.

stitutional holdings. Fifth, we exclude observations with extreme changes in split-adjusted total shares outstanding, where the share outstanding at quarter q is smaller than 50% or larger than 200% of the share outstanding at $q-1$. These observation account for 0.2% of the initial sample. Finally, we remove any observation with a Lend-to-IO ratio (*Lratio*) of less than 5%, trimming 1.3% of observations in the initial sample. Collectively, these filters reduce the number of observations by 19%.

3.2 Variable definitions

Our key variables include quarterly and daily (when possible) changes in insitutional ownership, lendable shares, as well as three existing proxies of institutional trading activity.

Quarterly measures for each stock are constructed as follows. The change in institutional holdings is:

$$dIO_q = \frac{IO_q - IO_{q-1}}{Shrout_{q-1}} \, ,$$

where, IO_q is the split-adjusted institutional holdings at the end of quarter q from 13-F, and *Shrout* is the number of shares outstanding from at the end of the previous quarter. Hence, dIO_q represents the change in the number of institutional shares normalized by the total shares outstanding. The change in lendable shares is defined similarly:

$$dLend_q = \frac{Lend_q - Lend_{q-1}}{Shrout_{q-1}} \, ,$$

where $Lend_q$ is the Markit's estimate for the quantity of lendable shares at the end of quarter q. The quarterly imbalance in BJZZ buy and sell volume is

$$Retail\_Trade_q = \frac{Retail\_Buy\_Shares_q - Retail\_Sell\_Shares_q}{Shrout_{q-1}} \, ,$$

where *Retail_Buy_Shares* and *Retail_Sell_Shares*, respectively, are the total amounts by buy and sell BJZZ share volumes, obtained from TAQ data and aggregated at the stock-quarter level. The quarterly imbalance in actual institutional activity using ANcerno data is

$$Institution\_Trade_q = \frac{Institution\_Buy\_Shares_q - Institution\_Sell\_Shares_q}{Shrout_{q-1}},$$

where, $Institution\_Buy\_Shares_q$ and $Institution\_Sell\_Shares_q$ are, respectively, the total share volumes of institutional buy and sell trades. Lastly, the quarterly imbalance in trades with values exceeding $50,000$ is

$$Trade50K_q = \frac{Trade50K\_Buy\_Shares_q - Trade50K\_Sell\_Shares_q}{Shrout_{q-1}},$$

where, $Trade50K\_Buy\_Shares_q$ and $Trade50K\_Sell\_Shares_q$ are, respectively, the total share volumes of large trades classified and buy and sell by the Lee-Ready algorithm.

We construct daily versions of these measures, with the exception of dIO. To maintain consistency, the daily version of $dLend$ is scaled by $Lend$ at the end of the previous calendar quarter. Similarly, the other three measures also scale daily changes by shares outstanding at the end of preceding calendar quarter.

We also construct the following stock characteristics at each quarter-end: (1) the number of institutional investors holding shares of a give stock obtained from 13F data, denoted **# Owners**; (2) the Herfindahl index of institutional ownership concentration calculated using 13F data, denoted **IOC_HHI**; (3) the natural log of firm size, measured by the product of closing price and the number of shares outstanding obtained from CRSP, denoted **log (Market Cap)**; (4) the book-to-market ratio reflecting the most recently observed book value and share price obtained from COMPUSTAT, denoted **BtoM**; (5) **Past Year Return**, calculated as the compound return of each stock stock over the preceding twelve months using CRSP; and (6) idiosyncratic volatility, which is the standard deviations of residuals of a market model estimated by WRDS Beta Suite using weekly data over the previous quarter, denoted **Idiosyncratic Vol**. Moreover, for each stock-quarter, we obtain the utilization rate, i.e., the ratio of shares lent divided by shares available averaged across lending programs, and average tenure, i.e., the average tenure across all outstanding security loans (in days), from Markit.

3.3 Summary statistics

Table 1 presents key summary statistics for the main variables of interest.

[Insert Table 1 here]

The mean and median of the fraction of institutionally-owned shares in total shares outstanding (IO) are 0.63 and 0.7, respectively. The average fraction of lendable equity in total shares outstanding (*Lend*) is 0.22. The lendable-to-IO ratio ($Lend/IO$) has a mean of 0.35, and a standard deviation of 0.14—indicative of its temporal and cross-sectional variation. The quarterly changes in both IO and Lend are close to zero on average. Their standard deviations are 0.06 and 0.03, respectively.

In terms of loan characteristics, on average, 17.42% of the lendable shares are lent out, for an average tenure of 88.73 days. We also find the institutional ownership to be quite dispersed in our sample with an average stock held by about 192 different institutions, with an average Herfindahl index of ownership concentration as little as 0.09. Table 1 also includes common stock characteristics such as the logarithm of market capitalization, the book-to-market ratio, average return over the past year, and idiosyncratic volatility.

The last three rows of Table 1 report the summary statistics of the alternative quarterly institutional trading proxies. The last column shows a much smaller number of observations when we examine actual institutional trading from ANcerno, which covers a shorter sample period from 2010 through 2014.

3.4 Lendable-to-IO ratio

Table 1 suggests that the lendable-to-IO ratio (*Lratio*) varies across stocks. Table 2 relates this variation to key stock characteristics.

[Insert Table 2 here]

Column (1) of Table 2 shows that *Lratio* is a persistent stock characteristic. Regressing *Lratio* on its own lag from the prior quarter yields a slop coefficient of 0.86, suggesting that *Lratio* is highly persistent from one quarter to the other for the same stock. This remarkable persistence in *Lratio* means that the change in *Lend* is highly correlated with the change in the underlying IO, even though institutions can make no more than 35% of their overall holdings available for lending.

Column (2) relates *Lratio* to the level and concentration of institutional ownership. The positive coefficient on # Owners and the negative coefficient IOC_HHI both suggest that the *Lratio* is higher for stocks with less concentrated institutional ownership. This finding is in line with the one-third cap on overall holdings that each institutional investor can make available for lending. Even though higher *Lratio* is associated with lower levels of institutional ownership, one should interpret this correlation cautiously. This negative association can be partially mechanical as IO appears in the denominator of the *Lratio*.

Column (3), examines the association between *Lratio* on some other key stock characteristics. *Lratio* tends to be higher for value stocks and stocks with lower idiosyncratic volatility, suggesting that such stocks are relatively more appealing from institutions perspective to be made available for lending. Column (4) includes all stock characteristics in one regression to demonstrate the robustness of associations documented in columns (2) and (3).

4 Tracking Quarterly Institutional Trading

This section validates the ability of our proxy of institutional trading by showing that it successfully tracks the changes in true institutional ownership. Since true institutional trading is difficult to observe at high frequencies, we use the quarterly changes in IO—obtained from 13F filings—as the benchmark. We show that our simple proxy is far superior to several alternatives in tracking institutional trading in terms of both in-sample association and out-of-sample predictive power. Our results obtain based on parsimonious uni- and multi-variate OLS estimates as well as sophisticated machine learning algorithms.

4.1 In-Sample Performance

We first evaluate the in-sample ability of the four proxies of institutional trading in tracking the actual change in IO at the quarterly frequency, i.e., dIO_q. We estimate

$$dIO^s_{jq} = a + bX^s_{jq} + u_{jq} \tag{1}$$

where dIO^s the standardized change in actual institutional ownership, i.e., dIO, and X^s is the standardized proxy $X \in \{dLend_q/Lratio_{q-1}, Retail\_Trade_q, Institution\_Trade_q, Trade50K_q\}$. In our baseline analysis, we conservatively assume a constant elasticity for IO_q with respect to $Lend_q$, leading us to scale $dLend_q$ by $Lratio$ from the previous quarter—later we show that relaxing this assumption only improves our results. Moreover, the use of both standardized dependent and independent variables in equation (1) facilitates straight forward comparisons slope coefficients (b) across the alternative proxies. These estimated slope coefficients capture the change in standardized dIO as a given proxy rises by one standard deviation. Hence a larger slope coefficient indicates the respective proxy's stronger ability to capture actual institutional trading. We examine specifications with or without firm fixed effects, and/or quarter effects, clustering standard errors by firm.

[Insert Table 3 here]

Table 3 shows that our proxy, $dLend/Lratio$, possesses the strongest association with actual changes in IO. Panel A, shows that when used as the sole explanatory variable, a one standard deviation increase in $dLend/Lratio$ is associated with 0.345 to 0.384 units of increase in standardized dIO depending on the set of fixed effects included. Moreover, the baseline adjusted-R^2 in the exercise is 15%. Panels B, C, and D report analogous results when standardized $Retail\_Trade_q$, $Institution\_Trade_q$, and $Trade50K_q$ are, respectively, used to explain standardized dIO. The b coefficient estimates for all of these proxies have the expected signs. However, their magnitudes are much smaller that those obtained for $dLend/Lratio$. In fact, the second best performing alternative is that constructed based on actual institutional trades obtained from ANcerno data, yielding b coefficients no greater than 0.193 and a baseline adjusted-R^2 of only 4%. The absolute values of the corresponding b coefficients for proxies based on BJZZ-identified and large trades never surpass 0.015 with negligible baseline adjusted-R^2. Panel E verifies that $dLend/Lratio$ maintains the strongest association with dIO when the other three proxies are also included as independent variables.[19]

The weak performance of alternatives relative to $dLend/Lratio$ should not surprise. First, ANcerno institutional volume accounts for 8-12% of the total trading volume (Hu et al. (2018)). Assuming that institutional volume accounts for 70% of the the total volume, it follows that ANcerno data covers only less than 20% of all institutional trades. Thus, as institutions can lend up to 30% of their holdings, our proxy likely offers a more accurate picture of overall institutional trading.

