Detecting Informed Trading Risk from Undercutting Activity in Limit Order Markets

Abstract

We develop simple measures of informed trading risk ($QIDRes$) that capture abnormal undercutting activity, reflecting the intuition that liquidity-providing algorithms compete less to fill incoming marketable orders when adverse selection exposure rises. Despite being conveniently computed from TAQ data, when examined around major information events, $QIDRes$ behaves similarly to existing measures of informed trading intensity/probability whose constructions are complex and demanding. Our measure predicts both arrivals and magnitudes of imminent information events. Moreover, episodes of high $QIDRes$ coincide with weaker subsequent price reversals, increased accumulation/covering of short interest, and more likely informed institutional trades. $QIDRes$ positively predicts stock returns up to six months forward, especially among stocks with tighter short sale constraints. Since $QIDRes$ is by construction orthogonal to stock liquidity and does not constitute a persistent stock characteristic, we attribute its return predictability to limits to arbitrage.

JEL Classification Codes: G14

Keywords: Informed Trading, Undercutting, Asset Pricing, Liquidity, Limits to Arbitrage

1 Introduction

Informed trading is a key concept in financial economics with implications for market efficiency, trading mechanisms, the cost of capital, etc. and interests academics and regulators alike. However, informed investors usually must conceal their intended trades before establishing desired positions to maximize rents from private information (see, e.g., Kyle (1985), Anand, Irvine, Puckett, and Venkataraman (2012)). Thus, measuring the extent of informed trading presence is inherently challenging, rendering many proposed measures ineffective (Ahren (2020)) and others computationally demanding to construct. We propose an easy-to-compute, intuitive measure of nondirectional informed trading risk that only requires trades and quotes data and thus can be computed for any security at the daily, or even finer, frequencies traded in any modern limit order market.

Our approach exploits the intuition that liquidity providers compete less over filling marketable orders they perceive to be informed. Specifically, a liquidity provider's willingness to "undercut" rivals should significantly drop when they expect arrivals of informed marketable orders. Hence, abnormally low undercutting activity reveals the extent of liquidity providers' concerns about anticipated incoming informed orders and hence is an indirect measure of informed trading risk.

The existing literature uses proprietary account level data to identify undercutting behavior as undercutting runs, or just runs, (Foley, Dyhrberg, and Svec (2022), Foley, Meling, and Ødegaard (2021)). Runs occur when multiple trading algorithms repeatedly undercut one another to get their quote to the front of the order book and to fill anticipated incoming marketable order flow. In the data, this translates to a sequence of rapid single tick improvements in the best quoted price on one side of the market followed by a sudden drop back to the pre-run prices as an incoming marketable order crosses the winner of the undercutting run's resetting quote. Put differently, undercutting runs are associated with successive best quote improvements that are followed a trade-driven quote deterioration. We relate undercutting activity to informed trading and seek to capture the prevalence and length of runs using aggregate market, rather than proprietary, data.

Increased informed trading risk negatively affects undercutting runs for at least two reasons. First, providing liquidity is only profitable when liquidity provider's counter parties are not privately information about fundamental asset values. Put differently, informed traders will only transact when the asset is mispriced, so transactions with informed traders can lead to losses for the liquidity provider, often referred to as adverse selection costs (see, e.g., Kyle (1985), Glosten and Milgrom (1985)). In modern limit order markets, liquidity-providing algorithms are not affirmatively obligated to continuously provide liquidity. Thus, when anticipating incoming informed, or 'toxic,' order flow, liquidity-providing algorithms can simply bow out and wait until order flow becomes less informed. Fewer algorithms competing to provide liquidity implies fewer runs. Second, when an increasing fraction of order flow is informed runs will stop sooner. This occurs because increased adverse selection costs limits how far an algorithm is willing to improve the quote in a run before determining that a trade is no longer profitable. Importantly, these algorithms operate with inventory holding horizons as short as a few seconds (Conrad and Wahal (2020)). Hence, when dodging directional informed flow expected to persist beyond these holding horizons, they limit providing liquidity and hence undercutting on *both sides* of the market to avoid unwanted inventory accumulation. In other words, despite the directional nature of informed trading, liquidity-providing algorithms react to increased informed trading risk by undercutting less on both sides of the market.

How to measure variations in undercutting using aggregate market data? In presence of undercutting runs, quote improvements must occur more frequently than trade-driven quote deteriorations, reflecting that each 'run' consists multiple quote improvements followed by one trade-driven quote deterioration. Such skewing is a systematic feature in the data. To capture this skewing, we calculate the QID ratio which is the total number of NBBO quote improvements observed on a given day minus the number of NBBO quote deteriorations associated with trades on that same day, all divided by the total number of NBBO changes of either type. The construction of QID imposes boundaries of -1 and 1, with near-zero QID reflecting modest undercutting.[1] Thus, QID moving closer to 1 signifies increases in undercutting runs; Figure 1 shows that *all* stocks feature positive average QID's, suggesting an overwhelming prevalence of undercutting runs.

We first exploit the exogenous variation in the cost of undercutting driven by the SEC's Tick Size Pilot program (TSP) to establish the validity of QID as a measure of undercutting. By temporarily raising the minimum tick size from 1¢ to 5¢ in a select group of stocks, the TSP quintupled the cost of undercutting born by liquidity-providing algorithms active in markets for these stocks (Werner, Rindi, Buti, and Wen (2022)). Using standard difference-in-difference analysis, we find that the

[1] Because (1) we exclude best quote deteriorations due to limit order cancellations and (2) executions of marketable orders likely lead to best quote deteriorations, we expect QID to be slightly negative in the absence of undercutting.

five-fold increase in the cost of undercutting when TSP was implemented *reduced* QID by about 0.44, which is quite close to the unconditional average of QID. Contrary to the TSP's effects on many other outcomes, this large effect on QID obtains regardless of how binding the 5¢ tick was. Moreover, TSP conclusion, which reduced the cost of undercutting in pilot stocks five-fold, caused an average *increase* of 0.42 in QID, virtually restoring its pre-TSP unconditional mean.[2] These findings confirm a remarkably strong inverse relationship between costs of undercutting and QID, indicating that QID is a sound measure of undercutting activity.

To interpret the variation in QID, one must distinguish between informed trading concerns and market liquidity. Importantly, wider dollar/relative bid-ask spreads in less liquid stocks create ample undercutting opportunities and lead to more prevalent runs. In fact, Figure 1 documents a positive association between QID and stock illiquidity, measured by relative quoted bid-ask spread.[3] Moreover, besides impacting liquidity provision strategies, e.g., undercutting, variations in liquidity can affect how informed traders transact, further highlighting the need to disentangle variations reflecting informed trading from those reflecting liquidity (Duarte and Young (2009)). To address these, we use a two-stage procedure to focus on *abnormal* undercutting activity, denoted $QIDRes$. For each quarter and each stock, we first fit a regression of daily QID on time-weighted relative bid-ask spread, accounting for the daily variation in liquidity at the stock level. Second, we apply the coefficients from the first stage to the following quarter's realizations to produce estimates of the unexpected (residual) QID, which we multiply by -1 to produce a positive, rather than an inverse, measure of informed trading risk. We also scale these stock-specific estimates by QID's intercept term from the previous quarter to remove any remaining cross-sectional variation associated with persistent stock characteristics such as liquidity. We dub the resulting ratio $QIDRes$ which we expect to rise when liquidity providers perceive informed trading risk to be higher.

We examine the behavior of $QIDRes$ around information events known to be associated with higher than average informed trading, including earnings announcements, unscheduled press re-

[2] Our analysis satisfies the heuristic hurdles when re-using experiments as all t-statistic in our TSP results range between 9–38, multiples of those proposed by Heath, Ringgenberg, Samadi, and Werner (2020).

[3] Relative quoted spread is particularly relevant for undercutting in U.S. equity markets. Dollar bid-ask spread together with the 1¢ tick size reflect the number of 1-¢-apart price levels potentially available for undercutting runs. However, the value per share of the stock, usually approximated by the quote midpoint in microstructure applications, together with the minimum lot size of 100 shares, required for any effective undercutting, reflect the minimum dollar value transferred per transaction as an undercutting run's winner trades. The minimum tick and lot size are fixed across all stocks, and relative bid-ask spread, defined as the ratio of dollar bid-ask spread to NBBO midpoint, controls for the two remaining relevant factors.

leases, and news arrivals. We document a significant spike in $QIDRes$ around such events that take up to 10 days to rebound. This is consistent with behaviors of other measures of informed trading risk such as Informed Trading Intensity (ITI) measures of Bogousslavsky, Fos, and Muravyev (2023); Probability of Informed Trading (PIN) measures—see Duarte, Hu, and Young (2020) for a discussion of the various PIN-based measures; and the multi-market information asymmetry measure of Johnson and So (2018). Consistent with the notion that market makers may learn from order flow about upcoming information events (Chae (2005)), we find using logistic regressions that increases in $QIDRes$ predict imminent arrivals of *unscheduled* information events. We further show that the magnitude and persistence of the spikes in $QIDRes$ are related to the size of the the post-event returns: information events with larger increases in $QIDRes$ are followed by larger post-event absolute returns; and for such events, post-event $QIDRes$ rebounds realize more slowly.

We then provide evidence against $QIDRes$ solely capturing 'sniping risk.' The literature pioneered by Budish, Cramton, and Shim (2015) shows in continuous-time limit order markets liquidity providers face adverse selection costs reflecting sniping risk, rather than information disadvantages of liquidity providers about fundamental values.[4] Relevant for our analysis is the intuition that liquidity providers should become reluctant to undercut when sniping risk rises due public news arrivals, leading to larger $QIDRes$ magnitudes around major information arrivals. To address this, instead of only relying on information events, we link $QIDRes$ to more direct sources of informed trading: (1) reflecting short-seller trades being often informed, $QIDRes$ is significantly higher during periods with large absolute changes in short interest, even after excluding periods that overlap with information events; (2) we also find significantly larger $QIDRes$ magnitudes on stock-days with informed mutual-fund trades, as identified by Barardehi, Da, and Warachka (2022), than on stock-days without. Other informed trading intensity/probability measures share these qualities.

We also provide evidence inconsistent with $QIDRes$ solely reflecting inventory management concerns of liquidity providers. Reflecting capital constraints, liquidity providers with unbalanced inventories avoid accumulating additional inventory or charge a premium to do so (Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010)). This can translate to reduced undercutting,

[4]With differences in order processing speeds across traders, the arrival of public news leaves some resting limit orders of slow trades stale not because there of information asymmetries but because these traders cannot cancel their orders fast enough. In turn, faster traders benefit from picking off (sniping) these stale orders at the loss of slow traders (also see Menkveld and Zoican (2017)).

i.e., higher $QIDRes$, by liquidity providers who now demand greater compensation for providing liquidity, with compensations manifested in short-term price pressure followed by price reversals (e.g., Hendershott and Menkveld (2014)). Contrary to this prediction, stock-days with higher $QIDRes$ are followed by *weaker* reversals. This finding also indicates that higher $QIDRes$ captures increased informed trading which should not precede price reversals (Bogousslavsky et al. (2023)). Moreover, stronger price reversals when undercutting activity is relatively higher is at odds with undercutting activity capturing informed trading via aggressive limit order submission strategies, where informed traders undercut others instead of removing liquidity through marketable orders.

We next document asset pricing implications of $QIDRes$. We first demonstrate that long-short portfolios that buy stocks with high $QIDRes$ and short stocks with low $QIDRes$ from two quarters earlier produce a statistically significant 4-factor alpha of 38 basis points per month. This finding extends the one-month return predictability of informed trading intensity measures, documented by Bogousslavsky et al. (2023), to longer horizons. Our additional asset pricing tests involve fixed-effect panel regressions that regress monthly excess returns on lagged $QIDRes$ and stock characteristics, including illiquidity measures. We find a positive association between monthly expected stock returns and $QIDRes$ from two quarters earlier, i.e., stocks with higher expected informed trading risk have higher returns. To highlight the incremental explanatory power of $QIDRes$ for expected returns, relative to existing informed trading intensity measures, we estimate "horse race" regressions that, in addition to $QIDRes$ from the preceding two quarters, include subsets of ten corresponding alternative measures (ITIs, PINs, and MIA) as independent variables. Not only $QIDRes$ maintains its explanatory power for expected returns when we control for existing measures, but $QIDRes$ is the *only* measure that significantly predicts future returns across all specifications.

Return predictability of informed trading risk as reflected in $QIDRes$ cannot be interpreted in the context of asset pricing theories such as Easley and O'Hara (2004) because $QIDRes$ *does not* constitute a stock characteristic. Specifically, $QIDRes$ exhibits no temporal persistence but rather displays modest mean reversion, if anything. In fact, reflecting its construction, $QIDRes$ should be orthogonal to persistent stock characteristics such as illiquidity, which we confirm empirically: in the cross-section, $QIDRes$ is minimally correlated with a host of stock characteristics as well as existing informed trading intensity proxies. For example, highlighting the contrast between $QIDRes$ and existing measures, the average absolute pairwise correlation coefficient between $QIDRes$ and five illiquidity measures is only 0.02; whereas the analogue for ITI and PIN measures is 0.15, with individual pairwise correlation coefficients as high as 0.53.

We interpret return predictability of $QIDRes$ in the context of insights from behavioral asset pricing literature and, more importantly, limits to arbitrage. First, long-only investors' buy trades are often motivated by positive information about fundamentals, whereas their selling activity often reflects pure liquidity trades (see, e.g., Akepanidtaworn, Di Mascio, Imas, and Schmidt (2020)). Second, short sellers who systematically investigate and trade on negative information face short sale constraints. To these ends, as Bogousslavsky et al. (2023) also argue, an informed trading risk measure is more likely to capture trading motivated by positive information, rather than negative information. Hence, increases in measures of informed trading risk such as $QIDRes$ should predict higher future returns. Our empirical findings confirm this. We control for short sale constraints (1) directly, using security lending fees, and (2) indirectly, using the level of institutional ownership adjusted for firm size (Nagel (2005)). Indeed, return predictability of $QIDRes$ concentrates among stocks with tight short sale constraints or those primarily owned by long-only-investors.

We contribute by developing an informed trading risk measure that is computed using aggregate frequencies of quote improvements and deteriorations. This simple construction offers several appealing features relative to existing measures: Our measures (1) are implementable for securities traded in any modern limit market; (2) do not require structural estimations as in, e.g., Easley, Hvidkjaer, and O'Hara (2002); (3) do not require hand-collected data and computationally demanding data-drive techniques as in Bogousslavsky et al. (2023); and (4) do not require observable trading activity in corresponding derivatives markets as in Johnson and So (2018).

Our methodology builds on the literature on order placement, including undercutting, strategies in modern limit order markets. Hasbrouck and Saar (2013) introduced the notion of 'strategic runs' to describe a sequence of order submission/cancellations by an *individual* trader. In this context, strategic runs that end with a trade may resemble successful undercutting efforts of an individual trader (Chordia and Miao (2020)). Bringing this idea to the market level, Foley et al. (2021) and Foley et al. (2022) directly examine undercutting 'runs' by identifying sequences of quote improvements, reflecting order submissions by *multiple* traders, that end with a trade. We posit that, in aggregate market data, best-quote improvements tend to capture undercutting efforts; whereas best-quote deteriorations due to trades tend to capture conclusions of undercutting runs. We measure aggregate undercutting intensity as the ratio of such quote improvements to quote deteriorations at the stock-day level. Intuitively, exposure to adverse-selection risk due to information asymmetry lowers liquidity providers' willingness to undercut, leading us to propose abnormally low undercutting as a new measure of increased informed trading risk.

We provide new evidence relevant for the debate about asset pricing implications of informed trading as $QIDRes$ predicts returns six months forward. Easley and O'Hara (2004) predict more frequent informed trading commands higher expected stock returns, with Easley et al. (2002), Kelly and Ljungqvist (2012), and Derrien and Kecskés (2013) providing supportive evidence in different settings. Hughes, Liu, and Liu (2007) and Petacchi (2015), respectively, link more frequent informed trading to higher cost of capital and higher cost of equity. However, Lambert, Leuz, and Verrecchia (2012) predict these links only exist in noncompetitive capital markets, with Armstrong, Taylor, Core, and Verrecchia (2011) providing supportive empirical evidence. In contrast, Wang (1993) posits that increased presence of informed investors reduces the cost of capital. Relatedly, Duarte and Young (2009) show that the ability of Easley et al. (2002)'s PIN measures to explain expected returns reflects the cross-sectional variation in liquidity, rather than that in prevalence of informed trading. However, our measure of informed trading risk is, by construction, unrelated to stock liquidity. In fact, our measure does not constitute a persistent stock characteristic, leading us to attribute its return predictability to limits to arbitrage such as short sale constraints.

