Detecting Informed Trading Risk from Undercutting Activity in Limit Order Markets *

Yashar H. Barardehi Peter Dixon Qiyu Liu

May 18, 2024

Abstract

We use abnormal undercutting activity (*QIDRes*) to measure informed trading risk, reflecting liquidity-providing algorithms competing less to fill marketable orders when adverse selection exposure rises. Despite its simple construction, when examined around information events, *QIDRes* behaves similarly to existing measures of informed trading intensity/probability whose constructions are complex. *QIDRes* predicts 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 increased informed institutional trades. *QIDRes* from prior quarters positively predicts monthly stock returns, especially among stocks with tighter short sale constraints. Since *QIDRes* is orthogonal to stock liquidity and is not 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

*We are grateful for comments and suggestions from Robert Battalio, Dan Bernhardt, Luis Ceballos, Zhi Da, Joel Hasbrouck, Tim Johnson, Travis Johnson, Albert Menkveld, Mitch Warachka, and Liyan Yang, as well as conference participants at LMU Corporate Finance Conference. Barardehi (barardehi@chapman.edu) is at the Argyros College of Business & Economics, Chapman University and the U.S. Securities and Exchange Commission. Dixon (DixonP@sec.gov), and Liu (LiuQi@sec.gov) are at the U.S. Securities and Exchange Commission.

1 Introduction

Informed trading is a key concept in various areas of financial economics including market efficiency, market structure, and the cost of capital. However, empirically measuring informed trading risk is difficult since informed traders conceal their presence by endogenously adjusting trading behavior with market conditions (Kyle (1985), Anand, Irvine, Puckett, and Venkataraman (2012)). This reality makes it difficult for researchers to empirically differentiate informed trading from other market conditions like liquidity (Ahren (2020), Duarte and Young (2009)). In this paper we propose an easy-to-compute, intuitive measure of nondirectional informed trading risk that is orthogonal to liquidity, performs at least as well as prominent existing measures in empirical tests, and only requires trades and quotes data. Importantly, our measure is computable at the daily, or even finer, frequencies for securities traded in any modern limit order market.

Our approach exploits the intuition that liquidity providers will compete less to trade against marketable orders they perceive to be informed. Specifically, we expect that the phenomena known as undercutting runs, or just runs (Foley, Dyhrberg, and Svec (2022), Foley, Meling, and Ødegaard (2021)), will decrease when informed trading risk is high and that this change in behavior will be observable in the trade and quote data. Undercutting refers to a trader using trivial price improvement to get their order to the front of the limit order queue. Undercutting runs occur when multiple trading algorithms repeatedly undercut each other as they compete to provide liquidity to an expected upcoming marketable order. In modern markets most liquidity providers have no affirmative obligation to provide liquidity in the face of informed or "toxic" order flow—liquidity provision that would lead them to incur losses ((Glosten and Milgrom (1985), Menkveld (2013)). Consequently, when informed trading risk is high, the willingness to "undercut" rivals will decrease, or disappear, reducing both the number and length of undercutting runs.[1]

The existing empirical literature on undercutting has primarily relied on proprietary account level data (e.g. Foley et al. (2021), Foley et al. (2022)) to identify runs. However, we observe that the nature of undercutting runs gives rise to patterns in the trades and quotes data that are

[1] Importantly, liquidity providing 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. Thus, 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. Put differently, abnormally low undercutting reveals the extent of liquidity providers' concerns about non-directional informed trading risk.

identifiable without proprietary data. Specifically, the hallmark of a run is a sequence of 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 the incoming marketable order executes the quote provided by the winner of the undercutting run. Empirically, this pattern can be analyzed by studying the difference between NBBO quote improvements and trade driven NBBO quote deteriorations. This leads us to measure undercutting activity by standardizing this difference at the stock-day level.

Specifically, our measure of undercutting, the QID ratio, reflects the total number of NBBO quote improvements observed on a given stock-day minus the corresponding number of trade-driven NBBO quote deteriorations on that same day, all divided by the sum of these two quantities. The construction of QID imposes boundaries of -1 and 1.[2] QID moving closer to 1 signifies increases in undercutting runs.