Second, trade sizes and inferred trading directions cannot effectively identify institutional trades. In today's order-driven fragmented markets, institutional investors employ sophisticated trade execution algorithms that split their intended (parent) orders dynamically and across trading venues and order types. As such institutional trades often appear in the form of small trades. Moreover, the frequent use of limit orders, low-latency, and prevalent trading at the quote midpoints renders the Lee-Ready unable to accurately sign errors. As such, classification of large trades into buy and sell becomes a challenge. See O'Hara (2015) for discussion of these issues.

Third, even though the imbalance in BJZZ-identified trades explain dIO with the expected negative sign, its explanatory power is minimal. This is consistent with Barardehi et al. (2024)'s finding that BJZZ-identified trades are inversely related to institutional trading

[19]Of note, the sample period for this analysis is from 2010 through 2014, reflecting limited ANcerno data.

only when liquidity is scarce. In such conditions, wholesalers internalize unequal amounts of retail buy and sell trades to provide liquidity to institutions, and the BJZZ algorithm picks up this imbalance. In normal times, however, institutions trade with other institutional counterparties at the midpoint, without a need to purchase liquidity from wholesalers. As a result, the inverse link between the imbalance in BJZZ-identified trades and institutional trading interest should be minimal in normal conditions.

4.2 In-Sample Conditional Performance

In this section, we revisit equation (1)'s assumption that IO's elasticity relative to *Lend* is constant across stocks. We investigate whether this is so by fitting equation (1) in different subsamples of stocks.

Specifically, in each quarter, we sort firms into two equally-large groups of each the following firm or security loan characteristics: **Lend/IO**, i.e.; *Lratio*, **Utilization**, which is the average ratio of shares on loan to lendable shares across security loans; **Average Tenure**, i.e., the average time duration for which loans were outstanding; **# Owners**, which is the number of institutional owners; **IOC_HHI**, denoting he Herfindahl index of institutional ownership concentration; **log (Market Cap)**, i.e., the natural log of the product of closing price and the number of shares outstanding; **BtoM**, defined as the book-to-market ratio based on the most recently observed book value and share price; **Past Year Return**, which is the average return of the stock over the preceding year; **Log(Institutional Holdings)**, i.e., the natural log of the number of shares held by institutional investors; and **Idiosyncratic Vol**, which is the standard deviations of residuals of market model estimated using weekly data over the previous quarter. In each subsample, we fit Fama-Macbeth estimates of equation (1) using standardized $dLend/Lratio$ as the independent variable and adopting Newey-West standard errors with 3 lags.

[Insert Table 4 here]

Panel A in Table 4 shows the association between our proxy and the actual quarterly changes in IO change is higher among stocks with more active security lending activity, reflected in higher utilization rates or lower average loan tenure. However, this association does not appear to significantly vary with *Lratio*. Since *Lratio* is highly persistent characteristic (see Table 2), this finding highlights the validity of our proxy regardless of the "economic importance" security lending at the individual stock level. Panel B shows sronger association between our proxy and dIO for stocks with more dispersed institutional ownership, where the lendable quantity is unlikely driven by the lending policies of a small number institutions holding a stock. Finally, Panel C reports stronger associations in large stocks, growth stocks, volatile stocks and recent winners, i.e., stocks that are likely to be associated with greater lending turnover.

4.3 Out-of-Sample Performance

We next turn to examining the abilities of the various proxies of daily institutional trading in predicting out-of-sample future institutional trading. As before, we aggregate these proxies at the stock-quarter level and then use quarterly changes in IO as a metric for actual institutional trading. For each proxy X in a quarter q^*, we estimate

$$dIO_{jq} = a + bX_{jq} + u_{jq} \tag{2}$$

using data from quarters $q^* - 20$ through $q^* - 1$. We then use the resulting parameter estimates and the observed X in quarter $q^* + 1$ to obtain the corresponding *predicted* change in institutional ownership, denoted $\widehat{dIO^*_{q+1}}$. Skipping one quarter ensures that our predictions are not subject to a potential look-a-ahead bias due to 2- or 3-day settlement-date adjustments. To examine the overall out-of-sample for a given proxy, we first regress the actual dIO_q on the predicted $\widehat{dIO_q}$ for each quarter where both quantities are available and store the resulting R^2_q. We then average each proxy's cross-sectional R^2_q's across quarters featuring predicted dIO_q and $\widehat{dIO_q}$. Of note, since our sample spans 2007-Q1 through 2021-Q4, employing 20 quarters to "train" equation (2) and skipping one quarter before making a prediction means that out-of-sample predicted dIO's are available as of 2013-Q2.[20]

Reflecting our findings in Section 4.2, we also allow a more flexible functional form when predicting dIO using $dLend$. In particular, while our baseline approach assumes a constant elasticity of IO relative to $Lend$, Table 4 suggest this elasticity may vary. We account for this possibility by implementing our out-of-sample prediction routine based on our proxy using the following model in the first step:

$$dIO_{jq} = a_0 + \sum_{k \in K} a_k \left[dLend_{jq} \times Char^k_{j,q-1} \right] + u_{jq}, \tag{3}$$

where $Char^k_{q-1}$ denotes the stock characteristics defined in Table 1. As before, for quarter q^* we use equation (3) parameter estimates using data from the preceding 20 quarters and the observed $dLend/Lratio$ in quarter $q^* + 1$ to obtain $\widehat{dIO_{q^*+1}}$.

[Insert Table 5 here]

Panel A in Table 5 highlights the superior performance of our proxy in forming out-of-sample predictions of institutional trading. When using equations (2) to form predictions, $dLend/Lratio$'s out-of-sample prediction R^2 averages at 13.8%. Reflecting the effectiveness of our proxy, even in a parsimonious setting, when we employ the more flexible equation (3) in the prediction process the average R^2 rises only to 17.7%. Both of these quantities are far larger for the analogues obtained using the other three proxies of institutional trading.

4.4 Out-of-Sample Performance: Machine Learning

In this section, we address the possibility that equation (3) is too parsimonious to capture the potentially complex and non-linearities in the relationships between IO's elasticity with respect to $Lend$ and stock characteristics. We employ machine learning algorithms to

[20] The exception is when we use ANcerno data covering 2010-2014, i.e., 20 quarters. Thus, we commence predicting dIO using $Institution\_Trade$ as of 2013-Q2 using data from the maximum number of past quarters available, excluding the immediately preceding quarter.

determine the "best" functional form governing these links. That is, we use

$$dIO_{jq} = Elasticity(Char^{1}_{j,q-1}, \ldots, Char^{k}_{j,q-1}, Lratio_{j,q-1}) \times dLend_{jq}, \tag{4}$$

where stock characteristics used are lendable shares, utilization rate, average tenure, number of institutional owners, institutional ownership concentration, market cap, book-to-market ratio, past year return, and idiosyncratic volatility, all obtained from the previous quarter. With the exception of changes in lendable shares, the remaining predictors are categorized into ten deciles each quarter and assigned values ranging from 1 to 10. Again, we use data from the preceding 20 quarters to train the model, skip one quarter, and then predict dIO one quarter ahead.

For a quarter q^*, we first use standalone nonlinear models, Elastic Net, Random Forest and Gradient Boosting to to train and validate $Elasticity(., q) \equiv dIO_{jq}/dLend_{jq}$ using the above stock characteristics and date spanning $q^* - 20$ through $q^* - 1$.[21] We then use the product of the **predicted** $Elasticity(., q^* + 1)$ and $dLend_{q^*+1}$ to obtain $d\widehat{IO_{q^*+1}}$. We also employ an ensemble of Elastic Net and Random Forest as well as an ensemble of Elastict Net, Random Forest, and Gradient Boosting. Ensemble predictions involve averaging the predictions generated by the underlying standalone models. For instance, the ensemble of Elastic Net, Random Forest, and Gradient Boosting averages the outputs of these three models for predicted $Elasticity(., q^* + 1)$ before forming predictions.

Reflecting the tendency of machine learning algorithm to over-fit outliers, we trim the most extreme top *and* bottom 5%, 2.5%, 1%, and 0.5% of elasticity observations from the training sample—but not from the validation and prediction samples. As Panel B in consistent with the sensitivity of machine learning algorithms to inclusion of outliers, Panel B in Table 5 clearly shoes the beneficial effects of outlier exclusions on the out-of-sample performance of our proxy when predictions are based on machine learning algorithms.

At a more general observation, however, is that the use of machine learning does not

[21] Training data covers quarters from $q^* - 20$ to $q^* - 2$, with quarter $q^* - 1$ reserved for validation.

appear to be decisively superior to the OLS-regression approaches reported in Panel A of Table 5. Specifically, average R^2's reported in Panel B of Table 5 indicate that the use machine learning would lead to minimal improvements in the out-of-sample performance of our proxy—average R^2s are only slightly larger than the 17.7% figure reported in Panel A of Table 5 *only if* we remove at least 10% of observations with most extreme quantities. These observations lead us to rely on our multi-variate OLS prediction approach, as opposed machine learning, when we analyze predicted institutional trading in the rest of the paper.

5 Daily IO Changes: Determinants and Applications

In this section, we first document the association between our proxy of daily changes in IO and several firm characteristics, as well as market conditions. We then apply our proxy to study aggregate institutional investor behavior in different contexts. Unless stated otherwise, throughout this section, we shift $dLend_{jt}$ backward by 2 or 3 days to account for settlement gaps between equity and security lending markets to capture the accurate timing of institutional trading (see Subsection 2.3 for details). Our findings from these applications align with the existing literature and uncover new patterns consistent with price under-reaction and strategic institutional investor trading.