2 Data and Methodology

2.1 Data

Our main sample runs from January 2010 thought December 2019 and includes NMS-listed common stocks whose share prices were at least $5 at the end of the preceding month. We obtain intraday quote and trade information from Daily TAQ; daily microstructure outcomes from WRDS Intraday Indicators; daily and monthly price and trade information from Daily and Monthly CRSP, respectively; Book-value information and earnings announcements dates from COMPUSTAT; earnings surprise scores from I/B/E/S; and news information from Ravenpack.

Our daily measure of undercutting, QID_{jt}, divides the difference between the number of best quote improvements, on either bid or ask side, and the number of trade-driven best quote deterioration, on either bid or ask side, by the total number of NBBO updates for stock j on day t. NBBOs are constructed following Holden and Jacobsen (2014) by merging Daily TAQ's NBBO and Quote files that are then matched with trades in the same millisecond obtained from Daily TAQ's Trade files. We match QID_{jt} with daily time-weighted dollar spreads (denoted qsp_{jt}) and percent quoted spreads (denoted psp_{jt}) as well as percent effective spreads (denoted $pefsp_{jt}$), realized spreads (denoted $prsp_{jt}$), price impacts (denoted $primp_{jt}$), regular-hour trading volume (denoted tv_{jt}), and volatility of 1-minute quote-midpoint returns (denoted $qvol_{jt}$) obtained from WRDS Intraday Indicators. We also match them with daily returns (denoted r_{jt}), reflecting quote midpoints at close, and trading volumes from Daily CRSP.[5] The CRSP-TAQ linking table provided by WRDS facilitates these mergers.

We then merge our daily data base with earnings announcements (EA), unscheduled corporate events (PR), and news arrivals unassociated with identifiable corporate events (NA), using the announcements' timing to identify the first trading day where trading takes place after an announcement. Earnings announcement dates are obtained from COMPUSTAT. Reflecting the findings of [cite] that the vast majority of such announcements arrive outside regular trading hours, we designate the trading day after the recorded announcement date as the effective announcement date. We obtain dates and timestamps of unscheduled press releases and news arrivals from Ravenpack. For press releases, we focus on Ravenpack "full-article" or "news-flash" observations with "news_relevance" scores of at least 90. For news arrivals, we focus on Ravenpack "full-article" or "news-flash" observations with "news_relevance" scores of at least 95 and no recorded "event_relevance" score. We construct event windows that span the 10 days prior to an announcement and 10 days after the announcement.[6]

We construct a set of stock characteristics for our asset pricing analysis using data from CRSP, COMPUSTAT, and 13F. For stock j in month m, $RET_{j,m-1}$ and $RET_{j,m-2}^{m-12}$, respectively, capture compound returns over the preceding month and the 11 months prior; $M_{j,m-12}$ is market-capitalization based on the closing price 12 months earlier; $DYD_{j,m-1}$ is dividend yield, i.e., the ratio of total dividend distributions over the 12 months ending in month $m-1$ divided by the closing price at the end of month $m-1$. The book-to-market ratio, $BM_{j,m-1}$, is the most recently reported

[5] Our daily return calculations account for dividend distributions and overnight adjustments such as stock splits.

[6] For each announcement type (EA, PR, or NA), we focus on the first announcement should multiple announcements cluster over a 20 day period. This endures non-overlapping event windows.

book value divided by market capitalization at the end of month $m-1$.[7] We obtain three-factor Fama-French betas for each stock from Beta Suite by WRDS. Our approach employs weekly data from rolling horizons that span the preceding 104 weeks, requiring a minimum of 52 weeks. For each stock month, the set of betas represent estimates from the estimation horizon ending in the last week of that month. As in Ang, Hodrick, Zhing, and Zhang (2006), we use a CAPM regression using daily observations in each month to construct monthly idiosyncratic volatility measures. We match each monthly observation with previous calendar quarter's fraction of institutionally owned shares outstanding ($IOShr$) and the concentration of such ownership based on a Herfindahl-Hirschman index ($IOShrHHI$) using 13F data.[8]

To control for stock illiquidity in each month m, we use five liquidity measures constructed using daily or intraday obsrvations from month $m-2$: (1) time-weighted dollar quoted spreads (QSP); (2) size-weighted dollar effective spread ($EFSP$); (3) monthly estimates of Kyle's λ, constructed by regressing 5-minute returns (calculated from quote midpoints) on the contemporaneous signed square root of net order flow (estimated using the Lee-Ready algorithm) from the respective month ($Lambda$); (4) a modified version of Amihud (2002)'s measure (AM);[9] and (5) Barardehi, Bernhardt, Da, and Warachka (2023)'s retail-based institutional liquidity measure ($ILMV$). We also construct turnover ratio (TO), defined as the average daily fraction of share volume to shares outstanding using observations from month $m-2$.

Finally, we obtain lending fee observations at the stock-day level for the 2009-2018 period from Financial Information Service (FIS) Astec Analytics. FIS compiles dollar-weighted average stock lending fees at daily frequencies. For each stock, we aggregate these lending fees annually to estimate expected lending fees over the following calendar year for the respective stock (see Dixon, Corbin, and Kelley (2021) for detailed descriptions of FIS data).

2.2 Abnormal Undercutting Activity and Informed Trading

This section describes the construction of our informed trading riskmeasure, $QIDRes$. The intuition behind our measure reflects market makers' efforts to avoid trading against informed investors. We

[7] Book value is defined as Compustat's shareholder equity value (seq) plus deferred taxes (txdb). We use the "linktable" from WRDS to match stocks across CRSP and Compustat, dropping stocks without links.

[8] We match CRSP with COMPUSTAT and 13F using the link tables and matching code provided by WRDS.

[9] Barardehi, Bernhardt, Ruchti, and Weidemier (2021) modify this measure by using open-to-close, instead of close-to-close, daily returns to construct Amihud measure's underlying daily liquidity proxy.

argue that market makers become less willing to undercut each others' quotes when they perceive incoming order flow to be informed. This notion is also consistent with marker makers' concerns about their limit orders becoming stale and picked off by faster traders, as first observed by Budish et al. (2015). Intuitively, an increased likelihood of informed trading raises the risk of a market maker's limit orders going stale and makes the market maker less willing to jump in front of the queue through undercutting.

It is important to observe that undercutting is more likely to occur in less liquid stocks, e.g., stocks with wider bid-ask spread, for two reasons. First, with a market maker's limit orders coinciding with the NBBO, a wider bid-ask spread provides larger profits per round-trip set of liquidity providing trades as market maker orders are filled by incoming marketable orders. Second, since trades need to improve the price by only 1¢ to undercut, a wider bid-ask spread implies a capacity for undercutting in terms of number ob available intra-spread price ticks. Moreover, undercutting by the best existing quotes by 1¢ is relatively cheaper for higher share prices (see Li and Ye (2023) for discussion on the relevance of the interaction share price and minimum tick size for liquidity provision). This leads us to use relative quoted bid-ask spread to control for the variation in undercutting capacities offered by market conditions. Figure 1 documents a strong positive association between our measure of undercutting, QID, and percent bid-ask spread that yields a R^2 of 54.41%.

To operationalize our intuition that informed trading riskdiscourages undercutting, we employ a backward-looking procedure to estimate abnormal undercutting activity at the stock-day level. We first estimate the following regression using daily observations of each stock in each quarter

$$QID^q_{jt} = a^q_j + b^q_j \ln(PESP)^q_{jt} + u^q_{jt}, \tag{1}$$

where QID^q_{jt} measures undercutting activity in stock j on day t of quarter q; $\ln(PESP)^q_{jt}$ is the natural log of the corresponding time-weighted percentage quoted spread; and u^q_{jt} is the error term. We then use estimated intercept and slope coefficients from the preceding quarter, i.e., $\widehat{a^{q-1}_j}$ and $\widehat{b^{q-1}_j}$, respectively, to construct daily estimates of unexpected undercutting activity in the current quarter. Finally, we scale unexpected undercutting by the unconditional average of QID from the previous quarter, i.e., $\widehat{a^{q-1}_j}$, to account for any systematic cross-sectional variation in the undercutting activity. Thus, abnormal undercutting activity for stock j on day t of quarter q is given by:

$$QIDRes_{jt}^{q} = -\frac{QID_{jt}^{q} - \left(\widehat{a_j^{q-1}} + \widehat{b_j^{q-1}} \ln(PESP)_{jt}^{q}\right)}{\widehat{a_j^{q-1}}}. \tag{2}$$

Since undercutting is expected to be abnormally low in presence of informed trading, higher $QIDRes$ reflects higher informed trading.

In Appendix A.2, we examine the qualitative robustness of our findings to two modified constructions of $QIDRes$. The first modification, denoted $QIDResSD$, reflects the undercutting restrictions sue to minimum tick sizes. Whenever the 1-¢ tick size binds, liquidity providing algorithms may not undercut on exchanges using non-marketable limit orders. As such, one can argue that for stocks where minimum tick more often binds the variation in QID, which we measure using the standard deviation of QID, is lower. Our qualitative findings extend if, instead of $\widehat{a_j^{q-1}}$, we use the corresponding standard deviation of QID from the previous quarter. The second alternative, denoted $QIDResV$, augments equations (1) and (2) with the volatility of 1-minute returns based on quote midpoints. This approach ensures that our measures do not conflate informed trading riskwith the effect of higher volatility, e.g., reflecting more frequently arrivals of purely public information, that can also deter liquidity provision and undercutting. This modification also leaves our qualitative findings unaffected.

3 Results

3.1 The Causal Impact of Undercutting Costs on *QID*

We begin our analysis by establishing the validity of the QID ratio as a measure of undercutting. To do so, we leverage the tick size pilot (TSP), during which a select number of stocks had their minimum tick sizes increased from 1¢ to 5¢—see, e.g., Werner et al. (2022), for a detailed description of the experiment. An increase in the tick size will decrease runs by making undercutting more expensive. For a TSP stock, the cost to undercut increased by five fold. Consequently, we expect the implementation of TSP to be associated with a decrease in QID and that the conclusion of the TSP will be associated with a reversal.

We study two TSP event windows: one around the imposition of TSP and the other around its conclusion. For our analysis of the imposition of the TSP, we examine the time window of 08/11/2016 through 12/15/2016. We follow Griffith and Roseman (2019) and exclude from this window the trading days spanning the staggered imposition of the TSP which comprise 10/03/2016–10/23/2016.[10] Our analysis of the imposition of the TSP has a pre-period where both the pilot and control stocks had a tick of 1¢, running from 8/11/2016 to 10/02/2016, and a treatment period where pilot stocks had a 5¢ tick and control stocks had a 1¢ tick, running from 10/24/2016 to 12/15/2016. Our analysis of the conclusion of the TSP runs from 08/07/2018 through 11/20/2018, during which the minimum tick size for stocks in TSP Test Groups was simultaneously reduced from 5¢ to 1¢ on 10/01/2018.[11]

We compare undercutting activity, QID, of control stocks, denoted C, to those of TSP Test Groups 1 and 2, denoted G1 and G2, respectively. Reflecting the similarities between G1 and G2 and to increase the statistical power of our tests, we combine G1 and G2 stocks together. The "`tick size pilot indicator`" flag in TAQ data identifies control and pilot stocks as well as the exact dates tick size changes were enforced for each pilot stock, facilitating accurate identifications of enforcement dates when tick changes were enforced or lifted with delays relative to the dates intended by the program. Stocks that changed test groups or that were removed from the TSP, for any reason, are excluded, as are stock-days with previous day's closing prices below $5.00.

Our estimation strategy is similar to Barardehi, Dixon, Liu, and Lohr (2023) who show that the same change in the tick size due to TSP had opposing impacts on certain outcomes depending on the extent to which minimum ticks were binding pre-shock. But more important for our analysis is that undercutting runs are affected by how tight the bid-ask spread is, and thus how many price levels competing liquidity providing algorithms can use to undercut. Hence, we assign each control stock to one of four bins based on their prevailing time-weighted quoted spread prior to the imposition and conclusion of the TSP. For the imposition window, stocks are classified into

[10] Some effects related to the tick size change may not occur instantaneously as market participants may need time to optimize systems and adapt behavior. Excluding the imposition period helps mitigate some of this noise that may muddle inference of the steady state effects of the tick size change.

[11] Following Rindi and Werner (2019), we remove trading days coinciding with Labor Day, Thanksgiving, and Black Friday from our sample. We also do not omit the period surrounding the conclusion of the TSP as we do with the imposition of the TSP because nearly all TSP stocks returned to a 1¢ tick simultaneously, with market participants returning to a familiar trading environment, i.e., one that had continued to operate on the majority of stocks. For these reasons, we generally view the conclusion of the TSP as a cleaner test than the TSP imposition.

four bins according to their quoted spreads in May and June of 2016:[12] : bin 1 (tick constrained) 5¢ or less quoted spread, bin 2 (near-tick constrained) greater than 5¢ but less than 10¢, bin 3 (intermediate spread) greater than 10¢ but less than 15¢, and bin 4 (wide spread) greater than 15¢. For the conclusion of the TSP, we assign stocks to bins reflecting average quoted spreads in May and June 2018: bin 1 (tick constrained) less than 5.5¢, bin 2 (near-tick constrained) greater than 5.5¢ but less than 10¢,[13] bin 3 (intermediate spread) greater than 10¢ but less than 15¢, and bin 4 (wide spread) greater than 15¢.

Our difference-in-difference strategy estimates the impact of an exogenous change in tick size, hence undercutting costs, on QID. We estimate

$$QID_t^j = \alpha_0 + \alpha_p Pilot^j + \alpha_e Event_t^j + \beta \left(Pilot^j \times Event_t^j \right) + u_t + \varepsilon_t^j, \tag{3}$$

by event window and bin, where QID_t^j is stock j's undercutting activity on day t; $Pilot_t$ is an indicator variable that equals 1 for treated stocks (G1 or G2) and equals 0 for control stocks; $Event_t^j$ of a treated stock equals 0 prior to a change in minimum tick size and equals 1 after the change, accounting for the enforcement date differences across stocks; $Event_t^j$ of a control stock in the imposition (conclusion) window equals zero before 10/03/2016 (10/01/2018) and equals 1 as of 10/24/2016 (10/01/2018); u_t is the date fixed effect; and $\varepsilon_{j,t}$ is the error term. Similar to Barardehi et al. (2023), we estimate the treatment effect β by fitting equation (3) using both quantile and OLS regressions, winsorizing QID_t^j at its 1^{st} and 99^{th} percentiles by tick constraint bin and treatment category. All of our estimates control for date fixed effects and double-clustered standard errors at the stock-date level.[14]

In Table 1 shows that our findings strongly align with the expected effect of a tick size change on undercutting. The first row of Panels A and B provide the difference-in-difference effect of the

[12] Specifically we use WRDS Intraday Indicators data for time-weighted average quoted spread for each stock during regular trading hours and compute a simple average across all trading days in May and June 2016.

[13] This slight modification of bin 1's threshold reflects the restrictions put in place by the TSP. The 5¢ tick size creates a floor on quoted spreads making it all but impossible for a TSP stock to have a time-weighted quoted spread less than 5¢, thus the threshold for tick constrained stocks is 5.5¢ for the conclusion of the TSP.

[14] Due to variation in the dates when the TSP was implemented across TSP stocks, simultaneous inclusion of variable $Event_{j,t}$ and date fixed effects do not lead to perfect co-linearity. The introduction of date fixed effects reflects the fact that for some stocks, the enforcement/lifting dates of TSP restrictions differ from the intended dates by the program. However, in unreported results, we verify robustness to, instead, the use of stock fixed effects or the use of both date and stock fixed effects. The robustness of results across these specifications is consistent with the findings of Rindi and Werner (2019), who also state that their results are virtually unchanged as they vary their fixed effects specifications.

TSP on QID for the various groups along with the median/mean value of QID for the control stocks in the sample. Consistent with tick constraints hindering undercutting, the median/mean value of QID increases as spreads get wider with the QID value for tick constrained stocks being very close to zero. Nonetheless, across all groups, and for the TSP Imposition and conclusion, the wider tick size is associated with a statistically negative shift in the QID ratio that reverses when tick sizes are returned to 1¢.

Our additional analyses attribute the TSP effects on QID to changes in the quoting behavior, consistent with the impact of a change in tick size on undercutting choices of liquidity providers. Rows two and three break down the effect of the TSP on the two aspects of the QID ratio. The second row shows the difference-in-difference effect of the TSP on the number of quote improvements divided by total number of quote updates ($Impr$). We find that increased tick size reduces the ratio of quote improvements to quote updates, consistent with reduced undercutting as it becomes more costly jump to the front of the queue. The third row shows that a wider tick size raises the ratio of trade driven quote deteriorations to all quote updates ($DeterTrade$). Existing literature establishes that the widening of tick size during TSP raised trade sizes but left trading volume unchanged (e.g., see Rindi and Werner (2019)), which suggests a reduction in the number of trades. As such, $DeterTrade$'s numerator likely declines as tick size widens, suggesting that the positive effect of a wider tick on $DeterTrade$ reflects reductions on the denominator, i.e., the number of quote updates, that more than offsets the decline in the numerator. These findings reinforce our interpretation that a larger tick size discourages undercutting as reflected in liquidity providers' less aggressive quoting behavior.