We establish the validity of QID as an undercutting measure by documenting its inverse relationship with undercutting costs. First, we exploit the exogenous change in the costs of undercutting driven by the SEC's Tick Size Pilot program (TSP). By temporarily raising the tick size from 1¢ to 5¢ for some stocks, the TSP quintupled a major component of undercutting costs (Werner, Rindi, Buti, and Wen (2022)). In standard difference-in-difference analysis, we find this exogenous increase in undercutting costs at TSP implementation ***reduced*** QID by about 0.44; whereas the TSP conclusion virtually mirrored the implementation results where QID ***increased*** by an average of 0.42.[3] Second, we exploit the positive impact of stock splits and the negative impact stock reverse splits on the costs of undercutting as reflected by relative tick sizes, i.e., 1¢ divided by share price. We find that QID significantly falls after stock splits, but significantly rises after stock reverse splits. Collectively, these findings highlight a strong inverse relationship between the costs of undercutting and QID, bolstering our interpretation of QID as a measure of undercutting activity.

We address two additional issues before employing QID to capture informed trading risk. First, prior literature demonstrates that informed trading risk measures tend to conflate informed trading and liquidity effects (Duarte and Young (2009), Ahren (2020)). Thus, QID could simply

[2]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.

[3]Unlike the TSP's heterogeneous effects on many other outcomes conditional on how binding the 5¢ tick was, its effect on QID are fairly homogeneous. Additionally, our analysis satisfies the heuristic hurdles when re-using experiments as all t-statistics range between 9–38, multiples of thresholds proposed by Heath, Ringgenberg, Samadi, and Werner (2020).

be capturing variations in liquidity that could also affect the willingness of a liquidity provider to undercut. In fact, Figure 2 documents a positive association between QID and stock illiquidity, measured by relative quoted bid-ask spread.[4] Second, the contributions of informed trading risk and liquidity to the variation in QID may vary in the cross-section. Thus, we must account for stock-specific effects to arrive at a measure that is comparable across stocks.

We address both concerns by orthogonalizing QID to liquidity and then standardizing it to make it comparable across stocks. Crucially, adopting this approach distinguishes our measure of informed trading risk from all existing measures, other relevant mircostructure outcomes, or stock characteristics. First, for each quarter and each stock, we fit a regression of daily QID on time-weighted relative bid-ask spread to control for liquidity conditions. Second, we apply the coefficients from the first step to the following quarter's realizations to produce estimates of the unexpected (residual) QID, i.e., undercutting activity that is orthogonal to liquidity. Next, we scale these stock-specific estimates of liquidity adjusted unexpected QID by the standard deviation of observed QID from the prior quarter. Lastly we multiply the resulting ratio by 1 to produce a positive, instead of inverse, measure of informed trading risk. We dub the resulting measure $QIDRes$. Consistent with its construction, $QIDRes$ satisfies two properties at the daily frequencies (1) it is distributed with a mean and a standard deviation close to 0 and 1, respectively; (2) it has nearly zero correlation with relevant contemporaneous mictostructure and liquidity outcomes such as quoted, effective, and realized spreads; price impact; volatility; and trading volume. This lack of correlation extends with respect to common measures of liquidity such as quoted spreads, effective spreads, and lambda as well as stock characteristics when we aggregate measures at quarterly frequencies.

We examine the behavior of $QIDRes$ around multiple information events known to be associated with informed trading. We also compare it's behavior to that of other prominent measures of informed trading such as the 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

[4] 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.

asymmetry (MIA) measure of Johnson and So (2018).

We document that around earnings announcements, unscheduled press releases, and news arrivals there is a significant spike in $QIDRes$ that takes up to 10 days to rebound. This is a pattern we also observe with the other measures of informed trading risk. Moreover, we find 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 more slowly. Finally, we test the predictive power of $QIDRes$ for informational events, exploring the notion that market makers may learn from order flow about upcoming information events (Chae (2005)). In fact, we find that increases in $QIDRes$ predict imminent arrivals of ***unscheduled*** information events.

We next provide evidence inconsistent with $QIDRes$ 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 due to sniping risk, rather than their information disadvantages about fundamental values.[5] 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 increases in $QIDRes$ around major information arrivals. To address this, instead of only relying on information events, we link $QIDRes$ to more direct sources of informed trading.