5.1 Determinants of Daily IO Changes

We relate our proxy of daily net changes in IO and its absolute values to several observable stock characteristics and market outcomes. For $Y_{jt} \in \{dLend_{jt}/Lratio_{j,q-1}, |dLend_{jt}/Lratio_{j,q-1}|\}$ on day t of month m, with $Lratio$ recorded that end of the previous quarter, we estimate

$$\begin{aligned} Y_{jt} &= \sum_{i=1}^{5} a_i Y_{j,t-i} + b_1 \times cret_{j,t-5}^{t-1} + b_2 \times cret_{j,t-10}^{t-6} + b_3 \times cret_{j,t-30}^{t-11} + b_4 \times cret_{j,t-125}^{t-31} \\ &+ b_5 \times cret_{j,t-250}^{t-126} + c_1 \times SIZE_{j,m-1} + c_2 \times BM_{j,m-1} + c_3 \times TO_{j,m-1} \qquad (5) \\ &+ c_4 \times OCAM_{j,m-1} + c_5 \times Volatility_{j,m-1} + c_6 \times MISP_{jm} + Y_0 + u_{jt}, \end{aligned}$$

where $Y_{j,t-i}$ denotes stock j's lagged Y quantities from days $t-1$ to $t-5$; $cret_l^h$ is the cumulative returns from day l to h; $SIZE$ is the natural log of market-capitalization; BM is the book-to-market ratio; TO is share turnover; $OCAM$ is the open-to-close Amihud illiquidity measure (Barardehi, Bernhardt, Ruchti, and Weidemier, 2021); $Volatility$ is the daily return standard deviation; and $MISP$ is Stambaugh, Yu, and Yuan (2012)'s monthly mis-pricing factor. We estimate the model using Fama-MacBeth regressions of daily cross sections, applying the Newey-West correction with 30 lags to standard errors.

Panel A in Table 6 shows that, consistent with institutional investors' market-making activities (Anand, Irvine, Puckett, and Venkataraman, 2012), $dLend_{jt}/Lratio_{j,q-1}$ displays a rapidly decaying negative auto-correlation—subsections 5.2 and 5.5 provide additional suggstive evidence of institutional market making. Moreover, consistent with Griffin, Harris, and Topaloglu (2003), net daily changes in institutional holdings tend to be higher in stocks with past higher returns. Lastly, $dLend_{jt}/Lratio_{j,q-1}$ tend to ne higher in less liquid and more volatile stocks, and institutions appear to be expanding net positions in undervalued stocks as reflected by Stambaugh et al. (2012)'s factor. Panel B in Table 6 reports on an analogous analysis for $|dLend_{jt}/Lratio_{j,q-1}|$, suggesting that the magnitudes of the net changes in IO holdings persists over time, tend to be higher in recent loser stocks, value stocks, less liquid stocks, and undervalued stocks.

5.2 Long-term return predictability

We first apply our proxy of daily IO changes in the context of the momentum anomaly, highlighting the role of net institutional trading for momentum. We motivate this analysis by the premise that institutional trading reveals fundamental information (Sias et al., 2006) and therefore leads prices to reflect fundamental asset values more precisely. Thus, market participants put some weight on net changes in institutional holdings when assessing whether associated price movements reflect information about fundamental asset values. Our daily proxy allows us to separate price movements that align with changes in IO from those that oppose them to examine whether these two types of price signals have different return predictability in the context of the momentum anomaly.

Since institutional investors primarily trade intraday (Lou et al., 2019), we focus on the alignment of *intraday* returns, denoted IDR, and proxied daily changes in IO, $dLend/Lratio$. Following Barardehi et al. (2024), we examine month-m's 3-factor alphas of trading strategies that buy past winners and sell past losers according to compound intraday returns from months $m-12$ to $m-2$.[22] However, before compounding past intraday returns, we split them into those that align with the corresponding daily change in IO, i.e., $\text{sign}(IDR) = \text{sign}(dLend)$, and those that do not, i.e., $\text{sign}(IDR) \neq \text{sign}(dLend)$. Tables 7 and 8 report portfolio sort results using, respectively, CRSP and NYSE breakpoints. We find that momentum strategies based on past intraday signals are only profitable when the signs of past intraday returns oppose those if daily changes in IO. Importantly, these findings are not driven by the variation in the distribution of days with $\text{sign}(IDR) = \text{sign}(dLend)$ versus those with $\text{sign}(IDR) \neq \text{sign}(dLend)$. In fact, the number of trading days entering the constructions of portfolio formation signals is evenly distributed across the two cases and remains roughly stable across portfolios.

[Insert Table 7 here]

[Insert Table 8 here]

We find that momentum strategies become unprofitable when positive (negative) past intraday return signals are associated with expansions (contractions) of institutional positions. For example, in Table 7, when $\text{sign}(IDR) = \text{sign}(dLend)$, average compound intraday return in the formation period rises from -72.43% at the bottom portfolio to 50.97% at the top portfolio (a spread of 123.40%), and aggregate $dLend/Lratio$ rises from -0.01 at the bottom portfolio to 0.16 at the top, with a statically insignificant holding-period 3-factor alpha. Of note, aggregate overnight returns that follow intraday signals, denoted $FONR$,

[22]Our qualitative findings extend if we instead use signals from months $m-7$ to $m-2$. Moreover, not only are past overnight signals irrelevant for our analysis, but Barardehi et al. (2024) show that momentum strategies based on past overnight signals are not profitable.

indicate nearly no immediate overnight price reversals at the top portfolios during the formation period, i.e., in the top portfolio intraday returns revers by only $0.31/50.97 = 0.6\%$. This lack of reversals suggests that the expansions of institutional positions that are associated with large positive intraday returns are likely motivated by information that is impounded in prices in the form of permanent price impacts.

By contrast, when $\text{sign}(IDR) \neq \text{sign}(dLend)$ in Table 7, average compound intraday return rises from -75.47% at the bottom portfolio to only 37.89% at the top (a spread of 113.35%), with aggregate $dLend/Lratio$ falling from 0.08 at the bottom portfolio to 0.02 at the top—that is, institutions tend to expand positions across all portfolios, just less so as we go from the bottom portfolio to the top. The smaller spread between intraday signals at the top and bottom portfolios is then followed by a statistically significant holding-period 3-factor alpha of 0.95%. Of note, aggregate $FONR$ suggests that institutions selling when intraday returns are at the top portfolio enjoy a $6.63/37.89 = 17.4\%$ immediately subsequent price reversal. This is consistent with institutional investors timing their liquidity consumption to save on trading costs.

Our collective findings are consistent with under-reaction theories of the momentum anomaly. To the extent that changes in net institutional holdings lead to incorporation of information into prices, market participants under-react to intraday price signals when these signals do not align with the corresponding changes in institutional holdings. This under-reaction manifests itself in the form of momentum as future prices adjust.

5.3 Short-term return predictability

We next examine the short-term return predictability of net changes in IO by conducting simple portfolio sorts and examining raw and risk-adjusted returns to long-short strategies. For stock j on day t of quarter q, we use two daily proxies of institutional trading: (1) $dLend_{jt}/Lratio_{q-1}$, i.e., the actual daily change in lendable quantity, scaled by the ratio of lendable quantity to shares outstanding from the previous quarter-end; and (2) our backward-looking *predicted* institutional trading measure obtained from the OLS estimates of equation (2). As described in Section 4.3, we allow a one-quarter gap between the trading and prediction samples. The only difference here is that we use daily $dLend_{jt}/Lratio_{q-1}$ to predict daily $\widehat{dIO_{jt}}$. Moreover, to ensure no look-ahead bias contaminates our findings, we do not shift observations backwards to account for the 2- or 3-day settlement gaps. Finally, to ensure that the samples based on these two measures are consistent, we use data post 2013-Q2 where both measures are available.

On each day t, we sort stocks into ten portfolios of each measure from $t-1$. We then estimate both equally-weighted and value-weighted future returns to a trading strategy and buys stocks in the top decile, reflecting extreme institutional buying pressure, and sells stocks in the bottom decile, reflecting extreme institutional selling pressure. We then calculate averages of raw and and five-factor risk-adjusted cumulative returns over the subsequent 1, 2, 3, 5, and 10 trading days.

[Insert Table 9 here]

[Insert Table 10 here]

Tables 9 and 10 show that extreme directions institutional trading negatively predicts future returns. This result obtains for both $dLend_{jt}/Lratio_{q-1}$ (Table 9) and $\widehat{dIO_{jt}}$ (Table 10), both raw and risk-adjusted returns, and using both equal and value weighted portfolio returns. The negative returns associated with our portfolio sorts are also economically sizable, ranging between -20 to -33 bps across different specifications. Overall, these robust findings are consistent with price reversals followed by directional institutional trading (Campbell et al., 1993; Hendershott and Menkveld, 2014).

5.4 Institutional trading around earnings announcements

The daily change in lendable shares also allows us to examine how institutions trade around important new events such as earnings announcements. Figure 2 Panel A reports daily means and 95% confidence intervals of $dLend_{jt}/Lratio$ in 21-day event windows around earnings announcement dates.[23]

[Insert Figure 2 here]

Panel A shows that institutions tend to unwind their holdings before earnings announcements and re-establish them afterwards. These patterns suggest that Di Magio et al. (2023)'s findings using actual trades of institutional investors in ANcerno data are likely generalizable to a much broader cross-section. Relatedly, Johnson and So (2018) report that trading costs are also higher prior to earnings announcements, especially cost of selling, and attribute the effect to financial intermediaries' tendencies to reduce their exposure to announcement risks. That is, an increased liquidity demand by institutions is coupled by a reduction in liquidity supply prior to earnings announcements. As such, price impacts of institutional trading prior and after the announcement could contribute to the increased return commonalities around earnings announcements as well as the well-known earnings announcement premium (see, e.g., Beaver (1968) and Yang et al. (2020)).