Our findings establish the impact of exogenous changes in the cost of undercutting on the level of QID, suggesting a strong positive link between QID and undercutting activity. We next relate abnormally low undercutting activity, i.e., high $QIDRes$, to increased informed trading.

3.2 *QIDRes* and Information Arrival

Our next analysis leverages the increased likelihood of informed trading around major instances of information arrival to highlight the correlation between abnormally low undercutting activity and informed trading. Specifically, we focus on earnings announcements (EA), unscheduled corporate events (PR), and news arrivals unassociated with identifiable corporate events (NA).

For each stock, we form twenty-day windows around each information event that occurs on day t, with pre-vent trading days $t-10$ through $t-1$ and post event trading days t through $t+10$. Whenever available, we use the exact time stamp of the information event to accurately identify the event day t; an event is matched with day t if the event took place after-hours on day $t-1$ or before the close on day t. For earnings announcements, where COMPUSTAT does not provide timestamps, we assume they all arrive after-hours. Moreover, to prevent contamination due to clustering of events, we focus on isolated events that do not follow a similar event in preceding 10 trading days, nor are followed by a similar event in the following 10 trading days.

To set up our analysis, we first explore the behavior of existing measures of informed trading intensity/probability around these events and confirm the findings in the literature. We analyze the behaviors of five different versions of Bogousslavsky et al. (2023)'s ITI measure,[15] as well as three versions of PIN, discussed by Duarte et al. (2020), the $OWRPIN$ measure of Odders-White and Ready (2008),[16] and MIA measures of Johnson and So (2018).[17] Figure 2 shows that all versions of ITI rise around these instances of information arrival, and that qualitatively similar results obtain using PIN and MIA, even though results vary across different versions of PIN and MIA and for different information events. Overall, these findings are consistent with increased informed trading riskaround instances of material information arrival.

Turing to $QIDRes$ in Figure 3 we document the same pattern. Across all information events we find that $QIDRes$ rises leading up to the event, peaking on the day of the event and reverting afterword. Consistent with adverse-section concerns underlying the abnormally low undercutting activity around information events, we find $QIDRes$ spikes are associated with significantly wider bid-ask spreads (in Panels A, C, and E). This short-term inverse relation between abnormal undercutting activity and spreads, i.e., the positive relation between $QIDRes$ and spreads, obtains despite the positive long-term relation shown in Figure 1—which reflects more ample undercutting opportunities when spreads are wide. Reduced undercutting in the face of widened bid-ask spreads can only reconcile with increased adverse-selection concerns of liquidity providers, suggesting that

[15]We thank authors of Bogousslavsky et al. (2023) for generously sharing with us 2010-2019 daily ITI measures.

[16]Estimates of PIN measures for all NMS stocks up to 2012 are available at Professor Edwin Wu's website.

[17]Estimates of MIA measures for qualifying stock-days up to December, 2018 are available at Professor Travis Johnson's website. Out of 5,940,019 stock-day $QIDRes$ observations in our 2010-2018 sub-sample, we can only match 446,066 stock-days featuring MIA measures. The number of missing observations reflect at least to constraints associated with MIA measures: (1) a common share must be optionable; and (2) to construct MIA for a given stock-day, Johnson and So (2018) require non-zero put and call option volume over the preciding 60 trading days.

$QIDRes$ captures informed trading. Further bolstering the idea that these events are associated with significant information we also find spikes in trading volume and abnormal absolute daily return around these events (Panels B, D, and F). Panels A and B present the results for earnings announcements. Panels C and D present the results for unscheduled corporate events, and Panels E and F present the results for other news arrivals. Across all events we observe that these days are associated with a spike in the bid ask spread, abnormal trading volume, and in absolute abnormal return. Importantly, the behavior of $QIDRes$ is distinct from that of volatility around information events. Figure A.1 shows that the qualitative behavior of $QIDRes$ around information events remains unaffected when we modify our measure to directly control for the the effect of volatility.

We next show that changes in $QIDRes$ predicts imminent upcoming *unscheduled* information arrival events, i.e., PRs and NAs defined earlier. To highlight the incremental predictive power of $QIDRes$, we control for other observables that, according to Figure 3, exhibit distinct behaviors prior to information arrival days. Specifically, we control for bid-ask spreads, trading volume, and absolute daily returns. Moreover, instead of focusing on isolated events, we control for information event clusters by observing that current information events can predict future information events.

Our analysis estimates the probabilities of unscheduled press releases (PR) and news arrivals (NA) using logistic regressions of these probabilities on past changes in undercutting behavior and a set of control variables, accounting for firm fixed effects. The dependent variable is defined as indicator function $I(z)_t^j$, with $z \in \{\mathrm{RP}, \mathrm{NA}\}$ that equals 1 when event z takes place on day t for stock j and equals 0 otherwise. The set of independent variables contain 5-day changes $\Delta x_{t-1}^j = x_{t-1}^j - x_{t-6}^j$, with $x \in \{QIDRes, qsp, tv, |r|\}$, in abnormal undercutting, quoted bid-ask spread, trading volume, and absolute returns. These variables, as shown in Figure 3 exhibit notable changes in the days leading up to an information event. To control for past relevant information events, additional independent variables are indicator functions $I(Inf)_s^j$ that equal 1 if an earning announcement (EA), an unscheduled press release (PR), *or* a news arrival (NA) event takes place on day s for stock j and equal 0 otherwise, with $s \in \{t-5, \ldots, t-1\}$.

We estimate the probability of event z to occur on day t for stock j using logistic regressions on a year-by-year basis.[18] We fit the models once only using $QIDRes$ and once using $QIDRes$ and all other controls. Tables 2 and 3 show that a 5-day change in $QIDRes$ positively predicts the

[18]Estimation by year reflects the computational burden when using the over 6 million observations from all years.

immediately upcoming unscheduled press release or news arrival. This is consistent with market makers learning from order flow about a an imminent information event (Chae (2005)). For press releases, this finding is robust to controlling for changes in trading and quoting outcomes, that correspond with the change in $QIDRes$, as well as clustering of information events. For news arrivals, the statistical significance is affected by controlling for this these outcomes, which is consistent with our earlier finding that $QIDRes$ spikes are smaller around NAs, relative to those observed around EAs and PRs. Overall, $QIDRes$ appears to possess an incremental predictive power for imminent information events relative other liquidity and information variables.

3.3 *QIDRes* and Information Content of Trades

In this section, we relate the spikes in $QIDRes$ around information arrivals, discussed in Section 3.2, to the extent of private information contained in the typical trade associated with these spikes. To do so, we first show that the magnitude and persistence of the increase in $QIDRes$ reflect the magnitude of the associated information event. Our tests are motivated by Kim and Verrecchia (1994)'s premise that more informative public news lead to greater post-event information asymmetries. For earnings announcements, we use SUE scores from I/B/E/S to capture the variation in the magnitude of events: in a given quarter, earnings announcement SUE scores in the top or bottom 20 percent—indicating that the announced earnings were significantly higher or lower than analyst consensus—are considered highly informative events. For press releases and news arrivals, we proxy for the information content using post-event realized price movements. For a day-t event, we simply divide each quarterly sample into those events associated with high versus low *absolute* compound post-event 10-day return.[19] Events in the top 40 percent are identified as highly informative events, and those in the bottom 60 percent are the less informative events.

Panels A through C of Figure 4 show that the magnitude of the increase in $QIDRes$ positively correlates with the magnitude of the information event. We first note that there is minimal pre-event variation in $QIDRes$ based on the magnitudes of information events, indicating that any post-event differences in abnormal undercutting may not be attributed to persistent stock characteristics such as volatility. Consistent with abnormally low undercutting activity, i.e., high $QIDRes$, capturing increased informed trading, we find in all cases that the event-day increase in

[19] Qualitative findings are robust to excluding event days from these return calculations

$QIDRes$ is larger for highly informative events than it is for less informative events. Moreover, undercutting activity appears to rebound more quickly toward pre-event levels following less informative events, suggesting that market-making algorithms return to "business as usual" as the risk of trading against informed investors drops. This pattern is remarkably stronger for news arrivals that are classified by Ravenpack as disassociated with any corporate events, suggesting that these events are highly unanticipated by market participants.

We further highlight the link between $QIDRes$ and informed trading riskby decomposing the transaction cost associated with each trade, as captured by effective spread, into permanent and temporary price impact components. This decomposition reflects the idea that the cost of consuming liquidity for incoming marketable order flow consists two components: (1) the compensation that liquidity providers demand for exposure to adverse-selection risk, captured by price impact and reflective of potential information advantages of liquidity consumers; and (2) the compensation that liquidity providers demand in return for facilitating "immediacy", captured by realized spreads that is generally attributed to operational costs incurred and revenues collected by market makers (see, e.g., Hendershott, Jones, and Menkveld (2011)). If the abnormally low undercutting documented in Figure 3 is due to informed trading, then any corresponding variation in effective spread should be primarily attributable to the price impact (adverse selection) component. Panels D, E, and F of Figure 4 show exactly this. Around the news events realized spreads are effectively unchanged and the entire observed increase in the effective spread is explained by an increase in the adverse selection component of the effective spread.

3.4 *QIDRes* and Direct Sources of Informed Trade

In this section, we address an alternative explanation for the association between abnormally low undercutting, i.e., high $QIDRes$, and the arrivals of information events. Specifically, we provide evidence that $QIDRes$ is unlikely to only capture increased 'sniping risk' around information events. Budish et al. (2015) show that in continuous-time limit order markets high-frequency traders engage in an arms race over the speed with which they can place/cancel orders. A key result in this literature is that differences in order processing speeds across traders lead limit orders of 'slower' traders to become stale for very short periods of time as the prices move against these resting orders upon arrivals of public information. These stale orders are then picked off, i.e. sniped, by 'faster' traders, leading to losses to slow traders. This phenomenon poses an adverse selection risk that is unrelated to information asymmetry about the fundamental value of the asset, but rather the speed with which different traders can respond to the arrivals of public information.[20] Relevant for our analysis is the possibility that information events that we study purely reflect increased 'sniping risk', as opposed to increased information asymmetry regarding fundamental value, leading to a reduction in the willingness of liquidity providers to undercut.

To address this concern, we use more direct measures of informed trading, as opposed to solely relying on variations around information events, to provide cross-sectional evidence that links increased informed trading riskto high $QIDRes$.[21] We first show that $QIDRes$ is higher when short sellers more activity take (accumulate) or leave (cover) short positions. The literature has provided robust evidence that short-seller trades are informed (see, e.g., Desai, Ramesh, Thiagarajan, and Balachandran (2002); Engelberg, Reed, and Ringgenberg (2012); Boehmer and Wu (2013), among others), so we expect to observe higher $QIDRes$ for stocks with high short selling activity.

We match each stock's bi-weekly percentage change in short interest to the corresponding averages of various informed trading riskmeasures, including $QIDRes$. We then sort each bi-weekly cross-section into ten portfolios (deciles) of signed percentage change in short interest, with the bottom decile containing stocks with largest coverings of short interest and the top portfolio containing stocks with largest 10% of short interest accumulations. We then calculate portfolio-level average informed trading riskmeasures in each bi-weekly period.[22] We finally plot the time-series means of these averages against change-in-short-interest portfolios.

Figure 5 shows that most measures of informed trading risk follow ∪-shaped patterns as we go from portfolio of stocks with largest coverings of short interest (declie 1) to stocks with largest accumulations of short interest (decile 10). This is consistent with private information underlying both buying and selling activity by short sellers and confirms Bogousslavsky et al. (2023)'s findings that relate ITIs o short interest. However, consistent with short sellers main focus on investigating negative information about asset values, most informed trading risk measures are highest when

[20] Menkveld and Zoican (2017) extent these insights by showing that exogenous increased in order processing speed offered by exchanges may exacerbate this issue and harm liquidity provision.

[21] Nonetheless, Appendix A.2 shows that a modified version of our measure $QIDResV$, which directly controls for the volatility of 1-minute quote midpoint returns exhibit patterns around information events that are qualitatively similar to those of $QIDRes$. This evidence suggests that pure sniping risk does drive the variation in $QIDRes$.

[22] To ensure that our findings doe not pick up any temporal variation in liquidity provision activities, for $QIDRes$, we first adjust each bi-weekly stock-specific average relative to the corresponding market-wide mean $QIDRes$.

short interest accumulations are largest. Panel A shows that all versions of ITI display these patterns; whereas Panel B and C show that even though PIN, $DYPIN$, $GPIN$, and MIA follow similar patters, $OWRPIN$ exhibits a $\cap$-shaped pattern. Panels D and E in Figure 5 document relationships between $QIDRes$ and short-seller activity conditioning on the past levels of short interest and firm size, respectively. Our findings suggest that (1) increased $QIDRes$ in times of high short-seller activity is more pronounced for stocks with higher levels of short interest, indicative of a higher likelihood that order flow contains orders from informed short sellers; and (2) the link between $QIDRes$ and the information content of short selling is not a small-stock phenomenon. Importantly, all these qualitative findings extend if we conservatively exclude biweekly periods that overlap with at least an EA, PR, *or* NA,[23] reinforcing the conclusion that informed trading risk identified by $QIDRes$ is likely distinct from increased sniping risk associated with public information arrival.

Second, we show that most measures indicate increased information asymmetry around a subset of informed mutual-fund trades. Barardehi et al. (2022) use ANcerno to identify industry-neutral self-financed trades of mutual funds, denoted INSFIT, and establish these trades are informed. We estimate the average incremental difference between informed trading risk measures around INSFIT days and non-INSFIT days, controlling for firm and date fixed effects.[24] We form 1-, 3-, and 5-day windows around stocks-days representing an INSFIT trade, examining INSFIT-bought and INSFIT-sold stocks separately. We then compare informed trading risk measures observed inside versus outside these windows.

Table 4 shows that stock-days featuring informed institutional trades are associated with statistically higher average informed trading risk measures. Specifically, with the exception of $ITI_{insider}$, $GPIN$, and $OWRPIN$, results based on all measures are consistent with increased informed trading risk on stock-days surrounding with INSFIT buy or INSFIT sell trades. Further highlighting the relevance of the information content of INSFIT trades, we find the largest differences on the "day of", i.e., 1-day INSFIT trade windows. Widening these windows to 3-day and 5-day horizons around the underlying INSFIT trades lead to smaller estimated differences that become statistically

[23]Such biweekly periods account for nearly half of the stock-days in our sample.

[24]We thank authors of Barardehi et al. (2022) for generously permitting us to use daily indicators that identify stocks bought and sold through INSFIT. Barardehi et al. (2022)'s sample spans January 1999 through September 2011, leaving us with the overlap period of January 2010 through September 2011 for our analysis.

insignificant for some existing measures.

In sum, we find a positive link between more direct, established sources of informed trading and various measures of informed trading risk used in our analysis. Our finding suggests that $QIDRes$ captures variation in the extent of information asymmetry, rather than solely that in sniping risk.

3.5 *QIDRes* and Compensation for Liquidity Provision

We next show that spikes in $QIDRes$ are hard to reconcile with inventory management concerns of liquidity providers driven by capital constraints. Comerton-Forde et al. (2010) show that liquidity providers with capital constraints become reluctant to accumulate additional inventory when their inventories are unbalanced; and So and Wang (2014) show that expected returns from liquidity provision significantly rise prior to earnings announcements reflecting increased inventory risk. Thus, a potential explanation for reductions in undercutting, i.e., $QIDRes$ spikes, may reflect inflated market maker inventories driven by increased liquidity demand that leads capital constraints to bind. Compensation for such liquidity provision is often reflected by short-term price pressure that is followed by price reversals (see, e.g., Campbell, Grossman, and Wang (1993); Hendershott and Menkveld (2014)). Thus, if inventory management concerns underlie the spikes in $QIDRes$, i.e., abnormally low undercutting, we should observe greater price reversals following high-$QIDRes$ days. We find the exact opposite.

Trading days with higher $QIDRes$ are followed by weaker price reversals. On each day t we sort stocks into quintiles of $QIDRes$. We then regress the cumulative returns from the close of day t through the close of day $t + n$, with $n \in \{1, \ldots, 10\}$, on day t returns, controlling for date and stock fixed effects. A negative slope coefficient indicates price reversal with the magnitude of this slope coefficient indicating the magnitude of this reversal. Table 5 shows that the high $QIDRes$ portfolio, containing stock-days with abnormally low undercutting activity, have coefficients closer to zero than the low $QIDRes$ portfolio. For all future return horizons, n, reversals grow nearly monotonically weaker, with the absolute values of slope coefficients shrinking by half, as we move from the low $QIDRes$ quintile to its high quintile. Hence, inventory management concerns of liquidity providers cannot drive the variation in $QIDRes$. In contrast, and consistent with our earlier findings, weaker price reversals that follow days with higher $QIDRes$ further reinforces that $QIDRes$ picks up informed trading. This finding is also consistent with Bogousslavsky et al. (2023)

who find that trading days with higher informed trading intensity (ITI) are followed by weaker price reversals.