Uncovering the the economics roots of these patterns is outside the scope of our paper. The literature offers a number economic explanations for this premium. For example, Patton and Verardo (2012) and Savor and Wilson (2016) offer explanations based on information spillovers and the resulting changes in systematic risk, while Barber et al. (2013) and Yang et al. (2020) document that idiosyncratic risk rises around announcements. Di Magio et al. (2023) use data from 331 mutual funds that that can be matched between ANcerno and CRSP to attribute institutional trading behavior around earnings announcements to post-announcement fund flow sensitivity. They argue that such flow sensitivity constitutes a source of limits to arbitrage for institutional investors that discourages institutions from taking advantage of earnings announcement premium.

Panels B and C plot estimates conditional on, respectively, negative and positive earnings surprises as reflected by by SUE scores obtained from I/B/E/S. We observe more persistent institutional buying post-announcement when the earnings surprise is positive. In this case,

[23]Unreported analysis confirms qualitatively similar patterns obtain if we use $\widehat{dIO_{jt}}$

the institutional trading measure is positive and significant during the entire 10-day window after the announcement. In contrast, in the event of negative earnings surprise, it quickly becomes indistinguishable from zero.

5.5 Institutional trading around stock splits

We finally examine institutional trading around stock splits. The literature has examined institutional trading activity around stock splits at monthly frequencies, e.g., Chemmanur, Hu, and Huang (2015) document increased *unsigned* monthly institutional trading volume following stock splits. Our proxy, $dLend_{jt}/Lratio$ allows us to shed new light on the directional institutional trading at daily frequencies around stock splits.[24]

We frame our high-frequency analysis by relating stock splits to strategic institutional-investor trading reflecting predictable variations in trading costs. We note that relative tick size, i.e., 1¢ divided by the share price, shapes the trading environment (O'Hara et al. (2019)). Specifically, a stock split when the minimum tick size is fixed at 1¢ raises the relative tick size by reducing the share price. This is similar a an increase in the minimum tick size without a stock split. Importantly, Chung et al. (2020) find that an increase in the minimum tick size, and hence an increase in the relative tick size, reduces the trading costs of institutional investors. Moreover, stock splits make ownership more accessible to retail investors by reducing the cost of purchasing each share. Hence, retail investors' demand to purchase a stock should increase following stock splits.

We document evidence consistent with institutional investors endogenously timing their trading relative to these predictable patterns in retail trade interest and institutional trading costs. Figure 3 shows that institutional flow significantly drops on the day of a split, consistent with institutions timing their selling to benefit from increased buying interest on the retail side, as first suggested by Kaniel et al. (2008). In the subsequent days, however, institutional trading reflects net buying, consistent with long only investors increasing their

[24] Unreported analysis confirms qualitatively similar patterns obtain if we use $\widehat{dIO_{jt}}$

positions to benefit from improved liquidity.

[Insert Figure 3 here]

6 Conclusion

Institutions can only lend what they currently own. Based on this simple intuition and the empirical fact that the ratio between lendable shares and institutional ownership is persistent, we propose to use the change in lendable shares to measure institutional trading.

At the quarterly frequency and during a more recent 2007-2021 sample period, we find the change in lendable shares to perform better in tracking institutional ownership change than alternatives based on large trades, non-retail trades and a subset of actual institutional trades. For example, a one standard deviation increase in lendable shares is associated with a 0.4 unit increase in the standardized actual change in institutional ownership. In out-of-sample prediction exercises using only past data, the change in lendable shares also perform better than these alternative. An OLS method that allows the elaticity between lendable share change and institutional ownership change to be a linear function of stock characteristics perform even better.

Importantly, lendable shares change at daily frequency, allowing us to track daily institutional trading. Daily analyses reveal three findings. First, daily institutional trading measures negatively and significantly predict future returns, consistent with the notion of a transitory price impact. Second, we find institutions unwind their holdings before the earnings announcement and re-establish them afterwards. The resulting price pressure contributes to the well-known earnings-announcement premium. Third, we find evidence consistent with institutions timing their liquidity consumption reflecting the predictable patterns in both retail trading interest and institutional trading costs around stock splits.

References

Aggarwal, R., P. A. Saffi, and J. Sturgess (2015). The role of institutional investors in voting: Evidence from the securities lending market. *The Journal of Finance 70*(5), 2309–2346.

Anand, A., P. Irvine, A. Puckett, and A. Venkataraman (2012). Performance of institutional trading desks: An analysis of persistence in trading costs. *Review of Financial Studies 25*, 557–598.

Barardehi, Y., V. Bogousslavsky, and D. Muravyev (2024). What drives momentum and reversal? evidence from day and night signals. Working paper.

Barardehi, Y. H., D. Bernhardt, Z. Da, and M. Warachka (2024). Institutional liquidity costs, internalized retail trade imbalances, and the cross-section of stock reutrns. *Journal of Financial and Quantitative Analysis, forthcoming*.

Barardehi, Y. H., D. Bernhardt, T. G. Ruchti, and M. Weidemier (2021). The night and day of Amihud's (2002) liquidity measure. *Review of Asset Pricing Studies 11*, 269–308.

Barber, B., E. T. De George, R. Lehavy, and B. Trueman (2013). The earnings announcement premium around the globe. *Journal of Financial Economics 108*, 118–138.

Barber, B. M., X. Huang, P. Jorion, T. Odean, and C. Schwarz (2023). A sub(penny) for your thoughts: Improving the identification of retail investors in TAQ. *Journal of Finance, forthcoming*.

Battalio, R., R. Jennings, M. Salgam, and J. Wu (2024). Identifying market maker trades as "retail" from TAQ: No shortage of false negatives and false positives. Working Paper.

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of Accounting Research 6*(1), 67–92.

Blocher, J., A. V. Reed, and E. D. Van Wesep (2013). Connecting two markets: An equilibrium framework for shorts, longs, and stock loans. *Journal of Financial Economics 108*(2), 302–322.

Blume, M. E. and B. D. Keim (2012). Institutional investors and stock market liquidity: Trends and relationships. Working Paper.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang (2021). Tracking retail investor activity. *Journal of Finance 76*, 2249–2305.

Campbell, J. Y., S. J. Grossman, and J. Wang (1993). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics 108*, 905–939.

Campbell, J. Y., T. Ramadorai, and A. Schwartz (2009). Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics 92*, 66–91.

Chemmanur, T. J., G. Hu, and J. Huang (2015). Institutional investors and the information production theory of stock splits. *Journal of Financial and Quantitative Analysis 50*, 413–445.

Chung, K. H., A. J. Lee, and D. Rösch (2020). Tick size, liquidity for small and large orders, and price informativeness: Evidence from the tick size pilot program. *Journal of Financial Economics 136*(3), 879–899.

Di Magio, M., F. Franzoni, S. Kogan, and R. Xing (2023). Avoiding idiosyncratic volatility: Flow sensitivity to individual stock returns. Working paper.

Dixon, P. N., C. A. Fox, and E. K. Kelley (2021). To own or not to own: Stock loans around dividend payments. *Journal of financial economics 140*(2), 539–559.

Dong, X. and Q. Zhu (2024). Equity lender base and limits to arbitrage: Position-level evidence from mutual funds. Working Paper.

Duffie, D., N. Garleanu, and L. H. Pedersen (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics 66*(2-3), 307–339.

D'avolio, G. (2002). The market for borrowing stock. *Journal of financial economics 66*(2-3), 271–306.

Easley, D., M. O'Hara, and G. Saar (2001). How stock splits affect trading: A microstructure approach. *Journal of Financial and Quantitative Analysis 31*(1), 25–51.

Frazzini, A. and O. Lamont (2007). 2007, the earnings announcement premium and trading volume. NBER Working Paper 13090.

Griffin, J. M., J. H. Harris, and S. Topaloglu (2003). The dynamics of institutional and individual trading. *Journal of Finance 53*, 2285–2320.

Hendershott, T. and A. J. Menkveld (2014). Price pressures. *Journal of Financial Economics 114*, 405–423.

Hu, G., K. M. Jo, Y. A. Wang, and J. Xie (2018). Institutional trading and abel noser data. *Journal of Corporate Finance 52*, 143–167.

Johnson, T. L. and E. C. So (2018). Asymmetric trading costs prior to earnings announcements: Implications for price discovery and returns. *Journal of Accounting Research 51*(1), 217–263.

Kaniel, R., G. Saar, and S. Titman (2008). Individual investor sentiment and stock returns. *Journal of Finance 63*, 273–310.

Kolasinski, A. C., A. V. Reed, and M. C. Ringgenberg (2013). A multiple lender approach to understanding supply and search in the equity lending market. *The Journal of Finance 68*(2), 559–595.

Lee, C. M. B. and B. Radhakrishna (2000). Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets 3*(2), 83–111.

Lee, C. M. C. and M. J. Ready (1991). Inferring trade direction from intraday data. *Journal of Finance 46*(2), 733–746.

Lou, D., C. Polk, and S. Skouras (2019). A tug of war: Overnight versus intraday expected returns. *Journal of Financial Economics 134*(1), 192–213.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics 116*, 257–270.

O'Hara, M., G. Saar, and Z. Zhong (2019). Relative tick size and the trading environment. *The Review of Asset Pricing Studies 9*(1), 47–90.