3.6 Asset Pricing Implications of *QIDRes*

The literature does not offer a theoretical or empirical consensus regarding the asset pricing implications of informed trading. For example, Easley and O'Hara (2004) argue that informed trading should be priced since the risk driven by information asymmetry is non-diversifiable; and so when a stock has more private information, and hence informed trading, investors will demand a premium to hold the stock because of the associated adverse selection risk. Consistent with this prediction, Easley et al. (2002) present evidence that PIN is priced in the cross-section (also see, e.g., Kelly and Ljungqvist (2012) and Derrien and Kecskés (2013))). In contrast, Lambert et al. (2012) argue that in a perfectly competitive market, information asymmetry risk is diversifiable and hence should not be priced, with Armstrong et al. (2011) providing empirical evidence supportive of this prediction. We show that $QIDRes$ predicts stock returns up to six months forward but, consistent with Bogousslavsky et al. (2023), we attribute this return predictability to limits to arbitrage.

A prominent drawback for the asset pricing implications of informed trading intensity/probability measures is that they are often correlated with stock illiquidity. This feature might conflate priced illiquidity (first shown by Amihud and Mendelson (1980)) with any potential return predictability associated with information asymmetry. For example, Duarte and Young (2009) show that PIN's cross-sectional return predictability primarily reflects a liquidity effect rather than informed trading. Responding to this concern, our measure of informed trading risk is, *by construction*, orthogonal to stock illiquidity. Recall from equation (2) that $QIDRes$ reflects the abnormal undercutting activity (1) relative to its expected levels *conditional* on percentage quoted spread and (2) after accounting for persistent cross-sectional variations in undercutting associated with any persistent stock charcteristic, e.g., illiquidity. To further bolster this virtue before testing the ability of $QIDRes$ to explain expect stock returns, we show our measure of informed trading risk is nearly orthogonal to existing measures, more importantly, to various stock illiquidity. Other measures of informed trading intensity/probability do not exhibit these properties.

Table 6 shows minimal cross-sectional correlation between monthly averages of $QIDRes$ vis à vis other measures of informed trading, ranging between 0 and 0.14, or stock illiquidity measures, ranging between 0 and 0.04. This lack of correlation reflects the construction of our informed trading risk measure, further distinguishing $QIDRes$ from existing measures of informed trading. Panel B in Table 6 suggests that, in contrast to $QIDRes$, different versions of ITI appear to be positively related to PIN and $DYPIN$ in the cross-section, with correlation coefficients that range from 0.19 to 0.53. In further contrast, Panels A and B in Table 6 also suggest that both ITI and PIN measures are related to stock illiquidity while $QIDRes$ (as expected) is not. For example, Panel A shows that in the 2010-2019 sample, the average of the absolute correlation coefficients obtained between different versions of ITI and various stock illiquidity is about 0.15, with the highest pairwise absolute correlation of 0.30; while the analogues average and highest value for $QIDRes$ are only about 0.01 and 0.02, respectively. Similarly, Panel B shows that the average absolute correlation between different versions of PIN and stock illiquidity measures is around 0.15, with a high pairwise absolute correlation of 0.23; while the analogues for $QIDRes$ are only 0.02 and 0.04, respectively.

We begin analyzing the asset pricing implication of $QIDRes$ by using simple portfolio sorts that suggest our measure of informed trading risk from quarter $q-2$ predicts monthly returns in quarter q.[25] For this analysis we work with a sample spanning January 2010 through August 2016, reflecting the significant impacts of TSP on the level of undercutting for a large group of stocks (see Section 3.1). These empirical choices allow us to examine the entire cross-section of NMS stocks with no TSP-driven gaps in the time-series of each stock.[26]

Table 7 shows that stocks with higher levels of informed trading risk feature higher expected returns. In particular, we find that average four-factor risk-adjusted monthly return of the portfolio of stocks with the the highest past levels of informed trading, i.e., stocks falling in the top $QIDRes$ quintile in quarter $q-2$, is 38pbs higher than that for the portfolio containing stocks with the lowest levels of informed trading, i.e., stocks falling in the bottom $QIDRes$ quintile in quarter $q-2$. Bogousslavsky et al. (2023) document next-month return predictability using ITIs; hence, complement their results, our finding that $QIDRes$ predicts monthly returns two quarters forward indicates that informed trading risk can predict future returns over longer horizons.

[25]Skipping one quarter allows enough time for dissipation of short-term price movements due to potential liquidity effects such as short-term price reversals.

[26]Unreported analysis insures that qualitative findings are robust to, instead, excluding TSP stocks between September 2016 through December 2018 when TSP was in effect, and using the remaining data in the 2010-2019 time period.

These portfolio return patterns cannot be explained based on a link between informed trading risk and stock illiquidity, as we earlier discussed minimal correlation between $QIDRes$ and stock liquidity. In addition, Panel A in Table A.1 documents that $QISRes$ is minimally correlated with a host of key stock characteristics, some of which like firm size, are known to explain the cross-section of expected returns. In addition, Panel B in Table A.1 shows that $QIDRes$ exhibit no temporal persistence; in contrast, it appears to be mean-reverting. In sum, consistent the random arrival/discovery of fundamental information, $QIDRes$ is not a persistent stock characteristic. As such, our portfolio sort findings are hard to reconcile with cross-sectional return differences related to known stock characteristics. Nevertheless, we next employ regression analysis to reinforce our portfolio result while controlling for a set of stock characteristics.

Our regression analysis estimates

$$RetRf_{j,q,m} = \gamma^0 + \gamma^1 \left(QIDRes_{j,q-1}\right) + +\gamma^2 \left(QIDRes_{j,q-2}\right) + \Lambda^\top \text{Control}_{j,q,m-1} + u_{j,q,m}, \qquad (4)$$

where $RetRf_{j,q,m}$ is stock j's return in month m of quarter q in excess of the corresponding 1-month T-Bill rate; $QIDRes_{j,q-2}$ denotes abnormal undercutting activity in quarter $q-2$ for stock j; $\text{Control}_{j,q,m-1}$ denotes the vector of controls including betas from the three-factor Fama-French model, book-to-market ratio, market capitalization, dividend yield, idiosyncratic volatility, previous month's return, the return from the prior 11 months, previous quarter's share of institutionally held shares, previous quarter's institutional ownership concentration, and share turnover in month $m-2$.

Table 8 summarizes our findings when we fit fixed-effect panel regressions based on equation (4): we find a statistically significant positive association between $QIDRes$ and expected stock returns, suggesting that stocks with higher expected informed trading risk also have higher returns. This finding is robust to (1) including year-month fixed effects only versus including both year-month and firm fixed effects, which we choose as our main specification; (2) to including institutional ownership concentration and share turnover, reflecting the extent of competition for liquidity between potentially informed investors (Lambert et al. (2012)); and (3) augmenting the set of controls with individual or all the five stock illiquidity measures, reflecting the main message of Duarte and Young (2009) as a general concern that may apply to any measure of informed trading.

Table 9 formally contrasts the abilities of different informed trading intensity/probability measures in explaining the cross-section of expected returns. We estimate horse race regressions based on modified specifications of equation (4) that include $QIDRes$ *and* different sets of alternative existing measures as independent variables subject to their availability. We find that the association between $QIDRes$ and expected returns remains in these regressions, and that most of the alternative measures do not load with a statistically significant coefficients. Notably, $QIDRes$ is *the* only measure that significantly predicts future returns in all specifications. We also note that ITIs are not completely backward-looking measures of informed trading risk as Bogousslavsky et al. (2023) train their machine learning algorithms using sub-sample of stock-days that are scatted over the entire time-series, and hence, quarter $q-2$ ITIs may, by construction, contain information about future returns. In sharp contrast, average $QIDRes$ from quarter $q-2$ is not conditional on any future trading or pricing outcome.

Our asset pricing evidence so far shows that our measure of informed trading, $QIDRes$, predicts stock returns two quarters forward. However, since $QIDRes$ does not constitute a persistent stock characteristic, this return predictability is hard to reconcile with the predictions of Easley and O'Hara (2004). Similarly, that $QIDRes$ is orthogonal to stock liquidity precludes the possibility that it is capturing some aspect of liquidity costs reflected in the cross-section of expected returns. We next show that limits to arbitrage can explain this robust return predictability.

We follow Bogousslavsky et al. (2023)'s logic that with random arrival/discovery of positive and negative news, measures of informed trading risk should not predict future stock returns. However, long-only investors are more likely to trade on positive news than on negative news, and hence measures of informed trading risk are more likely to pick up trading motivated by positive information that in turn are followed by price increases. For example, Akepanidtaworn et al. (2020) show that buy trades of fund managers are informed while their sell trades primarily mean to raise finance the informed purchases. Moreover, trading on negative information is subject short selling restriction such as security borrowing costs and regulatory constraints [CITE CITE CITE]. We focus on this latter channel to shed light on what underlies return predictability of $QIDRes$.

Our analysis is motivated by the literature that identifies short selling constraints as major limit to arbitrage that deters investigation and trading on negative information. We show that return predictability of $QIDRes$ is concentrated among stocks with tighter short sale constraints. We do so using two approaches, First, we use institutional ownership to proxy for short sale constraints.

Since institutional investors providing most of the lendable shares to potential short sellers who must borrow these shares before selling a stock short, studies like Chen, Hong, and Stein (2002) and Nagel (2005) suggest that short sale restrictions are tighter among stocks with lower institutional ownership. More recently, Sikorskaya (2023) shows that higher index benchmarked capital raises the cost of short selling, suggesting that high institutional ownership may also tighten short sale constrains.[27] Based on these insights from the literature, we expect the return predictability of $QIDRes$ to be concentrated among stocks with lowest and highest levels of institutional ownership. Our second approach is more direct and splits the sample based on observed equilibrium lending fees observed in the securities lending markets

Following Nagel (2005), we first orthogonalize institutional ownership relative to firm size, with both variables measured at the end of quarter $q-3$. That is, we calculate institutional ownership residuals with respect to a second-order polynomial of log market-capitalization. We then split each monthly cross-section into terciles of residual institutional ownership before estimating equation (4) in each tercile. Panel A in Table 10 shows that, consistent tight short sale constraints, return predictability of $QIDRes$ is concentrated among stocks with lowest institutional ownership. Moreover, consistent with the dominance of tendency of long-only investors to trade on positive, rather than negative, information, $QIDRes$ exhibits some return predictability among stocks with highest institutional ownership. To examine the effects of short sale constrains more directly, we also investigate $QISRes$'s return predictability on the level of lending fees, with higher such fees reflecting tightes short sale constraints. From FIS data, we calculate average lending fee of each stock in quarter $q-3$, and then sort monthly cross-section in the current year into terciles of this average security lending fee. Panel B in Table 10 clearly shows that only among stocks with high lending fees does $QIDRes$ exhibit statistically significant return predictability.

4 Conclusion

Despite the key importance of informed trading for different areas of financial economics, easy to implement empirical measures of informed trading have proven difficult to derive. In this paper, we propose an easy to compute and intuitive measure of informed trading risk which we refer to as

[27]Of note, Sikorskaya (2023)'s results are strongest for "special" stocks which often feature low institutional ownership.

$QIDRes$. Our measure only requires trades and quotes data and thus can be computed for almost all publicly traded stocks at the daily, or even finer, frequencies in any modern limit order market.

Our approach exploits the intuition that liquidity providers compete less to fill order flow if they perceive the incoming marketable orders to be informed. Specifically, a liquidity provider's appetite to "undercut" rivals should significantly drop when they expect arrivals of informed marketable orders. We argue that abnormally low undercutting activity reveals the concerns of liquidity providers about incoming informed orders and hence is an indirect measure of informed trading.

We contrast the $QIDRes$ with existing measures of informed trading intensity/probability whose constructions are computationally demanding, require proprietary data, or limit their applicability to certain stocks. We find that $QIDRes$ performs as well as or better than these alternative measures: (1) $QIDRes$ spikes around periods known to be associated with informed trading such as earnings announcements, unscheduled press releases, and news arrivals; (2) increases in $QIDRes$ predict imminent unscheduled information arrival events; (3) the magnitudes of the $QIDRes$ spikes are positively associated with the magnitudes of imminent information events; (4) stock prices reverse less on days when $QIDRes$ indicates more informed trading; (5) episodes of increased short selling activity are associated with higher $QIDRes$; and (6) stock-days with known informed mutual-fund trades exhibit higher $QIDRes$.

We also show that $QIDRes$ predicts stocks returns up to six months forward. However, $QIDRes$ is orthogonal to persistent stock characteristics, especially liquidity, indicating that its return predictability is distinct from liquidity premia as posited by Duarte and Young (2009) about PIN. Moreover, consistent with the notion that informed trading should not be predictable, $QIDRes$ does not constitute a persistent stock characteristic either. Hence, we attribute its return predictability to the asymmetry in limits to arbitrage that restrict trading based on negative information. In fact, return predictability of $QIDRes$ is concentrated among stocks with tightest short sale constraints.

References

Ahren, K. R. (2020). Do proxies for informed trading measure informed trading? evidence from illegal insider trades. *Review of Asset Pricing Studies 10*, 397–440.

Akepanidtaworn, K., R. Di Mascio, A. Imas, and L. Schmidt (2020). Selling fast and buying slow: Heuristics and trading performance of institutional investors. *Journal of Finance*. Forthcoming.

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

Amihud, Y. and H. Mendelson (1980). Market-making with inventory. *Journal of Financial Economics 8*, 31–53.

Anand, A., P. Irvine, A. Puckett, and K. Venkataraman (2012). Performance of institutional trading desks: An analysis of persistence in trading costs. *The Review of Financial Studies 25*(2), 557–598.

Ang, A., R. Hodrick, Y. Zhing, and x. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance 61*, 259–299.

Armstrong, C., D. Taylor, J. Core, and R. Verrecchia (2011). When does information asymmetry affect the cost of capital? *Journal of Accounting and Economics 49*, 1–40.

Barardehi, Y. H., D. Bernhardt, Z. Da, and M. Warachka (2023). Uncovering the liquidity premium in stock returns using retail liquidity provision. Working Paper.

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.

Barardehi, Y. H., Z. Da, and M. Warachka (2022). The information in industry-neutral self-financed trades. *Journal of Financial and Quantitative Analysis*. Forthcoming.

Barardehi, Y. H., P. Dixon, Q. Liu, and A. Lohr (2023). When does the tick size help or harm market quality? evidence from the tick size pilot. Working Paper.

Boehmer, E. and J. Wu (2013). Short selling and the price discovery process. *Review of Financial Studies 26*, 287–322.

Bogousslavsky, V., V. Fos, and D. Muravyev (2023). Informed trading intensity. *Journal of Finance*. Forthcoming.

Budish, E., P. Cramton, and J. Shim (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics 130*(4), 1547–1621.

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.

Chae, J. (2005). Trading volume, information asymmetry, and timing information. *Journal of Finance 60*(1), 413–442.

Chen, J., H. Hong, and J. Stein (2002). Breadth of ownership and stock returns. *Journal of Financial Economics 66*, 171–205.

Chordia, T. and B. Miao (2020). Market efficiency in real time: Evidence from low latency activity around earnings announcements. *Journal of Accounting and Economics 70*, 1013–35.

Comerton-Forde, C., T. Hendershott, C. Jones, P. Moulton, and M. Seasholes (2010). Time variation in liquidity: The role of market maker inventories and revenues. *Journal of Finance 65*(1), 295–331.

Conrad, J. and S. Wahal (2020). The term structure of liquidity provision. *Journal of Financial Economics 136*(1), 239–259.

Derrien, F. and A. Kecskés (2013). The real effects of financial shocks: Evidence from exogenous changes in analyst coverage. *Journal of Finance 68*(4), 1407–1440.

Desai, H., K. Ramesh, S. Thiagarajan, and B. Balachandran (2002). An investigation of the informational role of short interest in the NASDAQ market. *Journal of Finance 57*, 2263–2287.

Dixon, P. N., A. F. Corbin, and E. K. Kelley (2021). To own or not to own: Stock loans around dividend payments. *Journal of Financial Economics 140*, 539–559.

Duarte, J., E. Hu, and L. Young (2020). A comparison of some structural models of private information arrival. *Journal of Financial Economics 135*(3), 795–815.

Duarte, J. and L. Young (2009). Why is pin priced? *Journal of Financial Economics 91*(2), 119–138.

Easley, D., S. Hvidkjaer, and M. O'Hara (2002). Is information risk a determinant of asset returns? *Journal of Finance 57*(5), 2185–2221.

Easley, D. and M. O'Hara (2004). Information and the cost of capital. *Journal of Finance 59*(4), 1553–1583.

Engelberg, J., A. Reed, and M. Ringgenberg (2012). How are shorts informed? Short sellers, news, and information processing. *Journal of Financial Economics 105*, 260–278.

Foley, S., A. Dyhrberg, and J. Svec (2022). When bigger is better: the impact of a tiny tick size on undercutting behavior. *Journal of Financial and Quantitative Analysis*.

Foley, S., T. Meling, and B. A. Ødegaard (2021). Tick size wars: The market quality effects of pricing grid competition. *Available at SSRN 2866943*.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics 14*(1), 71–100.

Griffith, T. G. and B. S. Roseman (2019). Making cents of tick sizes: The effect of the 2016 us sec tick size pilot on limit order book liquidity. *Journal of Banking & Finance 101*, 104–121.

Hasbrouck, J. and G. Saar (2013). Low-latency trading. *Journal of Financial Markets 16*, 646–679.

Heath, D., M. C. Ringgenberg, M. Samadi, and I. M. Werner (2020). Reusing natural experiments.

Hendershott, T., C. Jones, and A. Menkveld (2011). "does algorithmic trading improve liquidity? *Journal of Finance 66*(1), 1–66.

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

Holden, C. W. and S. E. Jacobsen (2014). Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance 69*, 1747–85.

Hughes, J., J. Liu, and J. Liu (2007). Diversification and the cost of capital. *The Accounting Review 82*, 705–729.

Johnson, T. L. and E. C. So (2018). A simple multimarket measure of information asymmetry. *Management Science 64*(3), 1055–1080.

Kelly, B. and A. Ljungqvist (2012). Testing asymmetric-information asset pricing models. *The Review of Financial Studies 25*(5), 1366–1413.

Kim, O. and R. E. Verrecchia (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics 17*, 41–67.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315–1335.

Lambert, R., C. Leuz, and R. Verrecchia (2012). Information asymmetry, information precision, and the cost of capital. *The Review of Finance 16*(1), 1–29.

Li, S. and M. Ye (2023). Discrete price, discrete quantity, and the optimal nominal price of a stock. Working Paper.

Menkveld, A. J. and M. A. Zoican (2017). Need for speed? exchange latency and liquidity. *Review of Financial Studies 30*, 188–1228.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of financial economics 78*(2), 277–309.

Odders-White, E. R. and M. J. Ready (2008). The probability and magnitude of information events. *Journal of Financial Economics 87*(1), 227–248.

Petacchi, R. (2015). Information asymmetry and capital structure: Evidence from regulation fd. *Journal of Accounting and Economics 59*, 143–162.

Rindi, B. and I. M. Werner (2019). Us tick size pilot. Working Paper.

Sikorskaya, S. (2023). Institutional investors, securities lending, and short-selling constraints. Working paper.

So, E. and W. S. Wang (2014). News-driven return reversals: Liquidity provision ahead of earnings announcements. *Journal Financial Economics 114*, 20–35.

Wang, J. (1993). A model of intertemporal asset pricing under asymmetric information. *The Review of Economic Studies 60*, 249–282.

Werner, I. M., B. Rindi, S. Buti, and Y. Wen (2022). Tick size, trading strategies, and market quality. *Management Science*. Forthcoming.

Figures and Tables

Figure 1. Undercutting and Quoted Spreads.
The figure presents the relationship between undercutting activity, as measured by QID, and percent quoted bid-ask spread. for each stock, both QID and the natural log of time-weighted percent quoted bid-ask spread, constructed at the stock-day frequency, are averaged across all days in the sample. The scatter plot presents the correlation between these two averages across stocks. The sample includes stock-days of NMS-listed common shares between Jan 01, 2010 through Dec 31, 2019 with previous months' closing prices of at least \$5.

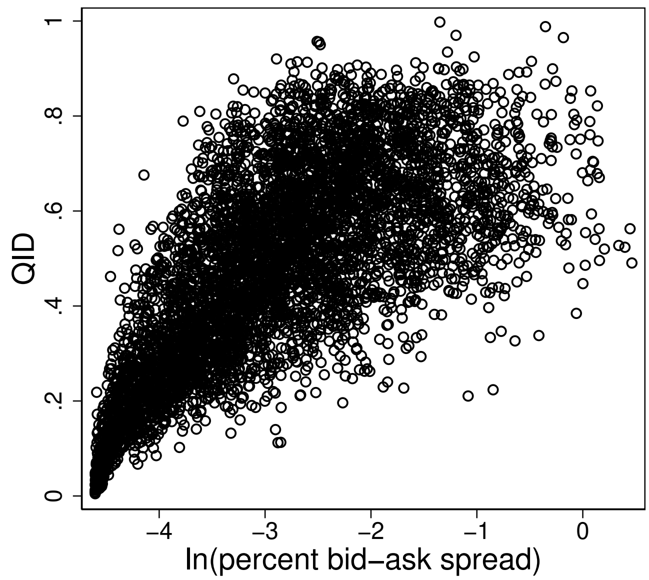


Figure 2. Existing Measures of Informed Trading around Unscheduled Corporate Announcements.
The figure presents medians of ITI, PIN, and MIA around around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Five versions of ITI and four PIN are considered. The sample includes all NMS-listed common stocks with previous quarter-end's share prices of at least \$5. Sample period is Jan, 2010 through Dec, 2019 for ITI; Jan, 2010 through Dec, 2012 for PIN; and Jan, 2010 through Dec, 2018 for MIA. Earnings announcement dates are obtained from COMPUSTAT; unscheduled press release dates and news arrivals not associated with any identified event are obtained from Ravenpack.

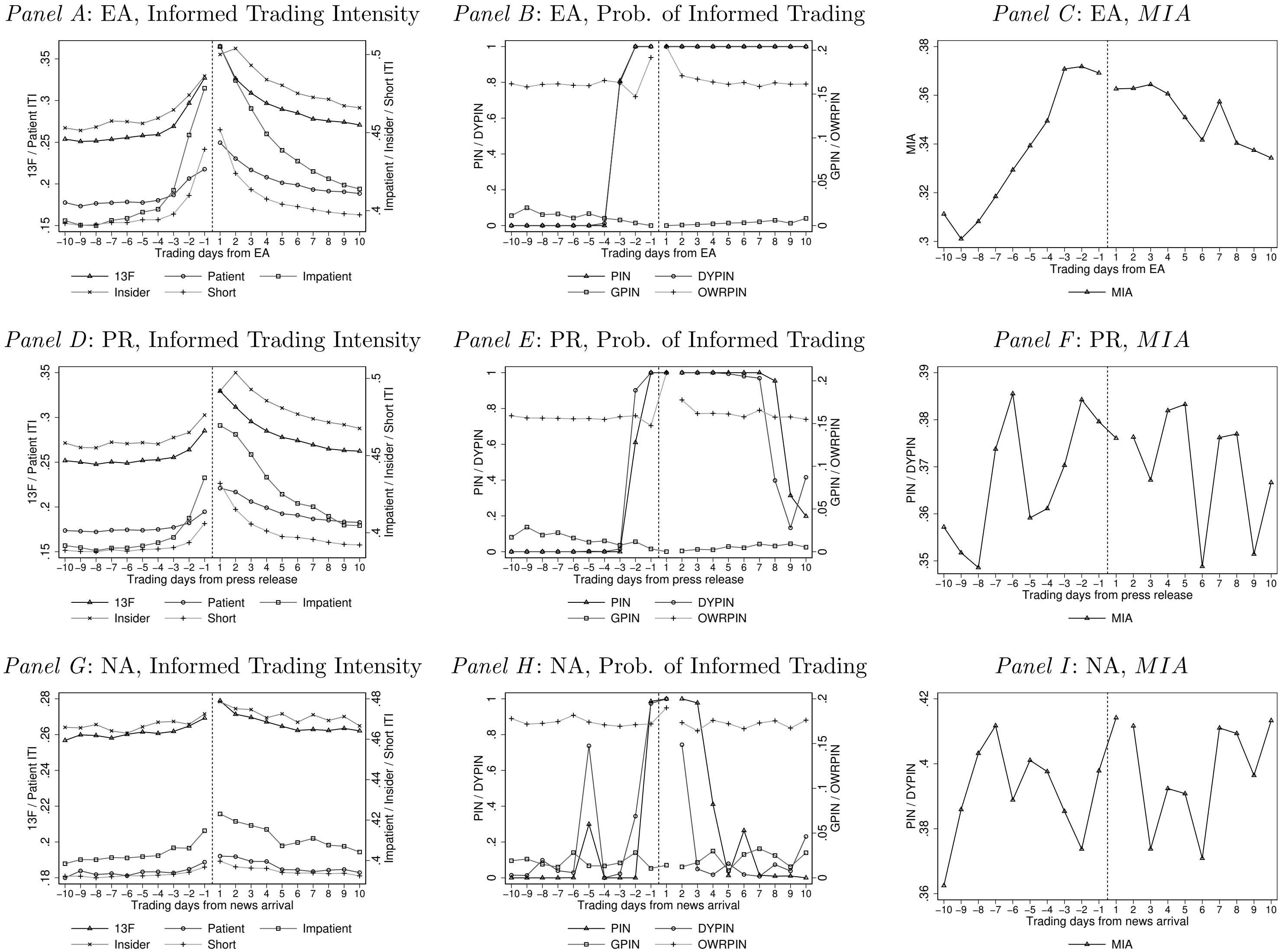


Figure 3. Undercutting Activity, Liquidity, and Information Asymmetry around Scheduled and Unscheduled Corporate Announcements.
The figure presents abnormal undercutting activity, dollar bid-ask spread, abnormal trading volume, and abnormal daily absolute return around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Daily abnormal undercutting values are calculated based on equation (2). Daily trading volume and absolute returns of each stock are normalized relative to the stock-specific median of each respective variable from the previous calendar quarter. The sample includes all NMS-listed common stocks between Jan, 2010 through Dec, 2019 with previous quarter-end's share prices of at least $5. Earnings announcement dates are obtained from COMPUSTAT; unscheduled press release dates and news arrivals not associated with any identified event are obtained from Ravenpack.

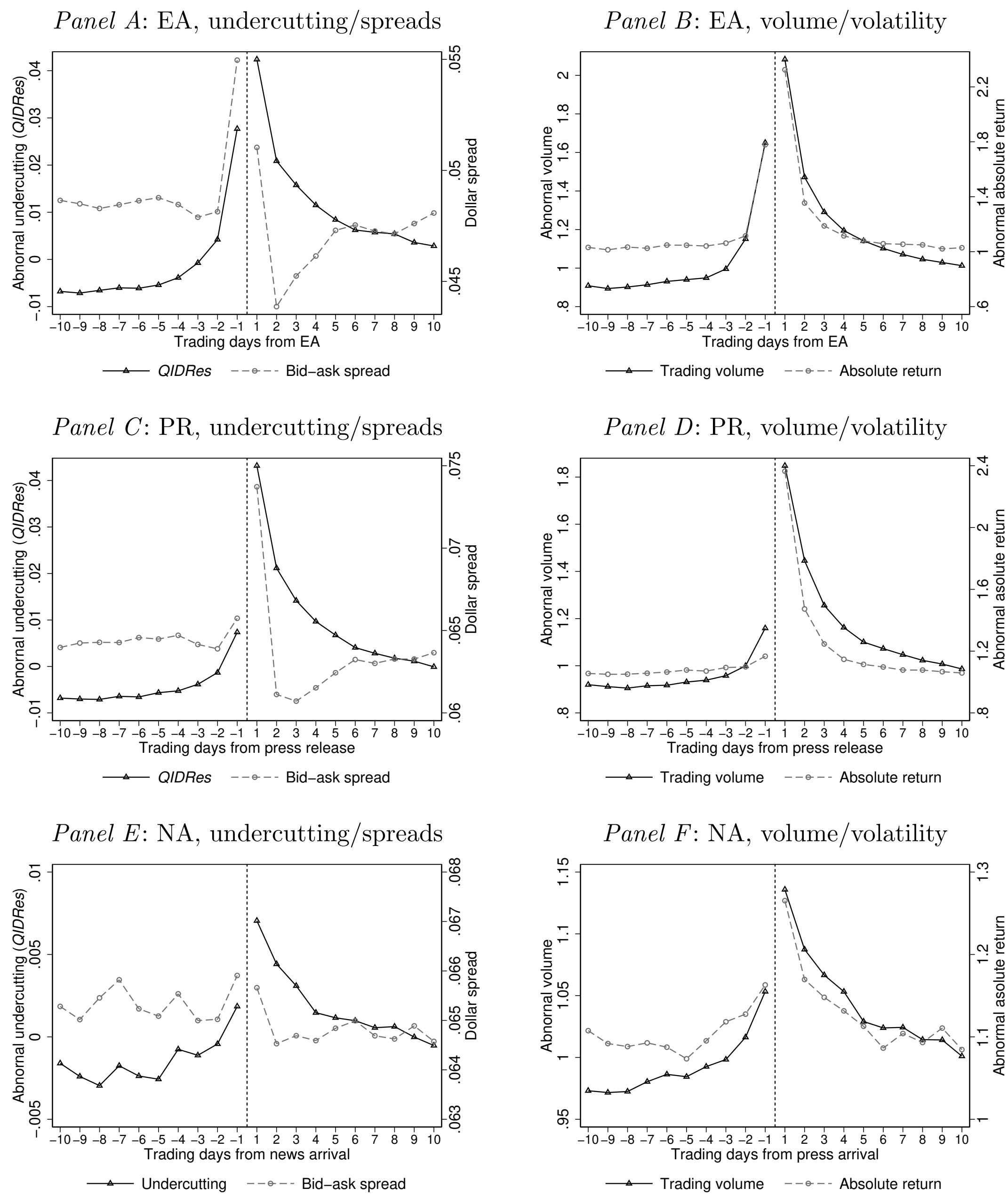


Figure 4. Undercutting Activity and Information Content of Trades, Events, and News.
Panels A through C present median abnormal undercutting activity around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Earnings announcements are classified into events with high earnings surprise score (SUE), i.e., top and bottom 20% of SUE scores in the respective quarter, and low/moderate SUE, i.e., the middle 60% of SUE scores in the respective quarter. Both unscheduled press releases (PR) and news arrivals (NA) are classified into high post-announcement/-news 10-day return, i.e., the top 40% of absolute 10-day compound return, and low post-announcement/-news 10-day return, i.e., the bottom 60% of absolute 10-day compound return. Daily abnormal undercutting values are calculated based on equation (2). Panels D through F present medians of daily percentage effective spreads, realized spreads and price impacts, all obtained from WRDS Intraday Indicators, around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). The sample includes all NMS-listed common stocks between Jan, 2010 through Dec, 2019 with previous quarter-end's share prices of at least $5. Earnings announcement dates are obtained from COMPUSTAT; SUE scores are obtained from I/B/E/S; unscheduled press release dates and news arrivals not associated with any identified event are obtained from Ravenpack.