Patton, A. J. and M. Verardo (2012). Does beta move with news? firm-specific information flows and learning about profitability. *The Review of Financial Studies 25*(9), 2789–2839.

Peirce, H. (2014). Securities lending and the untold story in the collapse of AIG. Working Paper.

Savor, P. and M. Wilson (2016). Earnings announcements and systematic risk. *Journal of Finance 71*, 83–138.

Sias, R. W., L. T. Starks, and S. Titman (2006). Changes in institutional ownership and stock returns: Assessment and methodology. *Journal of Business 79*, 2869–2910.

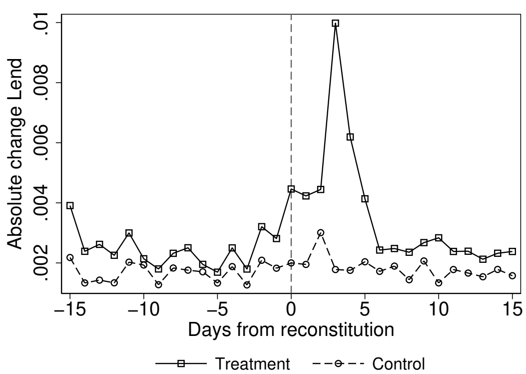
Stambaugh, R., J. Yu, and Y. Yuan (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics 104*, 288–302.

Yang, Y. C., B. Zhang, and C. Zhang (2020). Is information risk priced? evidence from abnormal idiosyncratic volatility. *Journal of Financial Economics 135*, 528–554.

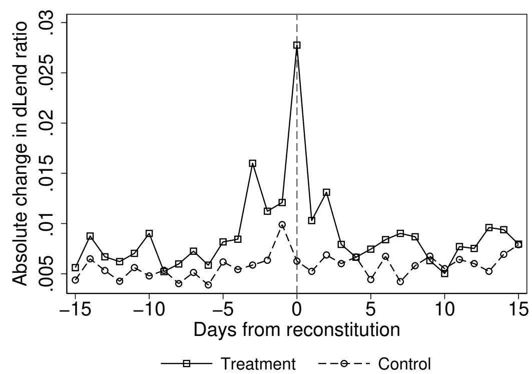
Figures and Tables

Figure 1. Daily Absolute Changes in Lendable Shares and Estimated Changes in Institutional Ownership around Index Reconstitution Dates. This figure reports on the variation in absolute changes in the number of lendabel shares as well as absolute estimated changes in institutional ownership around stock index reconstitution dates. Each year, index-switching stocks between Russell-1000 and Russell-1000 indexes on the last Friday of June are selected as "treatment" stocks. For each index-switching stock, the two stocks whose Russell-1000/2000 rankings in the preceding May fall immediately above and below the treated stock are used as control stocks. Panel A plots the medians of $|dLend_{jt}|$ for treated and control firms in 30-day event windows around reconstitution dates. Panel B plots the medians of estimates absolute changes in institutional ownership, i.e., $|dLend_{j,t+3}/Lratio|$, where $Lratio$ divides $Lend$ three days after the previous-quarter's end to IO at the end of the previous quarter. Estimates are shifted by three days reflecting the three-day gap between actual trade and settlement days in the security lending market. Panel C plots the median share of actual institutional trading volume, observed in ANcerno data, in total number of shares outstanding. The sample includes Russell-1000 and Russell-2000 common stocks from 2010 through 2016, with ANcerno data limited to 2010 through 2014.

Panel A: Change in lendable shares



Panel B: Estimated change in institutional ownership



Panel C: Institutional trading volume (ANcerno)

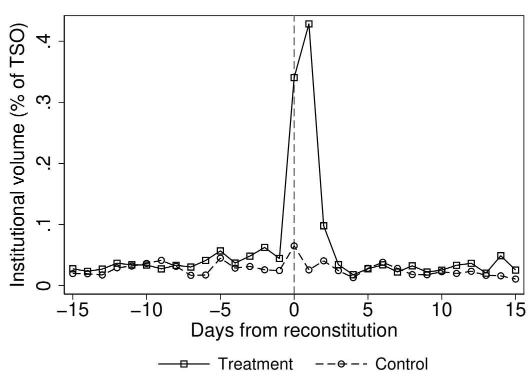
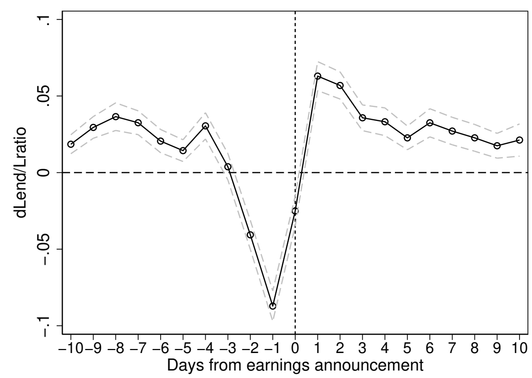
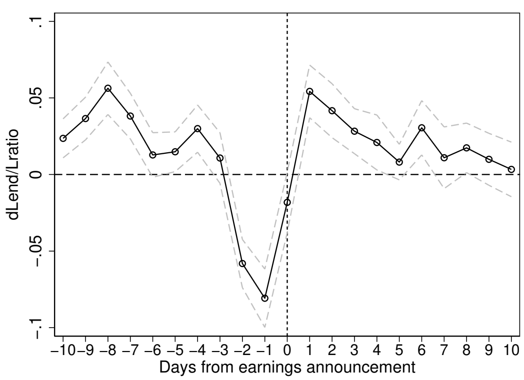


Figure 2. Estimated Daily Changes in Institutional Ownership around Earnings Announcements. This figure reports on the average net changes in estimated institutional ownership around earnings announcement dates. Panel A plots daily means and 95% confidence intervals of $dLend_{jt}/Lratio$ in 21-day event windows around earnings announcement dates. Panels B and C plot estimates conditional on, respectively, negative and positive earnings surprises, measures by SUE scores. Earnings announcement dates and SUE scores are obtained from I/B/E/S. To account for settlement gaps between equity and security lending markets, $dLend_{jt}/Lratio$ observations shifted backward three days prior to September, 6, 2017 and are shifted backward two days as of September, 6, 2017. The sample includes common NMS-listed stocks from January 2013 though December 2021. Daily $dLend_{jt}/Lratio$ observations are winsorized at 1% and 99%. Confidence intervals reflect standard errors that are clustered by date.

Panel A: All earnings announcements



Panel B: Earnings announcements with SUE< 0



Panel C: Earnings announcements with SUE> 0

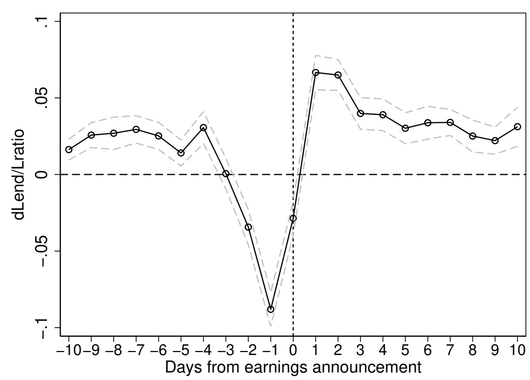


Figure 3. Estimated Daily Changes in Institutional Ownership around Stock Splits. This figure reports on average net changes in estimated institutional ownership around stock splits. It plots daily means and 95% confidence intervals of $dLend_{jt}/Lratio$ in 21-day event windows around stock-split dates. To account for settlement gaps between equity and security lending markets, $dLend_{jt}/Lratio$ observations are shifted backward three days prior to September, 6, 2017 and are shifted backward two days as of September, 6, 2017. The sample includes common NMS-listed stocks from January 2013 though December 2021. Daily $dLend_{jt}/Lratio$ observations are winsorized at 1% and 99%. Confidence intervals reflect standard errors that are clustered by date.

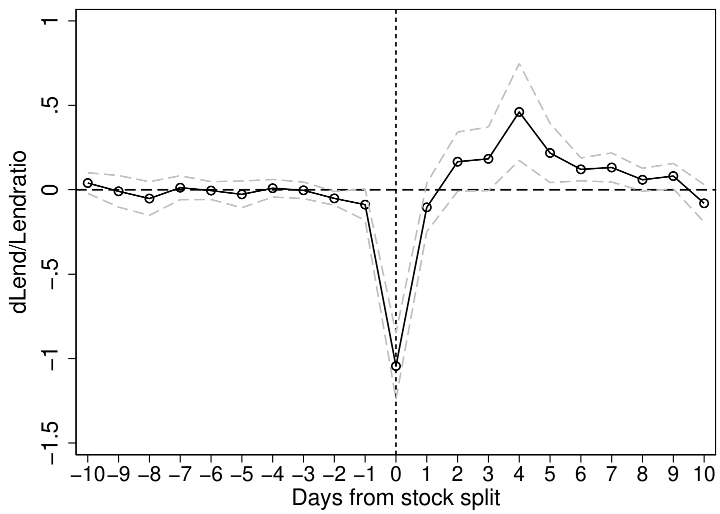


Table 1. Summary Statistics. This table reports mean, median, standard deviation, 5 percentile, 95 percentile, skewness, kurtosis, and the number of observations of the key variables, which are defined as follows: ***IO*** is the split-adjusted institutional holdings normalized by the total share outstanding; ***Lend*** is the split-adjusted lendable shares normalized by the total share outstanding; **Change in *IO*** is the split-adjusted change in institutional holdings normalized by the total share outstanding; **Change in *Lend*** is the split-adjusted change in lendable shares normalized by the total share outstanding; **Utilization** measures the average ratio of shares on loan to lendable shares across security loans,; **Average Tenure** measures the average time duration for which loans were outstanding; **# Owners** is the number institutional owners; **IOC_HHI** is the Herfindahl index of institutional ownership concentration; **log (Market Cap)** is the natural log of the product of closing price and the number of shares outstanding; **BtoM** is the book-to-market ratio based on the most recently observed book value and share price; **Past Year Return** is calculated by the average return of the stocks over past one year; **Idiosyncratic Vol** is the idiosyncratic volatility is the standard deviations of residuals of market model estimated using weekly data over the previous quarter; **Retail Trade** represents the imbalance between buyer- vs. seller- initiated internalized retail trades identified by the BJZZ algorithm in TAQ; **Institution Trade** is the institutional order flow obtained from ANcerno; **Trade>50K** represents imbalance between buyer- vs. seller-initiated trades with dollar volumes of at least $50,000 obtained from TAQ; and ***Lend/IO*** is the ratio of lendable share and institutional holdings.