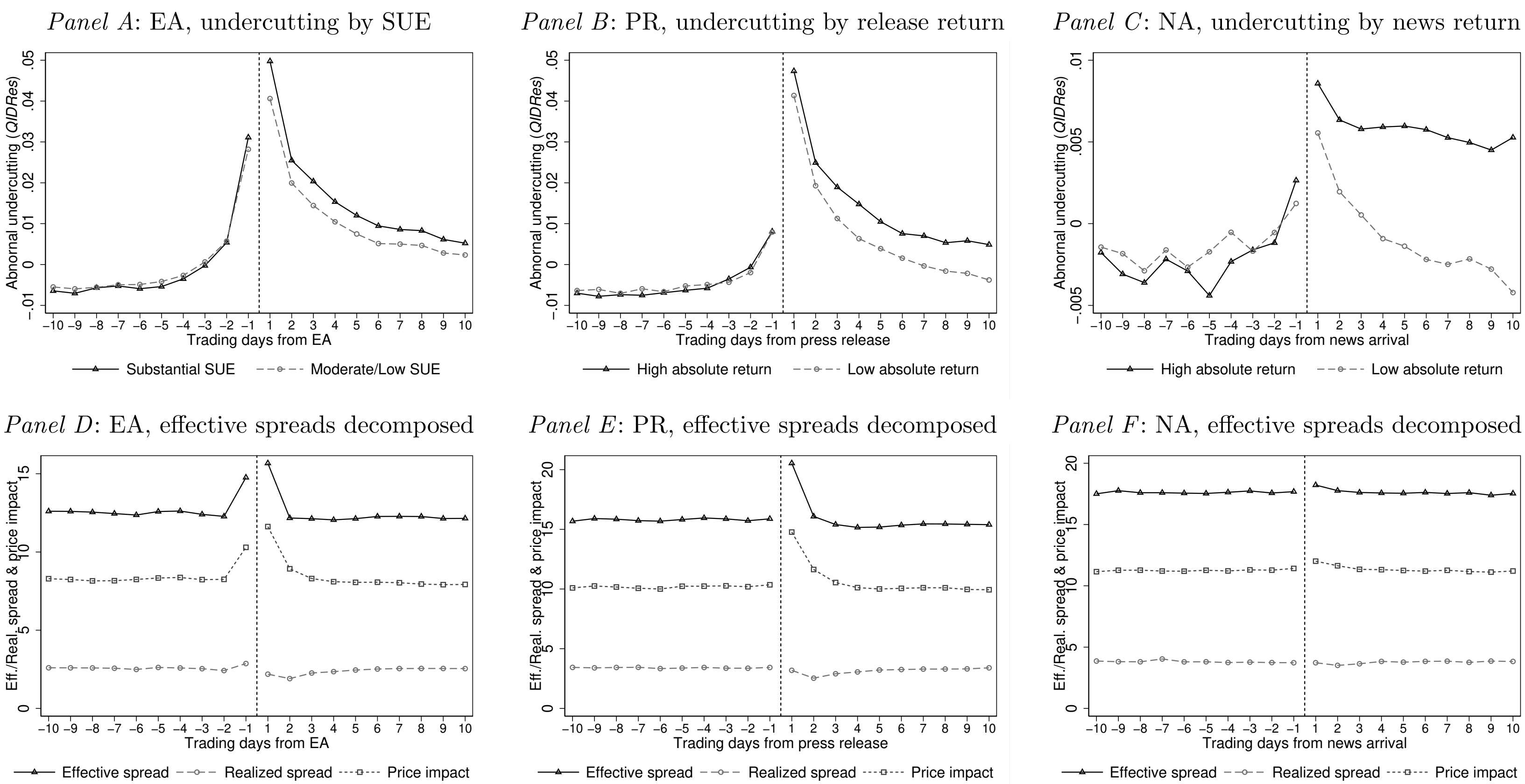
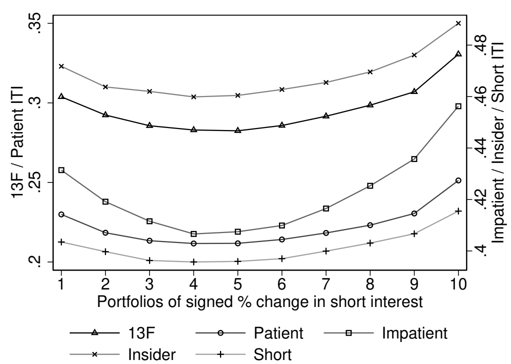
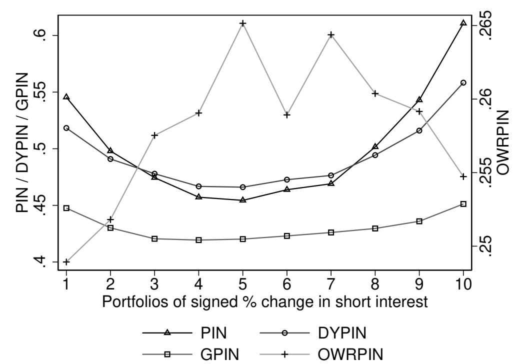


Figure 5. Informed Trading measures and Short Selling Activity.
The figure presents averages of various informed trading measures across levels of short selling activity. For averages of daily informed trading measures are calculated over bi-weekly intervals and matched with corresponding percentage change in short interest. each bi-weekly cross-section is sorted into portfolio (deciles) of signed percentage change in short interest. Equal weighted means of informed trading measures are calculated across stocks in each portfolio at the bi-weekly frequencies. The time-series averages are these means are plotted portfolio indexes, with 1 and 10 indexing the portfolios of stocks with largest declines and increased, respectively, in short interest. Panel A, B, and C present results for ITI, PIN, and MIA measures, respectively. Panel D presents results based on $QIDRes$ where each bi-weekly cross-section is decomposed in to terciles of the most recent short interest levels (defined as the most recent number of shares sold short by the total number of shares outstanding) before portfolios of percentage change in short interest are formed within each tercile. Panel E presents results based on $QIDRes$ where each bi-weekly cross-section is decomposed in to terciles of market-capitalization (defined as the product of the most recent share price and the total number of shares outstanding) before portfolios of percentage change in short interest are formed within each tercile. The sample includes all NMS-listed common stocks between Jan, 2010 through Dec, 2019 with previous quarter-end's share prices of at least $5.

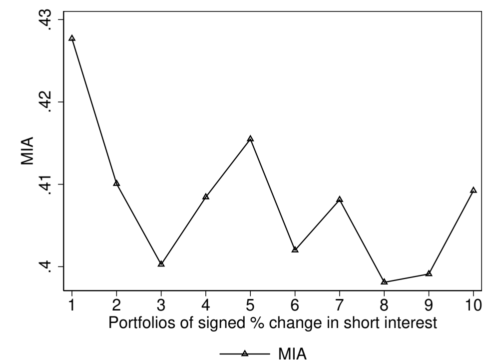
Panel A: ITI measures



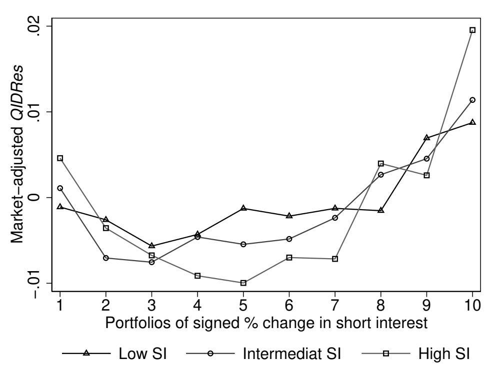
Panel B: PIN measures



Panel C: MIA measure



Panel D: $QIDRes$ by short interest level



Panel E: $QIDRes$ by market-capitalization

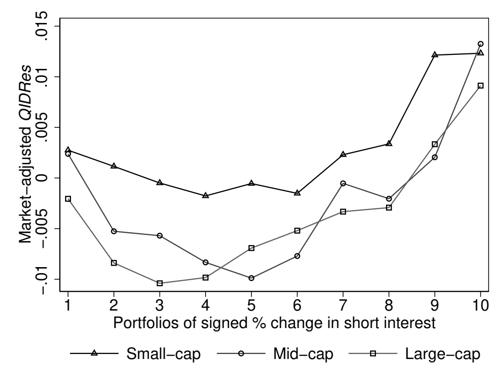


Table 1. Minimum Tick Size and the Undercutting Activity.
The table presents estimated impacts of an exogenous change in the minimum quoting and trading increment, i.e., tick size, on undercutting activity for differentially tick-constrained stocks. QID is the difference between the daily number of NBBO improvements and the number of trade-driven NBBO deteriorations, divided by the total number of NBBO updates. $Impr$ divides the number of NBBO improvements by the number of NBBO updates. $DeterTrade$ divides the number of *trade-driven* NBBO deteriorations by the number of NBBO updates Panel A presents the impacts of an increase in tick size from 1¢ to 5¢, using data from 08/12/2016-12/14/2016, for stocks with different tick constraint status prior to tick size increase. Stocks are classified into four tick constraint bins according to the average May and June 2016 quoted spreads of: (1) no more than 5¢, (2) 5¢ to 10¢, (3) 10¢ to 15¢, and (4) greater than 15¢. Panel B presents the impacts of a reduction in tick size from 5¢ to 1¢, using data from 08/08/2018-11/20/2018, for stocks with different tick constraint status prior to tick size reduction. Stocks are classified into four tick constraint bins according to the average May and June 2018 quoted spreads of: (1) no more than 5.5¢, (2) 5.5¢ to 10¢, (3) 10¢ to 15¢, and (4) greater than 15¢. Equation (3) is estimated using median (quantile) and OLS regressions. Estimates control for date fixed effects and double-cluster standard errors by stock and date. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: TSP imposition

	QR				OLS			
Dependent variable:	May & June 2016 quoted spread group				May & June 2016 quoted spread group			
QID	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
$Pilot \times Event$	−0.36***	−0.51***	−0.37***	−0.29***	−0.38***	−0.60***	−0.52***	−0.32***
	[−18.07]	[−17.01]	[−11.80]	[−11.80]	[−20.89]	[−35.33]	[−20.12]	[−13.18]
Median/Mean of control	0.11	0.54	0.70	0.74	0.16	0.50	0.64	0.65
$Impr$								
$Pilot \times Event$	−0.043***	−0.061***	−0.074***	−0.079***	−0.030***	−0.065***	−0.075***	−0.054***
	[−19.83]	[−18.52]	[−16.33]	[−12.43]	[−19.96]	[−30.41]	[−22.36]	[−11.00]
$DeterTade$								
$Pilot \times Event$	0.090***	0.098***	0.075***	0.052***	0.087***	0.12***	0.100***	0.056***
	[18.24]	[15.58]	[11.02]	[9.51]	[24.14]	[33.99]	[17.75]	[10.64]

Panel B: TSP conclusion

	QR				OLS			
Dependent variable:	May & June 2018 quoted spread bin				May & June 2018 quoted spread bin			
QID	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
$Pilot \times Event$	0.23***	0.64***	0.54***	0.27***	0.33***	0.52***	0.54***	0.27***
	[9.54]	[26.90]	[18.30]	[11.63]	[15.51]	[38.45]	[28.99]	[13.21]
Median/Mean of control	−0.01	0.35	0.38	0.46	0.02	0.33	0.37	0.42
$Impr$								
$Pilot \times Event$	0.010***	0.059***	0.078***	0.060***	0.0059***	0.032***	0.061***	0.048***
	[5.52]	[14.48]	[16.23]	[9.88]	[4.55]	[15.68]	[16.38]	[9.82]
$DeterTade$								
$Pilot \times Event$	−0.053***	−0.12***	−0.11***	−0.049***	−0.078***	−0.12***	−0.11***	−0.051***
	[−8.98]	[−21.92]	[−18.14]	[−10.57]	[−15.86]	[−39.42]	[−25.58]	[−12.39]

Table 2. Probability of Unscheduled Press Releases and Recent *QIDRes*.
This table reports in the predictive power of $QIDRes$ for the likelihood of imminent unscheduled press releases (PR). Panel A fit logit regressions of day t probability of PR conditional on the most recent 5-day change in $QIDRes$. Panel A fit logit regressions of day t probability of PR conditional on the most recent 5-day changes in $QIDRes$, bid-ask spread (qsp), trading volume (tv), and absolute daily return $|r|$ as well as arrivals of information events (Inf) including earnings announcements (EA), press releases (PR), or news arrivals (NA) over days $t-5$ through $t-1$. All estimates control for firm fixed effects. The sample includes NMS common shares from January 2010 to December 2019, excluding stocks whose previous month-end's closing price is below \$5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Logit estimates of the probability of PR conditional on $QIDRes$

Independent variable	Year 2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
$\boldsymbol{\Delta QIDRes_{t-1}}$	**0.27*****	**0.38*****	**0.23*****	**0.24*****	**0.20*****	**0.40*****	**0.26*****	**0.039****	**0.077*****	**0.19*****
	[3.84]	**[8.50]**	**[5.43]**	**[5.43]**	**[6.32]**	**[8.81]**	**[7.26]**	**[2.30]**	**[5.16]**	**[9.39]**
Observations	283,372	405,818	398,901	431,975	478,124	496,973	501,138	524,777	555,070	568,553

Panel B: Logit estimates of the probability of PR conditional on $QIDRes$ and controls

Independent variable	Year 2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
$\boldsymbol{\Delta QIDRes_{t-1}}$	**0.19*****	**0.29*****	**0.16*****	**0.16*****	**0.11*****	**0.28*****	**0.20*****	**0.026**	**0.055*****	**0.15*****
	[2.60]	**[6.45]**	**[3.56]**	**[3.56]**	**[3.38]**	**[5.99]**	**[5.36]**	**[1.43]**	**[3.62]**	**[7.06]**
Δqsp_{t-1}	−0.21*	−0.14	−0.21***	−0.010	−0.0032	−0.090	0.0032	0.19***	−0.021	0.059
	[−1.65]	[−1.41]	[−3.06]	[−0.16]	[−0.06]	[−1.38]	[0.04]	[3.22]	[−0.42]	[1.17]
Δtv_{t-1}	0.041***	0.034***	0.067***	0.046***	0.052***	0.049***	0.055***	0.052***	0.049***	0.067***
	[10.44]	[10.41]	[16.57]	[11.29]	[13.36]	[12.63]	[13.37]	[11.68]	[10.96]	[15.01]
$\Delta \|r\|_{t-1}$	0.019***	0.023***	0.016***	0.0051	0.0067**	−0.0022	−0.0067**	−0.0049	−0.0083***	−0.00026
	[5.64]	[8.03]	[3.98]	[1.35]	[2.01]	[−0.70]	[−2.02]	[−1.30]	[−2.64]	[−0.09]
$I(Inf)_t - 1$	0.73***	0.63***	0.84***	0.64***	0.71***	0.91***	1.02***	0.97***	0.97***	1.03***
	[34.79]	[37.76]	[44.08]	[37.81]	[46.44]	[57.62]	[59.65]	[58.88]	[59.63]	[66.37]
$I(Inf)_t - 2$	0.098***	0.038**	0.0047	0.0022	0.031*	0.073***	0.046**	0.0094	0.086***	0.12***
	[4.25]	[2.08]	[0.22]	[0.12]	[1.86]	[4.15]	[2.34]	[0.50]	[4.78]	[7.15]
$I(Inf)_t - 3$	0.045*	0.048***	−0.010	0.0041	0.037**	0.053***	0.086***	0.042**	0.099***	0.11***
	[1.92]	[2.64]	[−0.48]	[0.22]	[2.19]	[2.97]	[4.40]	[2.27]	[5.41]	[6.35]
$I(Inf)_t - 4$	0.041*	0.056***	0.0031	−0.043**	0.0060	0.044**	0.050**	0.054***	0.052***	0.16***
	[1.75]	[3.06]	[0.15]	[−2.33]	[0.35]	[2.43]	[2.52]	[2.88]	[2.85]	[9.16]
$I(Inf)_t - 5$	0.10***	0.17***	0.036*	0.053***	0.070***	0.058***	0.063***	0.051***	0.12***	0.19***
	[4.31]	[9.63]	[1.67]	[2.91]	[4.18]	[3.20]	[3.19]	[2.70]	[6.41]	[10.95]
Observations	273,678	393,227	384,672	416,006	462,636	481,954	483,768	494,925	530,147	553,131

Table 3. Probability of news arrivals and Recent *QIDRes*.
This table reports in the predictive power of $QIDRes$ for the likelihood of imminent news arrivals (NA). Panel A fit logit regressions of day t probability of NA conditional on the most recent 5-day change in $QIDRes$. Panel B fit logit regressions of day t probability of NA conditional on the most recent 5-day changes in $QIDRes$, bid-ask spread (qsp), trading volume (tv), and absolute daily return $|r|$ as well as arrivals of information events (Inf) including earnings announcements (EA), press releases (PR), or news arrivals (NA) over days $t-5$ through $t-1$. All estimates control for firm fixed effects. The sample includes NMS common shares from January 2010 to December 2019, excluding stocks whose previous month-end's closing price is below $5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Logit estimates of the probability of NA conditional on $QIDRes$

Independent variable	Year 2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
$\boldsymbol{\Delta QIDRes_{t-1}}$	**0.21*** [3.25]**	**0.100*** [2.66]**	**0.048 [1.40]**	**0.078** [2.23]**	**0.12*** [4.05]**	**0.098** [2.44]**	**0.079*** [2.73]**	**0.010 [0.78]**	**0.039*** [3.78]**	**0.054*** [3.67]**
Observations	262,975	390,686	387,258	432,219	464,996	480,602	480,382	515,661	557,768	576,516

Panel B: Logit estimates of the probability of NA conditional on $QIDRes$ and controls

Independent variable	Year 2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
$\boldsymbol{\Delta QIDRes_{t-1}}$	**0.19*** [2.81]**	**0.059 [1.52]**	**0.0060 [0.17]**	**0.023 [0.65]**	**0.035 [1.16]**	**0.013 [0.30]**	**0.038 [1.30]**	**−0.0020 [−0.15]**	**0.023** [2.16]**	**0.030* [1.95]**
Δqsp_{t-1}	−0.27** [−2.12]	−0.060 [−0.67]	−0.23*** [−4.13]	−0.14*** [−2.84]	−0.031 [−0.68]	−0.076 [−1.32]	−0.018 [−0.28]	−0.026 [−0.53]	−0.17*** [−4.76]	−0.049 [−1.51]
Δtv_{t-1}	0.021*** [6.50]	0.012*** [4.35]	0.026*** [7.74]	0.026*** [7.62]	0.040*** [11.61]	0.029*** [8.91]	0.035*** [10.82]	0.035*** [10.13]	0.027*** [8.17]	0.032*** [9.67]
$\Delta \lvert r \rvert_{t-1}$	0.012*** [3.63]	0.014*** [5.64]	0.012*** [3.55]	0.022*** [7.30]	0.017*** [5.53]	0.019*** [6.79]	0.0070** [2.52]	0.0059** [1.97]	−0.00078 [−0.36]	0.010*** [4.78]
$I(Inf)_t-1$	0.17*** [9.15]	0.21*** [15.08]	0.25*** [16.90]	0.21*** [15.89]	0.27*** [20.03]	0.27*** [19.71]	0.25*** [17.75]	0.30*** [23.30]	0.39*** [35.71]	0.38*** [36.18]
$I(Inf)_t-2$	0.16*** [8.63]	0.13*** [9.09]	0.19*** [13.00]	0.14*** [11.13]	0.24*** [17.80]	0.18*** [13.14]	0.13*** [9.27]	0.19*** [14.95]	0.23*** [20.84]	0.23*** [21.83]
$I(Inf)_t-3$	0.054*** [2.82]	0.074*** [5.13]	0.15*** [10.04]	0.15*** [11.51]	0.23*** [17.58]	0.12*** [8.63]	0.13*** [9.33]	0.12*** [9.15]	0.21*** [18.85]	0.16*** [14.59]
$I(Inf)_t-4$	0.052*** [2.74]	0.100*** [6.90]	0.10*** [6.75]	0.049*** [3.71]	0.11*** [8.16]	0.11*** [8.16]	0.083*** [5.79]	0.12*** [9.00]	0.20*** [17.49]	0.15*** [14.21]
$I(Inf)_t-5$	0.13*** [6.91]	0.16*** [11.27]	0.16*** [10.62]	0.12*** [9.10]	0.11*** [8.25]	0.16*** [11.59]	0.15*** [10.58]	0.14*** [10.94]	0.25*** [22.31]	0.18*** [16.96]
Observations	253,672	379,163	373,326	416,926	447,561	465,097	463,472	485,791	532,537	560427

Table 4. Informed Trading Measures around Informed Trades of Mutual Funds.
The table reports the incremental differences in various measures of informed trading around informed trades of mutual funds. Measures of informed trading are compared between stock-days around institutional buys and sells involved in Industry-Neutral Self-Financed Informed-Trades of Barardehi et al. (2022) and other stock-days. For each informed trading measure Y_t^j, the η_i coefficient from the following regression is reported: $Y_t^j = \eta_0 + \eta_i \times I(t-i, t+i)_t^j + \epsilon_t^j$, where $I(t-i, t+i)_t^j$ is an indicator function that equals 1 in the $i \in \{0, 1, 2\}$ days surrounding an INSFIT trade on t, and equals 0 otherwise. The model is fit once using INSFIT buy trade indicators and once using INSFIT sell trade indicators. All estimates control for firm and date fixed effects. The sample includes NMS common shares from January 2010 to September 2011, excluding stocks whose previous month-end's closing price is below $5. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Difference in informed trading measures around INSFIT buys trades

INSFIT	Informed trading measure										
trade window	***QIDRes***	ITI_{13D}	$ITI_{patient}$	$ITI_{impatient}$	$ITI_{insider}$	ITI_{short}	PIN	$DYPIN$	$GPIN$	$OWRPIN$	MIA
t	**0.0072*** **	0.014***	0.0058***	0.015***	0.0054***	0.0074***	0.022***	0.031***	−0.0043	−0.0054	−0.0035
	[5.91]	[6.50]	[3.10]	[9.01]	[3.08]	[9.24]	[3.43]	[4.47]	[−0.63]	[−1.62]	[−0.52]
$[t-1, t+1]$	**0.0060*** **	0.0072***	0.0028**	0.0084***	0.0030**	0.0045***	0.016***	0.019***	−0.0014	−0.0067**	−0.00097
	[5.60]	[4.73]	[2.03]	[6.72]	[2.44]	[8.03]	[3.03]	[3.82]	[−0.31]	[−1.98]	[−0.25]
$[t-2, t+2]$	**0.0057*** **	0.0056***	0.0015	0.0068***	0.0025**	0.0036***	0.012**	0.017***	−0.0015	−0.0067**	−0.0040
	[5.35]	[4.17]	[1.22]	[5.92]	[2.40]	[7.23]	[2.31]	[3.92]	[−0.38]	[−2.03]	[−1.17]
Sample mean	**0.0049**	0.2895	0.2241	0.4132	0.4490	0.4096	0.5131	0.4962	0.4277	0.2703	0.3885

Panel B: Difference in informed trading measures around INSFIT sell trades

INSFIT	Informed trading measure										
trade window	***QIDRes***	ITI_{13D}	$ITI_{patient}$	$ITI_{impatient}$	$ITI_{insider}$	ITI_{short}	PIN	$DYPIN$	$GPIN$	$OWRPIN$	MIA
t	**0.0072*** **	0.013***	0.0080***	0.013***	−0.0011	0.0054***	0.040***	0.033***	0.011	0.0061*	−0.0095
	[5.61]	[6.08]	[3.79]	[7.38]	[−0.50]	[5.91]	[5.95]	[3.99]	[1.50]	[1.82]	[−1.18]
$[t-1, t+1]$	**0.0057*** **	0.0097***	0.0049***	0.0091***	0.00056	0.0042***	0.022***	0.015***	0.0034	0.0035	−0.0045
	[5.17]	[6.31]	[3.63]	[7.27]	[0.45]	[6.59]	[4.19]	[2.79]	[0.71]	[1.32]	[−0.94]
$[t-2, t+2]$	**0.0052*** **	0.0075***	0.0045***	0.0077***	0.00061	0.0032***	0.017***	0.015***	0.0034	0.0019	−0.0047
	[4.99]	[5.51]	[3.69]	[6.75]	[0.57]	[5.86]	[3.52]	[3.20]	[0.86]	[0.74]	[−1.17]
Sample mean	**0.0049**	0.2895	0.2241	0.4132	0.4490	0.4096	0.5131	0.4962	0.4277	0.2703	0.3885

Table 5. Price Reversals by Abnormal Undercutting Activity and Time Horizon.
This table reports the extent of price reversal over the next 10 trading days as a function of abnormal undercutting activity and time horizon. Each daily cross-section is sorted into quintiles of $QIDRes$. For each such quintile panel regressions of compound returns over the next $n \in \{1, 2, \ldots, 10\}$ days, denoted $CR^j_{t+1,t+n}$, on current day's returns, denoted R^j_t, are estimated. Regressions control for stock and date fixed effects and double-cluster standard errors at both date and stock levels. All return cross-sections are winsorized at 1% and 99%. Estimates are reported by $QIDRes$ quintile and n. The sample includes all NMS-listed common stocks between Jan, 2010 through Dec, 2019 with previous quarter-end's share prices of at least \$5. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

		Dependent Variable									
$QIDRes$ quintile		$CR_{t,t+1}$	$CR_{t,t+2}$	$CR_{t,t+3}$	$CR_{t,t+4}$	$CR_{t,t+5}$	$CR_{t,t+6}$	$CR_{t,t+7}$	$CR_{t,t+8}$	$CR_{t,t+9}$	$CR_{t,t+10}$
Low	Slope	−0.057***	−0.065***	−0.077***	−0.087***	−0.092***	−0.096***	−0.098***	−0.10***	−0.11***	−0.11***
		[−3.96]	[−4.28]	[−4.84]	[−5.44]	[−5.81]	[−5.98]	[−6.14]	[−6.58]	[−7.49]	[−7.01]
	Observations	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520	1,071,520
2	Slope	−0.043***	−0.046***	−0.055***	−0.065***	−0.068***	−0.074***	−0.076***	−0.081***	−0.085***	−0.080***
		[−3.13]	[−3.32]	[−3.73]	[−4.34]	[−4.71]	[−5.01]	[−5.26]	[−5.69]	[−6.27]	[−5.73]
	Observations	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173	1,078,173
3	Slope	−0.036***	−0.040***	−0.050***	−0.056***	−0.056***	−0.059***	−0.059***	−0.063***	−0.069***	−0.066***
		[−2.99]	[−3.23]	[−3.80]	[−4.22]	[−4.25]	[−4.42]	[−4.44]	[−4.75]	[−5.30]	[−5.01]
	Observations	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529	1,078,529
4	Slope	−0.029***	−0.036***	−0.045***	−0.051***	−0.053***	−0.057***	−0.062***	−0.066***	−0.069***	−0.066***
		[−3.27]	[−3.81]	[−4.58]	[−5.00]	[−5.26]	[−5.49]	[−5.88]	[−6.31]	[−6.67]	[−6.23]
	Observations	1,078,338	1078338	1078338	1078338	1078338	1078338	1078338	1078338	1078338	1078338
High	Slope	−0.026***	−0.030***	−0.037***	−0.044***	−0.047***	−0.050***	−0.056***	−0.057***	−0.062***	−0.064***
		[−5.27]	[−5.46]	[−6.31]	[−7.10]	[−7.40]	[−7.64]	[−8.33]	[−8.39]	[−8.87]	[−8.92]
	Observations	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142	1,075,142

Table 6. Correlation between Informed Trading Measures and Stock Illiquidity.
This table presents the correlations matrices of informed trading measures and stock illiquidity. Panel A reports on the correlations between $QIDRes$ (indexed 1); five versions of ITI (indexed 2 through 6); and five illiquidity measures, time-weighted dollar quoted spread (QSP), size-weighted dollar effective srpead ($EFSP$), Kyle's λ ($Lambda$), Barardehi et al. (2021)'s open-to-close Amihud measure (AM), and Barardehi et al. (2023)'s retail-based institutional liquidity measure ($ILMV$), indexed 11 through 15, for the 2010-2019 sample sample. Panel B reports on the correlations between $QIDRes$, indexed 1; five versions of ITI, indexed 2 through 6; four versions of PIN, indexed 7 through 10; and five illiquidity measures, QSP, $EFSP$. $Lambda$, AM, and $ILMV$, indexed 7 through 11, for the 2010-2012 sample, where we have access to PIN measures. All measures are constructed at the monthly frequency by averaging daily observations.

Panel A: Correlation between, $QIDRes$, ITI, and illiquidity, the 2010-2019 sample

Variable index		Variable index 1	2	3	4	5	6	7	8	9	10
1	$QIDRes$										
2	ITI_{13D}	0.04									
3	$ITI_{patient}$	0.06	0.78								
4	$ITI_{impatient}$	0.04	0.73	0.55							
5	$ITI_{insider}$	0.00	0.35	0.39	0.38						
6	ITI_{short}	0.08	0.46	0.38	0.65	0.19					
7	QSP	0.01	−0.11	−0.07	−0.16	0.06	−0.24				
8	$EFSP$	0.01	−0.12	−0.08	−0.19	0.05	−0.26	0.97			
9	$Lambda$	0.01	−0.09	0.00	−0.25	0.11	−0.30	0.24	0.29		
10	AM	0.01	−0.08	−0.03	−0.19	0.00	−0.20	0.26	0.31	0.56	
11	ILM	0.02	−0.17	−0.04	−0.34	0.06	−0.37	0.37	0.42	0.59	0.45

Panel B: Correlation between, $QIDRes$, ITI, PIN and illiquidity, the 2010-2012 sample

Variable index		Variable index 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	$QIDRes$														
2	ITI_{13D}	0.11													
3	$ITI_{patient}$	0.14	0.80												
4	$ITI_{impatient}$	0.09	0.75	0.58											
5	$ITI_{insider}$	0.05	0.36	0.33	0.44										
6	ITI_{short}	0.12	0.52	0.49	0.69	0.31									
7	PIN	0.10	0.34	0.32	0.42	0.17	0.53								
8	$DYPIN$	0.11	0.32	0.30	0.38	0.19	0.43	0.63							
9	$GPIN$	0.00	−0.02	−0.01	0.00	−0.08	0.08	0.02	0.02						
10	$OWRPIN$	0.00	−0.01	0.00	−0.03	0.01	−0.06	−0.05	−0.02	−0.02					
11	QSP	0.02	−0.03	−0.05	−0.03	0.08	−0.21	−0.17	−0.10	−0.16	0.07				
12	$EFSP$	0.00	−0.03	−0.05	−0.04	0.07	−0.23	−0.18	−0.11	−0.17	0.09	0.93			
13	$Lambda$	0.04	0.00	0.02	−0.09	0.19	−0.23	−0.20	−0.10	−0.19	0.21	0.27	0.27		
14	AM	0.01	0.00	0.02	−0.08	0.05	−0.17	−0.11	−0.07	−0.10	0.13	0.13	0.15	0.66	
15	ILM	0.04	0.05	0.06	−0.06	0.17	−0.26	−0.21	−0.11	−0.22	0.10	0.43	0.42	0.63	0.40

Table 7. Informed Trading Alphas.

This table presents excess returns as well as three-, four-, and six-factor alphas conditional on our measure of informed trading. Each month m cross-section in quarter q is sorted into quintiles of $QIDRes$ from quarter $q-2$, with quintles formed based in NYSE breakpoints. The time series averages of monthly equally weighted portfolio returns as well that for the long-short (High−Low) portfolio, after subtracting the 1-month Treasury-bill rate, are reported as "excess returns." The 3-factor alphas reflect the intercept of time-series regressions of portfolio excess returns on Fama-French three factors. The 4-factor alphas reflect the intercepts when the 3-factor models are augmented with the momentum factor. The 6-factor alphas reflect the intercepts when 4-factor models are augmented by profitability and investment factors. The sample contains NMS common shares with previous month-end's closing prices of at least $5 from the January 2010 through August 2016. Standard errors are Newey-West-corrected using 12 lags. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

	$QIDRes$ quintile					
Portfolio return	Low	2	3	4	High	High−Low
Excess return	0.93**	1.24***	1.29***	1.32***	1.22**	0.29
	[2.23]	[3.26]	[3.03]	[3.36]	[2.57]	[1.66]
3-factor alpha	−0.17*	0.12	0.14***	0.14**	0.057	0.22
	[−1.79]	[1.55]	[2.67]	[2.33]	[0.54]	[1.30]
4-factor alpha	−0.18**	0.085	0.13**	0.18***	0.21**	0.38***
	[−2.01]	[1.34]	[2.31]	[3.47]	[2.38]	[3.13]
6-factor alpha	−0.13	0.10	0.13**	0.19***	0.26***	0.39***
	[−1.59]	[1.63]	[2.30]	[3.40]	[3.51]	[2.85]