	Mean	Median	Std	p5	P95	skew	kurt	N
IO	0.64	0.70	0.26	0.12	0.96	-0.66	-0.66	105169
Lend	0.22	0.22	0.12	0.03	0.42	0.20	-0.25	105169
Lend/IO	0.35	0.34	0.14	0.13	0.56	3.53	116.82	105169
Change in *IO*	0.01	0.00	0.06	-0.06	0.09	2.81	32.50	105169
Change in *Lend*	0.00	0.00	0.03	-0.03	0.05	0.08	46.12	105169
Utilization	17.42	9.55	19.81	1.45	64.43	1.86	3.12	105169
Average Tenure	88.73	70.23	77.26	17.14	217.97	3.82	31.11	105167
# Owners	191.81	124.00	221.06	21.00	606.00	3.43	17.83	105169
IOC_HHI	0.09	0.06	0.09	0.03	0.26	3.60	17.78	105169
log (Market Cap)	20.47	20.36	1.83	17.55	23.63	0.15	-0.03	105169
BtoM	3.14	3.02	1.34	1.13	5.58	0.58	0.68	105169
Past Year Return	0.15	0.12	0.64	-0.70	1.05	4.51	83.73	105167
Idiosyncratic Vol	0.06	0.05	0.05	0.02	0.13	9.44	296.05	103559
Retail Trade	0.00	0.00	0.02	-0.01	0.01	19.16	1244.01	85368
Institution Trade	0.00	0.00	0.02	-0.02	0.03	5.52	449.97	32420
Trade>50K	-0.02	-0.01	0.22	-0.37	0.30	-3.18	154.95	96977

Table 2. Impact of Stock/Security Loan Characteristics on the Ratio of Lendable share and Institutional Ownership. This table presents the associations between the ratio of quarterly lendable shares and institutional ownership and security loan and stock characteristics as defined in Table 1. Institutional ownership, defined as the split-adjusted number of shares owned by institutional investors, is obtained from 13F filings. The values of lendable shares are obtained from Markit. Both Institutional ownership and lendable shares are normalized relative to the total number of shares outstanding obtained from CRSP. Fama-MacBeth regressions are applied with the following specifications. Specification (1) regresses ***Lend/IO*** on ***Lend/IO*** in previous quarter; Specification (2) regresses ***Lend/IO*** on the institutional characteristics including **# Owners**, **IOC_HHI**, and **IO**; specification (3) regresses ***Lend/IO*** on firm characteristics including **Market Cap**, **BtoM**, **Past Year Return**, and **Idiosyncratic Vol**; specification (4) regresses ***Lend/IO*** on all characteristics above, respectively. The sample includes all NMS-listed common shares covered by Markit in 2007-Q4 through 2021-Q4. Standard errors are Newey-West adjusted with 3 lags. The numbers in parentheses are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent variable = ***Lend/IO***			
	(1)	(2)	(3)	(4)
Lagged *Lend/IO*	0.858***			
	(64.86)			
# Owners		0.000*		0.000
		(1.94)		(1.17)
IOC_HHI		−0.455***		−0.447***
		(−30.86)		(−22.27)
IO		−0.114***		−0.120***
		(−10.35)		(−13.57)
Market Cap			0.000	−0.001
			(0.05)	(−0.27)
BtoM			0.011***	0.010***
			(9.45)	(10.92)
Past Year Return			0.002	−0.001
			(0.79)	(−0.34)
Idiosyncratic Vol			−0.231***	−0.152**
			(−3.88)	(−2.54)
Observations	103,557	103,557	103,557	103,557
Number of groups	57	57	57	57
Adjusted R-squared	0.74	0.11	0.05	0.13

Table 3. Correlations Between Changes in Institutional Ownership and Measures of Institutional Flow. This table presents the associations between the quarterly changes in institutional ownership (***dIO***), defined as the split-adjusted number of shares owned by institutional investors, obtained from 13F filings and four daily measures of institutional flow aggregated at the stock-quarter level, using panel regression estimates of equation 1. Panel A reports the correlation with the corresponding change in the number of lendable shares (***dLend***) obtained from Markit in 2007-Q4 through 2021-Q4, divided by ***Lratio*** (defined as ***IO/Lend***). Panel B reports the correlation with the corresponding imbalance between buyer- vs. seller- initiated internalized retail trades obtained from TAQ in 2010-Q1 through 2021-Q4. Panel C reports the correlation between the change in IO and the corresponding institutional order flow obtained from ANcerno in 2010-Q1 through 2014-Q3. Panel D reports the correlation with the corresponding imbalance between buyer- vs. seller- initiated trades with dollar volumes of at least $50,000 obtained from TAQ in 2007-Q4 through 2021-Q3. Panel E reports the correlation with all the four measures of institutional flow in 2010-Q1 through 2014-Q3. In each panel, Columns 1–3 report results from panel regressions, where the change in IO serves as the dependent variable and the institutional flow measure serves as the independent variable. Column 4 presents results from a Fama-MacBeth regression using the same dependent and independent variables. The sample includes all NMS-listed common shares. The standard error is clustered by firms. The numbers in parentheses are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Change in *IO* vs. change in *Lend*

	(1)	(2)	(3)	(4)
dLend/Lratio	0.384***	0.345***	0.386***	0.406***
	(29.48)	(26.25)	(28.24)	(14.35)
Firm FE	No	Yes	No	N/A
Time FE	No	No	Yes	N/A
Observations	105,169	105,169	105,169	105,169
Adj-R^2	0.15	0.19	0.19	0.16

Panel B: Change in *IO* vs. minus institutional flow (BJZZ)

	(1)	(2)	(3)	(4)
Retail_trade	−0.005	−0.010	−0.005	−0.105***
	(−0.68)	(−1.13)	(−0.68)	(−3.52)
Firm FE	No	Yes	No	N/A
Time FE	No	No	Yes	N/A
Observations	86,761	86,761	86,761	86,761
Adj-R^2	0.00	0.09	0.04	0.00

Panel C: Change in *IO* vs. institutional flow (Ancerno)

	(1)	(2)	(3)	(4)
Institution_Trade	0.193***	0.186***	0.190***	0.210***
	(8.17)	(7.86)	(7.90)	(11.54)
Firm FE	No	Yes	No	N/A
Time FE	No	No	Yes	N/A
Observations	32,771	32,771	32,771	32,771
Adj-R^2	0.04	0.10	0.12	0.05

Panel D: Change in *IO* vs. innstitutional flow ($50k+ trades)

	(1)	(2)	(3)	(4)
Trade>50K	0.007	0.015***	0.011**	0.014**
	(1.36)	(3.08)	(2.32)	(2.30)
Firm FE	No	Yes	No	N/A
Time FE	No	No	Yes	N/A
Observations	98,143	98,143	98,143	98,143
Adj-R^2	0.00	0.08	0.05	0.00

Panel E: Change in *IO* vs. different measures of institutional flow

	(1)	(2)	(3)	(4)
dLend/Lratio	0.278***	0.234***	0.302***	0.340***
	(12.19)	(10.23)	(12.31)	(12.58)
Retail Trade	−0.036*	−0.057	−0.039*	−0.217***
	(−1.84)	(−1.60)	(−1.91)	(−3.49)
Institution_Trade	0.158***	0.168***	0.151***	0.157***
	(8.80)	(8.11)	(8.58)	(12.35)
Trade>50K	−0.003	0.002	−0.003	−0.001
	(−0.30)	(0.18)	(−0.35)	(−0.07)
Firm FE	No	Yes	No	N/A
Time FE	No	No	Yes	N/A
Observations	30,700	30,700	30,700	30,700
Adj-R^2	0.12	0.15	0.21	0.18

Table 4. Correlations Between Changes in Institutional Ownership and Changes in Lendable Shares: Conditional. This table presents Fama-MacBeth estimation results of equation (1), with X^s being the standardized $dLend/Lratio$, conditional on end-of-quarter security loan and stock characteristics defined in Table 1. For each characteristic and in each quarter, the sample is sorted into two equally-large subsamples. The sample includes all NMS-listed common shares covered by Markit in 2007-Q4 through 2021-Q4. Standard errors are Newey-West adjusted with 3 lags. The numbers in parentheses are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Correlations Conditional on Loan Characteristics

	Lend/IO		**Utilization**		**Average Tenure**	
	High	Low	High	Low	High	Low
dLend/Lratio	0.408***	0.419***	0.426***	0.376***	0.313***	0.449***
	(16.06)	(13.69)	(16.40)	(10.28)	(13.76)	(13.99)
Observations	52,599	52,570	52,599	52,570	52,597	52,570
Number of quarters	57	57	57	57	57	57
Adjusted R-squared	0.16	0.17	0.17	0.13	0.09	0.19