Table 8. The Cross-Section of Expected Returns and Abnormal Undercutting Activity. This table reports on the relation between undercutting activity and the cross-section of expected returns. Equation (4) is estimated using $QIDRes$ constructed in the preceding two quarters and 5 liquidity measures constructed in month $m-2$. Other controls include three-factor Fama-French betas three-factor Fama-French betas ($\beta^{mkt}_{j,m-1}$, $\beta^{hml}_{j,m-1}$, $\beta^{smb}_{j,m-1}$), estimated using weekly observations from the two-year period ending in the final full week of month $m-1$, book-to-market ratio, ($\mathrm{BM}_{j,m-1}$), natural log of market capitalization, ($\ln(\mathrm{Mcap}_{j,m-12})$), dividend yield ($\mathrm{DYD}_{j,m-1}$), defined as total dividends over the past 12 months divided by the share price at the end of month $m-1$, idiosyncratic volatility ($\mathrm{IdVol}_{j,m-1}$), previous month's return ($RET_{j,m-1}$), preceding return from the prior 11 months ($RET_{j,(m-12,m-2)}$), and previous quarter's fraction institutionally owned shares outstanding ($IOShr_{j,q-1}$). The previous quarter's Herfindahl-Hirschman index for institutional ownership ($IOShrHHI_{j,q-1}$) and month $m-2$ share turnover ($TO_{j,m-2}$) serve as measures of market competition. Estimates are from panel regressions that control for firm and month-year fixed effects, double clustering standard errors by these two dimensions. The sample includes NMS common shares from January 2010 to December 2019, excluding stocks whose previous month-end's closing price is below $5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Independent Variable				Illiquidity measures						
				QSP	$EFSP$	Lambda	AM	ILM		
$QIDRes_{q-1}$	**0.29**	**0.50**	**0.55***	**0.52**	**0.52**	**0.51**	**0.50**	**0.50**	**0.54***	**0.58***
	[1.04]	**[1.60]**	**[1.70]**	**[1.65]**	**[1.65]**	**[1.62]**	**[1.60]**	**[1.60]**	**[1.68]**	**[1.78]**
$QIDRes_{q-2}$	**0.30***	**0.52*****	**0.53*****	**0.53*****	**0.53*****	**0.52*****	**0.52*****	**0.52*****	**0.54*****	**0.55*****
	[1.84]	**[3.04]**	**[3.12]**	**[3.10]**	**[3.10]**	**[3.04]**	**[3.03]**	**[3.05]**	**[3.12]**	**[3.20]**
Illiquidity	None	None	None	−1.10**	−2.33***	0.0014	−0.39	0.15	All	All
				[−2.30]	[−2.83]	[0.01]	[−1.09]	[0.30]		
β^{mkt}	−0.13	0.25	0.28	0.25	0.25	0.25	0.25	0.26	0.25	0.27
	[−0.45]	[1.26]	[1.41]	[1.26]	[1.24]	[1.26]	[1.24]	[1.28]	[1.27]	[1.39]
β^{hml}	−0.18	−0.15	−0.15	−0.15	−0.15	−0.15	−0.15	−0.15	−0.16	−0.15
	[−1.11]	[−0.98]	[−0.94]	[−0.99]	[−1.00]	[−0.99]	[−0.99]	[−0.98]	[−1.01]	[−0.97]
β^{smb}	0.065	0.071	0.082	0.070	0.070	0.069	0.067	0.072	0.067	0.078
	[0.44]	[0.48]	[0.56]	[0.48]	[0.48]	[0.47]	[0.46]	[0.49]	[0.46]	[0.53]
BM	0.27**	1.03***	1.10***	1.01***	1.01***	1.03***	1.05***	1.03***	1.01***	1.08***
	[2.15]	[3.07]	[3.33]	[3.02]	[3.02]	[3.04]	[3.11]	[3.07]	[3.01]	[3.29]
ln(Mcap)	−0.011	−2.45***	−2.43***	−2.41***	−2.41***	−2.45***	−2.46***	−2.44***	−2.40***	−2.40***
	[−0.25]	[−10.09]	[−10.02]	[−10.12]	[−10.07]	[−10.05]	[−10.08]	[−10.02]	[−9.98]	[−10.0]
DYD	0.48	−0.19	0.068	−0.45	−0.46	−0.20	−0.22	−0.17	−0.43	−0.23
	[0.25]	[−0.10]	[0.03]	[−0.22]	[−0.23]	[−0.10]	[−0.11]	[−0.09]	[−0.22]	[−0.11]
Id. Vol.	−0.21**	−0.065	−0.052	−0.060	−0.058	−0.066	−0.063	−0.064	−0.057	−0.045
	[−2.29]	[−0.90]	[−0.73]	[−0.83]	[−0.80]	[−0.90]	[−0.87]	[−0.89]	[−0.78]	[−0.63]
RET_{-1}	−1.10	−4.45***	−4.45***	−4.47***	−4.48***	−4.44***	−4.44***	−4.45***	−4.47***	−4.48***
	[−1.01]	[−4.08]	[−4.08]	[−4.10]	[−4.11]	[−4.08]	[−4.09]	[−4.08]	[−4.11]	[−4.11]
$RET_{(-12,-2)}$	0.31	−1.83***	−1.81***	−1.79***	−1.78***	−1.83***	−1.83***	−1.82***	−1.77***	−1.77***
	[1.11]	[−6.20]	[−6.04]	[−6.15]	[−6.12]	[−6.16]	[−6.19]	[−6.03]	[−5.89]	[−5.87]
$IOShr$	0.44***	−0.94***	−1.42***	−0.97***	−0.99***	−0.94***	−0.95***	−0.94***	−0.97***	−1.43***
	[2.81]	[−3.20]	[−4.28]	[−3.31]	[−3.36]	[−3.19]	[−3.25]	[−3.20]	[−3.32]	[−4.31]
$IOShr_HHI$			−1.78***							−1.71***
			[−3.49]							[−3.29]
TO			−28.1**							−30.2**
			[−2.24]							[−2.42]
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	234,110	234,026	234,026	234,026	234,026	233,564	234,026	234,026	233,564	233,564

Table 9. The Cross-Section of Expected Returns and Informed Trading: A Horse Race. This table reports on the relation informed trading measures and the cross-section of expected returns. equation (4) is estimated using $QIDRes$, along with different subsets of other informed trading measures, from the preceding two quarters. Control variables contain the full set of controls used in Table 8. The sample periods 2010-2019, 2010-2018, and 2010-2012 reflect the availability of alternative measures ITIs, MIA, and PIN, respectively. The samples include all NMS common shares, excluding stocks whose previous month-end's closing price is below \$5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. All estimates control for year-month and stock fixed effects, and standard errors are double-clustered at both levels. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

RHS variable	2010-2019 sample	2010-2018 sample		2010-2012 sample			
$\boldsymbol{QIDRes_{q-1}}$	**0.51**	**0.40**	**0.41**	**1.96**	**1.82**	**1.84**	**2.13**
	[1.49]	**[1.04]**	**[1.04]**	**[1.44]**	**[1.33]**	**[1.35]**	**[1.42]**
$\boldsymbol{QIDRes_{q-2}}$	**0.52*****	**0.54****	**0.49****	**2.64****	**2.44****	**2.52****	**2.45****
	[2.93]	**[2.48]**	**[2.29]**	**[2.36]**	**[2.23]**	**[2.33]**	**[1.99]**
$ITI_{13D,q-1}$	0.33		0.46		−1.04	−0.94	−3.91
	[0.28]		[0.30]		[−0.31]	[−0.28]	[−1.11]
$ITI_{13D,q-2}$	−2.05		−2.19		−4.62*	−4.54*	−2.82
	[−1.61]		[−1.42]		[−1.73]	[−1.71]	[−0.87]
$ITI_{patient,q-1}$	1.21		1.35		**5.54***	**5.70***	**7.00***
	[0.92]		[0.74]		**[1.81]**	**[1.83]**	**[1.92]**
$ITI_{patient,q-2}$	0.14		1.21		−0.42	−0.50	−0.34
	[0.10]		[0.70]		[−0.15]	[−0.18]	[−0.10]
$ITI_{impatient,q-1}$	−0.58		−1.86		−3.78	−3.20	−4.57
	[−0.39]		[−1.04]		[−0.93]	[−0.82]	[−0.87]
$ITI_{impatient,q-2}$	1.02		**3.18***		−0.41	−0.71	−4.23
	[0.74]		**[1.87]**		[−0.14]	[−0.24]	[−1.32]
$ITI_{insider,q-1}$	1.50		**2.52***		−0.49	−0.11	1.30
	[1.34]		**[1.72]**		[−0.13]	[−0.03]	[0.28]
$ITI_{insider,q-2}$	**2.48*****		**2.60****		**5.18***	**4.86***	3.19
	[2.74]		**[2.18]**		**[2.01]**	**[1.86]**	[0.89]
$ITI_{short,q-1}$	−0.44		0.082		3.92	5.49	9.83
	[−0.16]		[0.02]		[0.51]	[0.68]	[0.88]
$ITI_{short,q-2}$	0.99		−2.50		**11.0***	10.8	**13.6***
	[0.38]		[−0.73]		**[1.72]**	[1.65]	**[1.76]**
MIA_{q-1}		**1.68*****	**1.54*****				1.03
		[3.15]	**[2.95]**				[0.65]
MIA_{q-2}		0.22	0.21				0.55
		[0.45]	[0.45]				[0.46]
PIN_{q-1}				−0.013		−0.19	−0.11
				[−0.02]		[−0.28]	[−0.13]
PIN_{q-2}				0.075		0.0068	−0.11
				[0.16]		[0.01]	[−0.18]
$DYPIN_{q-1}$				−0.60		−0.70	−0.16
				[−0.98]		[−1.12]	[−0.22]
$DYPIN_{q-2}$				0.62		0.53	0.68
				[1.02]		[0.84]	[0.85]
$GPIN_{q-1}$				0.51		0.46	0.39
				[1.04]		[0.95]	[0.58]
$GPIN_{q-2}$				−1.01*		−1.01*	−1.13
				[−1.84]		[−1.75]	[−1.40]
$OWRPIN_{q-1}$				−0.72		−0.75	−0.47
				[−1.41]		[−1.29]	[−1.10]
$OWRPIN_{q-2}$				0.80		0.80	**0.65***
				[1.67]		[1.68]	**[1.71]**
Observations	216,077	119,098	118,113	25,045	25,045	25,045	16,065

Table 10. Return Predictability of Informed Trading Measures and Short Sale Constraints. This table reports on the relation between $QIDRes$ and the cross-section of expected returns by level of short sale constraints. Panel A reports estimation results of equation (4) within monthly terciles of residual institutional ownership, defined each month as the residual share of institutionally owned shares is orthogonolized relative firm size following Nagel (2005), with institutional ownership and form size measured at the end of quarter $q-3$. Panel B reports estimation results of equation (4) within terciles of quarter $q-3$'s average security lending fees obtained from FIS database. The sample includes NMS common shares from January 2010 to December 2019, excluding stocks whose previous month-end's closing price is below $5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. Estimates control for stock and year-month (year-quarter) fixed effects, and standard errors are double-clustered at both levels. The numbers in brackets are t-statistics with ***, **, and * identifying statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Return predictability of $QIDres$ by institutional ownership

Independent Variable	Tercile of residual institutional ownership: Low		Intermediate		High	
$QIDRes_{q-1}$	0.87**	0.85**	0.18	0.16	0.85**	0.79**
	[2.22]	[2.24]	[0.42]	[0.37]	[2.29]	[2.14]
$QIDRes_{q-2}$	0.84***	0.84***	0.047	0.036	0.53*	0.50*
	[3.10]	[3.10]	[0.17]	[0.13]	[1.80]	[1.72]
Stock characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Liquidity controls	Yes	No	Yes	No	Yes	No
Observations	77,127	77,279	77,777	77,985	78,247	78,349

Panel B: Return predictability of $QIDres$ by lending fee

Independent Variable	Tercile of security lending fee: Low		Intermediate		High	
$QIDRes_{q-1}$	0.57	0.54	0.43	0.40	1.38**	1.26*
	[1.50]	[1.45]	[1.06]	[0.98]	[2.02]	[1.90]
$QIDRes_{q-2}$	−0.14	−0.17	0.37	0.34	1.60**	1.53*
	[−0.27]	[−0.32]	[0.71]	[0.66]	[2.05]	[1.98]
Stock characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Liquidity controls	Yes	No	Yes	No	Yes	No
Observations	78,226	78,392	78,053	78,193	76,511	76,670

A Appendix

A.1 Is *QIDRes* a Stock Characteristic?

Table A.1. Correlations Between Current $QIDRes$, Past $QIDRes$, and Stock Characteristics. Panel A presents pairwise correlations between variables used in asset pricing tests. These variables include our measures of informed trading from the two preceding quarters, i.e., $QIDRes_{j,q-1}$ and $QIDRes_{j,q-2}$, three-factor Fama-French betas ($\beta^{mkt}_{j,m-1}$, $\beta^{hml}_{j,m-1}$, $\beta^{smb}_{j,m-1}$), estimated using weekly observations from the two-year period ending in the final full week of month $m-1$, book-to-market ratio, ($\text{BM}_{j,m-1}$), natural log of market capitalization, ($\ln(\text{Mcap}_{j,m-12})$), dividend yield ($\text{DYD}_{j,m-1}$), defined as total dividends over the past 12 months divided by the share price at the end of month $m-1$, idiosyncratic volatility ($\text{IdVol}_{j,m-1}$), previous month's return ($RET_{j,m-1}$), preceding return from the prior 11 months ($RET_{j,(m-12,m-2)}$), previous quarter's fraction institutionally owned shares outstanding ($IOShr_{j,q-1}$), previous quarter's Herfindahl-Hirschman index for institutional ownership ($IOShrHHI_{j,q-1}$), and month $m-2$ share turnover ($TO_{j,m-2}$). Panel B presents estimates of the AR(2) models the regress $QIDRes_{j,q}$ on $QIDRes_{j,q-1}$ and $QIDRes_{j,q-2}$ using different specifications with and without double-clustered standard errors at year-quarter and stock levels. The sample includes NMS common shares from January 2010 to December 2019, excluding stocks whose previous month-end's closing price is below \$5 as well as stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment.

Panel A: Correlations between current/past $QIDRes$ and stock characteristics

Variable index		Variable index 1	2	3	4	5	6	7	8	9	10	11	12	13
1	$QIDRes_{q-1}$													
2	$QIDRes_{q-2}$	−0.042												
3	β^{mkt}	0.005	−0.001											
4	β^{hml}	0.003	0.003	−0.03										
5	β^{smb}	0.021	0.007	0.12	0.15									
6	BM	0.034	0.027	−0.09	0.33	0.05								
7	ln(Mcap)	−0.026	−0.016	0.26	−0.10	−0.40	−0.27							
8	DYD	0.007	0.012	−0.13	0.10	−0.16	0.10	0.10						
9	Id. Vol.	0.030	0.014	0.14	−0.07	0.32	0.06	−0.31	−0.15					
10	RET_{-1}	−0.011	0.018	0.00	0.00	0.00	−0.09	−0.02	0.01	0.03				
11	$RET_{(-12,-2)}$	−0.064	−0.056	−0.01	−0.09	−0.04	−0.25	−0.07	−0.08	−0.07	−0.03			
12	$IOShr$	−0.005	−0.013	0.29	−0.03	0.01	−0.20	0.41	−0.12	−0.08	0.00	−0.03		
13	$IOShr_HHI$	0.010	0.011	−0.18	0.02	0.06	0.18	−0.35	0.01	0.14	−0.01	−0.01	−0.60	
14	TO	0.03	0.01	0.35	−0.09	0.09	−0.10	0.21	−0.11	0.24	−0.01	0.02	0.31	−0.16

Panel B: AR(2) models of $QIDRes$

	(1)	(2)	(3)	(4)
Constant	−0.0021***	−0.0021	−0.0021***	−0.0023***
	[−5.48]	[−0.62]	[−16.83]	[−22.84]
$QIDRes_{q-1}$	−0.035***	−0.035***	−0.020***	−0.046***
	[−13.42]	[−7.38]	[−4.11]	[−3.97]
$QIDRes_{q-2}$	−0.014***	−0.014	−0.015*	−0.039**
	[−5.17]	[−1.38]	[−1.82]	[−2.37]
Quarter FE	No	No	Yes	Yes
Stock FE	No	No	No	Yes
Clustered Errors	N/A	Quarter & Stock	Quarter & Stock	Quarter & Stock
Observations	75,018	75,018	75,018	74,793

This section present evidence that $QIDRes$ is not persistent stock/firm characteristic. Panel

A in Figure A.1 presents pairwise correlation coefficients between $QIDRes_{q-1}$, $QIDRes_{q-2}$ and an array of stock characteristics. $QIDRes$ is nearly orthogonal to all these stocks characteristics. Panel B present estimates of an AR(2) model that regresses $QIDRes_q$ on $QIDRes_{q-1}$ and $QIDRes_{q-2}$ using the panel of stock-quarter observations in our sample. $QIDRes$ exhibits no temporal persistence; if anything, it exhibit some degree of mean reversion, which consistent with its "residual" nature.

A.2 Modified Constructions of *QIDRes*

This section provides documents the robustness of our main findings to controlling for binding tick sizes and the effects of intraday volatility on undercutting. We construct two modified versions of $QIDRes$. The first modification uses equation (1) to fit parameters from the previous quarter, but it defines $QIDResSD$ as follows

$$QIDResSD_{jt}^{q} = -\frac{QID_{jt}^{q} - \left(\widehat{a_j^{q-1}} + \widehat{b_j^{q-1}}\ln(PESP)_{jt}^{q}\right)}{S(QID)_j^{q-1}}, \tag{5}$$

where $S(QID)_j^q$ denotes the standard deviation of daily QID_{jt}^q observations. This modification accounts for the more tightly bounded undercutting in stocks with binding minimum tick sizes, which in turn reduces the variation in QID in these stocks. The second modification accounts for the possibility that liquidity providing algorithms with very short holding periods avoid undercutting in more volatile stocks/markets, for a any given level of information asymmetry. Hence, the first stage in this modification involves modeling QID as a function of both spreads and volatility. That is, we first fit

$$QID_{jt}^{q} = \alpha_j^q + \beta_j^q \ln(PESP)_{jt}^{q} + \gamma_j^q qvol_{jt}^{q} + v_{jt}^{q}, \tag{6}$$

where $qvol_{jt}^q$ is the daily standard deviation of 1-minute quote-midpoint returns. Thus, a modified abnormal undercutting activity—that accounts for high-frequency volatility—for stock j on day t of quarter q is given by:

$$QIDResV_{jt}^{q} = -\frac{QID_{jt}^{q} - \left(\widehat{\alpha_j^{q-1}} + \widehat{\beta_j^{q-1}}\ln(PESP)_{jt}^{q} + \widehat{\gamma_j^{q-1}} qvol_{jt}^{q}\right)}{\widehat{a_j^{q-1}}}. \tag{7}$$

Figure A.1 shows that $QIDResSD$ and $QIDResV$ behave qualitatively very similarly to the basking $QIDRes$ around major information events.

Figure A.1. Abnormal Undercutting Activity around Scheduled and Unscheduled Corporate Announcements: Robustness.

The figure presents alternative versions of abnormal undercutting activity, $QIDResSD$ and $QIDResV$, around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). The sample includes all NMS-listed common stocks between Jan, 2010 through Dec, 2019 with previous quarter-end's share prices of at least \$5. Earnings announcement dates are obtained from COMPUSTAT; unscheduled press release dates and news arrivals not associated with any identified event are obtained from Ravenpack.