Panel B: Correlations Conditional on Institutional Characteristics

	# Owners		**IOC_HHI**		**IO**	
	High	Low	High	Low	High	Low
dLend/Lratio	0.431***	0.399***	0.368***	0.464***	0.425***	0.398***
	(15.89)	(12.20)	(11.05)	(15.24)	(12.92)	(14.76)
Observations	52,595	52,574	52,599	52,570	52,599	52,570
Number of quarters	57	57	57	57	57	57
Adjusted R-squared	0.18	0.16	0.13	0.21	0.17	0.15

Panel C: Correlations Conditional on Firm Characteristics

	log (Market Cap)		**BtoM**		**Past Year Return**		**Idiosyncratic Vol**	
	High	Low	High	Low	High	Low	High	Low
dLend/Lratio	0.454***	0.386***	0.371***	0.433***	0.445***	0.357***	0.410***	0.364***
	(15.08)	(12.45)	(12.98)	(13.28)	(14.05)	(13.18)	(13.38)	(13.73)
Observations	52,599	52,570	52,599	52,570	52,599	52,568	51,796	51,763
Number of quarters	57	57	57	57	57	57	57	57
Adjusted R-squared	0.20	0.14	0.13	0.18	0.19	0.12	0.16	0.13

Table 5. Predictive Power of Changes in Lendable Shares and Existing Institutional Flow Measures for Changes in Institutional Ownership. This table presents the out-of-sample performance of various proxies of institutional trading to predict the cross-section of actual institutional trading. Each quarter, the actual change in institutional ownership is regressed on its predicted change in institutional ownership ($\widehat{dIO}$), based on a proxy of institutional trading. Average R^2 of each proxy is calculated across quarters. The first four rows in Panel A present predictive average R^2's for **Retail Trade**, **Institution Trade**, **Trade>50K**, or ***dLend/Lratio***, where $\widehat{dIO}$ is constructed using equation (2). The last row in Panel A presents the predictive average R^2 when $\widehat{dIO}$ is constructed using equation (3), interacting $dLend$ with the following characteristics from the previous quarter: **Lratio**, **Utilization**, **Change in Utilization**, **Average Tenure**, **# Owners**, **IOC_HHI**, **log (Market Cap)**, **BtoM**, **Past Year Return**, **Log(Institutional Holdings)**, **Idiosyncratic Vol**, **Utilization**, **Change in Utilization**, and **Average Tenure**. All characteristics are defined in Table 1. Panel B presents average predictive R^2s when machine learning algorithms are used to form $\widehat{dIO}$ based on $dLend$ and characteristics, as in equation (4). The model is trained using Elastic Net, Random Forest, Gradient Boosting, and ensemble methods of the three. In the training samples of predictions based on machine learning, the top *and* bottom x% of $dLend_{jq}/Lratio_{q-1}$ are excluded, with $x \in \{0.5, 1, 2.5., 5\}$

Panel A: Out-of-Sample correlations between dIO_q and $\widehat{dIO}_q$ predicted by *Ins_Flow* using OLS

Institutional flow measure (Ins_Flow)	**Predictive %R^2**
BJZZ flow	0.34
Ancerno flow	5.80
$50K+ flow	0.29
***dLend*/Lratio**	13.80
Multivariate OLS with *dLend*	17.70

Panel B: Out-of-Sample correlations between dIO_q and $\widehat{dIO}_q = Elasticity(Chars) \times dLend_q$

	Predictive %R^2s			
	Trim the highest and lowest y percent elasticities			
Estimation method	$y = 5\%$	$y = 2.5\%$	$y = 1\%$	$y = 0.5\%$
Ensemble of Enet and RF	18.46	17.80	16.37	14.54
Ensemble of Enet, RF, and GBRT	18.12	17.82	17.05	15.37
Elastic Net (Enet)	15.98	16.29	16.39	14.91
Random Forest (RF)	18.45	17.60	15.93	13.94
Gradient Boosting (GBRT)	18.19	17.69	16.32	15.04

Table 6. Determinants of Daily Institutional Trading. This table documents the association between daily institutional trading activity and past daily outcomes. Panel A and B examine these associations for $dLend_t/Lratio_{q-1}$ and $|dLend_t/Lratio_{q-1}|$, respectively, reporting estimates of equation (5). Each institutional trading variable from day t, scaled by 100, is regressed on its lags from days $t-5$ through $t-1$; as well as a combination of the following variables: compound returns over the preceding 5 trading days, $cret_{t-5}^{t-1}$, the 5 days before them, $cret_{t-10}^{t-6}$, the 20 days before them, $cret_{t-30}^{t-11}$, the 95 days before them, $cret_{t-125}^{t-31}$, and the 125 trading days before them, $cret_{t-250}^{t-126}$; previous month-end's log market-capitalization (SIZE) and the book-to-market ratio (BM); previous month's open-to-close Amihud liquidity measure (OCAM) of Barardehi et al. (2021) and daily return standard deviation (Volatility); and the current month's aggregate mispricing factor (MISP) of Stambaugh et al. (2012). Estimates reflect Fama-MacBeth regression of daily cross-sections with Newey-West standard errors based on 30 lags. Daily cross-section of each vriable in winsorized at percentiles 1 and 99. The numbers in parenthesis reflect t-statistics, and symbols ***, **, and * identify statistical significance at the 1%, 5%, and 10% type one errors, respectively. The sample includes all NMS-listed common shares in 2007-Q4 through 2022-Q1, but—reflecting the availability of MISP—-limited to 2007-Q4 through 2013-Q4 when MISP is used.

	Panel A: Dependent variable: $Y = dLend/Lratio$				Panel B: Dependent variable: $Y = \|dLend/Lratio\|$			
Constant	0.048***	0.038***	0.085***	−0.027	0.38***	0.35***	0.36***	0.65***
	(8.03)	(6.06)	(2.70)	(−0.36)	(47.13)	(55.34)	(8.75)	(8.37)
Y_{t-1}	−0.46***	−0.47***	−0.47***	−0.44***	0.34***	0.33***	0.33***	0.30***
	(−83.57)	(−80.95)	(−81.23)	(−43.94)	(99.07)	(98.11)	(94.81)	(45.51)
Y_{t-2}	−0.21***	−0.23***	−0.23***	−0.20***	0.060***	0.057***	0.049***	0.046***
	(−40.48)	(−37.61)	(−37.36)	(−22.18)	(26.15)	(24.95)	(22.09)	(10.85)
Y_{t-3}	−0.10***	−0.12***	−0.12***	−0.11***	0.097***	0.094***	0.087***	0.078***
	(−22.51)	(−24.79)	(−24.62)	(−14.66)	(58.54)	(58.57)	(50.31)	(40.52)
Y_{t-4}	−0.049***	−0.057***	−0.055***	−0.050***	0.079***	0.076***	0.068***	0.064***
	(−15.19)	(−16.61)	(−17.16)	(−10.95)	(35.38)	(33.47)	(29.66)	(33.78)
Y_{t-5}	−0.017***	−0.020***	−0.019***	−0.017***	0.10***	0.099***	0.088***	0.079***
	(−8.11)	(−9.62)	(−9.78)	(−4.86)	(73.09)	(70.43)	(64.96)	(36.78)
$cret_{t-5}^{t-1}$		1.61***	1.60***	1.61***		−0.099***	−0.11***	−0.13***
		(28.60)	(28.56)	(18.19)		(−4.19)	(−5.02)	(−4.28)
$cret_{t-10}^{t-6}$		0.22***	0.23***	0.15***		−0.079***	−0.10***	−0.063
		(9.56)	(9.94)	(3.91)		(−3.21)	(−4.63)	(−1.46)
$cret_{t-30}^{t-11}$		0.069***	0.065***	0.056***		−0.046***	−0.093***	−0.027
		(6.18)	(6.38)	(3.71)		(−2.69)	(−6.61)	(−1.11)
$cret_{t-125}^{t-31}$		0.099***	0.10***	0.12***		−0.015	−0.034***	0.0092
		(15.49)	(16.42)	(13.12)		(−1.40)	(−3.89)	(0.51)
$cret_{t-250}^{t-126}$		0.042***	0.050***	0.040***		0.0076	−0.012*	−0.0061
		(7.41)	(9.00)	(4.85)		(0.98)	(−1.89)	(−0.49)
SIZE			−0.0033**	−0.0025			−0.0039*	−0.018***
			(−2.21)	(−0.74)			(−1.83)	(−4.93)
BM			−0.000012	−0.0000048			−0.00012***	−0.000072***
			(−1.46)	(−0.41)			(−13.23)	(−7.13)
TO			0.00088***	0.00070**			0.0076***	0.0093***
			(4.52)	(2.03)			(20.92)	(21.34)
OCAM			−0.0091*	−0.032*			−0.15***	−0.34***
			(−1.79)	(−1.75)			(−21.11)	(−10.39)
Volatility			0.61***	0.51***			2.81***	3.34***
			(6.79)	(3.26)			(23.01)	(15.85)
MISP				0.0015***				0.00033***
				(11.16)				(2.64)
R-squared	0.215	0.235	0.252	0.240	0.250	0.260	0.275	0.252
Observations	6,054,524	6,054,524	6,054,524	1,938,596	6,054,524	6,054,524	6,054,524	1938596

Table 7. Three-Factor Alphas of Portfolios Based on Past Intraday Signals Conditional on the Signs of Past Intraday Returns and *dLend*: CRSP Breakpoints This table reports three-factor alphas of trading strategies that buy past winners and sell past losers according to compound past intraday return (IDR). Following Barardehi et al. (2024), in each month m, a stock's signals are constructed using intraday returns from months $m-12$ through $m-2$. The trading days underlying the past return signals are decomposed into those where IDR and $dLend$ have the same sign and those where IDR and $dLend$ have opposite signs. The time-series equally-weighted month m returns of each of the ten portfolios (deciles with CRSP breakpoints), net of the 1-month T-Bill trade, as well as the high-minus-low (10−1) portfolio returns are regressed on market, size, and value factor returns to estimate the corresponding three-factor alpha, i.e., the intercept. Standard errors are Newey-West adjusted with 3 lags, and the numbers in parentheses are t-statistics. For either portfolio sort, the average number of trading days used to form portfolios; the average past intraday return; the average overnight return following each relevant intraday signal periods ($FONR$); average $dLend/Lratio$ on signal days; and average $dLend/Lratio$ on non-signal days, from months $m-12$ through $m-2$. The sample includes all NMS-listed common shares covered by Markit in 2007-Q4 through 2021-Q4.

	Portfolios of months $m-12$ to $m-2$ intraday returns										
Signal	1	2	3	4	5	6	7	8	9	10	10−1
sign(IDR) = sign($dLend$)											
Month-m alpha	−0.26	0.066	0.42***	0.21*	0.15	0.16*	0.055	0.18**	0.17	0.24*	0.51
	(−0.70)	(0.33)	(2.63)	(1.70)	(1.47)	(1.96)	(0.76)	(2.23)	(1.46)	(1.68)	(1.32)
Signal-day count	113.15	111.33	108.93	106.53	105.29	105.68	108.21	111.03	113.43	114.87	1.72
Signal-day compound IDR	−72.43	−26.36	−13.37	−5.76	0.05	5.31	10.75	17.06	26.06	50.97	123.40
Signal-day compound $FONR$	28.59	9.82	6.00	4.52	3.69	2.90	2.60	2.31	1.57	−0.31	−28.89
Aggregate $dLend/Lratio$	−0.01	0.02	0.03	0.03	0.03	0.04	0.05	0.06	0.09	0.15	0.16
sign(IDR) ≠ sign($dLend$)											
Month-m alpha	−0.63*	0.047	0.20	0.21*	0.18**	0.24**	0.23***	0.29***	0.28***	0.33***	0.95***
	(−1.88)	(0.22)	(1.41)	(1.79)	(2.26)	(2.39)	(2.88)	(3.23)	(2.80)	(2.66)	(2.68)
Signal-day count	120.35	116.79	113.87	110.82	108.77	107.14	107.23	109.32	112.96	117.57	−2.78
Signal-day compound IDR	−75.47	−33.22	−20.08	−12.01	−5.90	−0.66	4.28	9.72	17.05	37.89	113.35
Signal-day compound $FONR$	33.61	12.26	8.04	5.99	4.33	3.05	1.92	0.95	−0.24	−6.63	−0.33
Aggregate $dLend/Lratio$	0.08	0.08	0.07	0.06	0.05	0.04	0.03	0.03	0.03	0.02	−0.06

Table 8. Three-Factor Alphas of Portfolios Based on Past Intraday Signals Conditional on the Signs of Past Intraday Returns and *dLend*: NYSE Breakpoints This table reports three-factor alphas of trading strategies that buy past winners and sell past losers according to average past intraday return (IDR). Following Barardehi et al. (2024), in each month m, a stock's signals are constructed using intraday returns from months $m-12$ through $m-2$. The trading days underlying the past return signals are decomposed into those where IDR and $dLend$ have the same sign and those where IDR and $dLend$ have opposite signs. The time-series equally-weighted month m returns of each of the ten portfolios (deciles with NYSE breakpoints), net of the 1-month T-Bill trade, as well as the high-minus-low (10−1) portfolio returns are regressed on market, size, and value factor returns to estimate the corresponding three-factor alpha, i.e., the intercept. Standard errors are Newey-West adjusted with 3 lags, and the numbers in parentheses are t-statistics. For either portfolio sort, the average number of trading days used to form portfolios; the average past intraday return; the average overnight return following each relevant intraday signal periods ($FONR$); average $dLend/Lratio$ on signal days; and average $dLend/Lratio$ on non-signal days, from months $m-12$ through $m-2$. The sample includes all NMS-listed common shares covered by Markit in 2007-Q4 through 2021-Q4.

	Portfolios of months $m-12$ to $m-2$ intraday returns										
Signal	1	2	3	4	5	6	7	8	9	10	10−1
sign(IDR) = sign($dLend$)											
month-m alpha	−0.087	0.29*	0.38***	0.18	0.14	0.17*	0.050	0.15*	0.20*	0.19	0.28
	(−0.29)	(1.79)	(3.00)	(1.59)	(1.22)	(1.97)	(0.60)	(1.89)	(1.96)	(1.47)	(0.84)
Signal-day count	112.65	109.94	107.73	105.60	104.85	105.75	107.84	110.34	112.93	114.64	1.98
Signal-day compound IDR	−58.50	−18.48	−9.37	−3.59	1.09	5.37	9.85	15.07	22.35	45.34	103.84
Signal-day compound $FONR$	22.80	7.49	5.09	4.12	3.54	3.00	2.64	2.49	1.93	0.42	−22.38
Aggregate $dLend/Lratio$	0.00	0.02	0.03	0.03	0.03	0.04	0.05	0.06	0.08	0.14	0.14
sign(IDR) $\neq$ sign($dLend$)											
month-m alpha	−0.42	0.27*	0.15	0.22*	0.29***	0.21**	0.27***	0.28***	0.20**	0.36***	0.78**
	(−1.49)	(1.73)	(1.38)	(1.94)	(2.77)	(2.42)	(2.93)	(3.15)	(1.98)	(3.36)	(2.52)
Signal-day count	119.24	114.88	112.06	109.52	107.72	106.77	107.21	108.94	112.33	116.94	−2.30
Signal-day compound IDR	−61.82	−24.37	−14.58	−8.37	−3.49	0.77	4.85	9.44	15.58	35.03	96.85
Signal-day compound $FONR$	26.63	9.37	6.66	4.96	3.75	2.64	1.73	1.06	0.11	−5.78	−0.33
Aggregate $dLend/Lratio$	0.08	0.07	0.06	0.05	0.04	0.04	0.03	0.03	0.03	0.03	−0.06

Table 9. Return Predictability of the Daily Change in Lendable Shares. This table reports on the return predictability of daily changes in lendable equity. On each day t in quarter q, stocks are sorted into 10 groups based on the average of $dLend_{jt}/Lratio_{q-1}$ on day $t-1$. High-minus-low cumulative returns are constructed for 1-, 2-, 3-, 5-, and 10-day horizons. Panels A and B present, respectively, equally-weighted and value-weighted cumulative returns with and without Fama-French 5-factor risk-adjustments. The sample period is from 04/01/2013 to 12/31/2021, excluding stocks not covered by Markit. The numbers in parentheses are t-statistics using Newey-West standard error with 30 lags, with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Long minus Short Returns of all sample - Equally weighted

	1 day	**2 days**	**3 days**	**5 days**	**10 days**
Average return	−7.57***	−15.73***	−23.41***	−27.92***	−35.98***
	(−5.77)	(−4.78)	(−6.48)	(−6.08)	(−5.89)
FF5 Alpha	−7.54***	−15.70***	−23.13***	−26.11***	−31.45***
	(−5.78)	(−4.59)	(−6.10)	(−5.61)	(−5.88)

Panel B: Long minus Short Returns of all sample - Value weighted

	1 day	**2 days**	**3 days**	**5 days**	**10 days**
Average return	−5.29***	−10.99***	−16.35***	−16.39***	−19.30***
	(−4.07)	(−4.32)	(−5.27)	(−4.04)	(−3.80)
FF5 Alpha	−4.89***	−10.12***	−15.26***	−15.00***	−17.77***
	(−3.87)	(−4.09)	(−5.01)	(−3.99)	(−3.94)

Table 10. Return Predictability of the Predicted Daily Change in Institutional Ownership. This table reports on the return predictability of daily changes in predicted daily institutional trading, $\widehat{dIO_{jt}}$. Predictions are forms based on equation (2). On each day t in quarter q, stocks are sorted into 10 groups based on the average of $\widehat{dIO_{jt}}$ on days $t-1$. High-minus-low cumulative returns are constructed for 1-, 2-, 3-, 5-, and 10-day horizons. Panels A and B present, respectively, equally-weighted and value-weighted cumulative returns with and without Fama-French 5-factor risk-adjustments. The sample period is from 04/01/2013 to 12/31/2021, excluding stocks not covered by Markit. The numbers in parentheses are t-statistics using Newey-West standard error with 30 lags, with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Long minus Short Returns of all sample - Equally weighted

	1 day	**2 days**	**3 days**	**5 days**	**10 days**
Average return	−7.92***	−16.68***	−22.66***	−26.91***	−33.38***
	(−6.47)	(−5.39)	(−6.73)	(−6.19)	(−5.85)
FF5 Alpha	−7.90***	−16.83***	−22.18***	−24.89***	−28.96***
	(−6.52)	(−5.26)	(−6.49)	(−5.79)	(−5.79)

Panel B: Long minus Short Returns of all sample - Value weighted

	1 day	**2 days**	**3 days**	**5 days**	**10 days**
Average return	−6.48***	−13.25***	−17.68***	−18.97***	−21.63***
	(−5.15)	(−5.30)	(−6.31)	(−5.62)	(−4.71)
FF5 Alpha	−6.15***	−12.58***	−16.52***	−16.84***	−19.97***
	(−4.94)	(−5.04)	(−6.09)	(−5.25)	(−4.40)