First, we explore the relation between $QIDRes$ and changes in short interest. A large literature documents that changes in short interest are strong predictors of future stock performance (see, e.g., Boehmer, Huszar, and Jordan (2010), Dixon and Kelley (2022)). Consistent with $QIDRes$ being linked to informed trading, we document that $QIDRes$ is significantly higher during periods with large absolute changes in short interest, even after excluding periods that overlap with information events. Second, we examine the behavior of $QIDRes$ around stock-days with informed mutual-fund trades, as identified by Barardehi, Da, and Warachka (2022). Again, we document significantly larger $QIDRes$ on informed mutual fund trading days than on days without such trading.[6]

We also provide evidence suggesting that $QIDRes$ is not simply reflecting inventory manage-

[5] With differences in order processing speeds across traders, the arrival of public news leaves some resting limit orders of slow trades stale not because 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)).

[6] We document similar behavior for both short interest and informed trading days with most other informed trading intensity/probability measures.

ment concerns of liquidity providers. Due to 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, i.e., higher $QIDRes$, as such liquidity providers demand greater compensation for providing liquidity. As demonstrated by Hendershott and Menkveld (2014), these inventory dynamics give rise to short-term price pressure followed by price reversals. Thus, if $QIDRes$ were reflecting inventory management concerns then stock-days with higher $QIDRes$ would be associated with stronger price reversals. However, we find the exact opposite occurs. Stock-days with higher $QIDRes$ are followed by ***weaker*** reversals which further suggests informed trading. Moreover, we show that these reversals patterns persist when we control for realized volatility—in fact, $QIDRes$ is nearly orthogonal to contemporaneous realized volatility

The inverse link between $QIDRes$ and subsequent reversals is also at odds with increased undercutting reflecting the increased use of limit orders by informed investors. Traders possessing a positive (negative) signal may undercut more on the bid (ask) side, instead of using marketable buy (sell) orders that would reveal their trading intentions. If this mechanism underlies the primitive motive for undercutting, then price reversals should be ***weaker*** following abnormally high undercutting, i.e., when $QIDRes$ is low, the exact opposite of our findings.

We next document asset pricing implications of $QIDRes$. We first demonstrate that long-short portfolios formed using $QIDRes$ from the prior two quarters earlier produce statistically significant risk-adjusted returns of over 30 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 positive associations between monthly expected stock returns and $QIDRes$ from the preceding two quarters, i.e., stocks with higher expected informed trading risk have higher returns. We highlight the incremental explanatory power of $QIDRes$ for returns, relative to existing informed trading intensity measures, in "horse race" regressions. These regressions, in addition to $QIDRes$ from the preceding two quarters, include subsets of ten corresponding alternative measures as independent variables. Not only does $QIDRes$ maintain 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 existing theories such as Easley and O'Hara (2004), interpreting informed trading risk as a stock characteristic, or Duarte and Young (2009), arguing that informed trading risk is correlated with illiquidity. Among unique features of $QIDRes$ are (1) it ***does not*** constitute a stock characteristic; and (2) it is orthogonal to stock illiquidity. Specifically, $QIDRes$ exhibits no temporal persistence but rather displays modest mean reversion, if anything. Moreoever, 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 limits to arbitrage. Short sellers who systematically investigate and trade on negative information face short sale constraints, while trading on positive information is not subject to such 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. When we control for short sale constraints using security lending fees indeed we observe that the return predictability of $QIDRes$ is stronger among stocks with tighter short sale constraints.

2 Linking Undercutting Runs to Informed Trading Risk

We next provide a simple framework that formally links the risk of trading against informed investors from liquidity providers' perspectives to their tendencies to participate undercutting runs. Consider the setup of a simple one period rational expectation equilibrium model based off of Glosten and Milgrom (1985). An asset takes the equally likely value of 0 or 1. The fraction π of liquidity demanders are informed and know the true value of the asset only buying when the the value equals 1 and only selling when the value equals 0. π captures informed trading risk in the market. The remaining 1 π fraction of liquidity demanders are uninformed and buy and sell with equal probability. Liquidity providers come in two types: sophisticated and unsophisticated. Unsophisticated liquidity providers, denoted $ULPs$, are passive, competitive, and set prices equal to the conditional expected value of the asset. Sophisticated liquidity providers, denoted $SLPs$, pay a cost c which will, with probability ρ inform them about whether the next trade to arrive is informed or uninformed and on which side of the market the trade will arrive. It does not inform them about the arrival time of the upcoming trade which is random.[7] There are m SLPs where the value m is determined in equilibrium such that the expected profit associated with being an SLP is equal to the cost c. The likelihood that at least one of the m $SLPs$ receives a signal is $\phi = 1 \quad (1 \quad \rho)^m$.

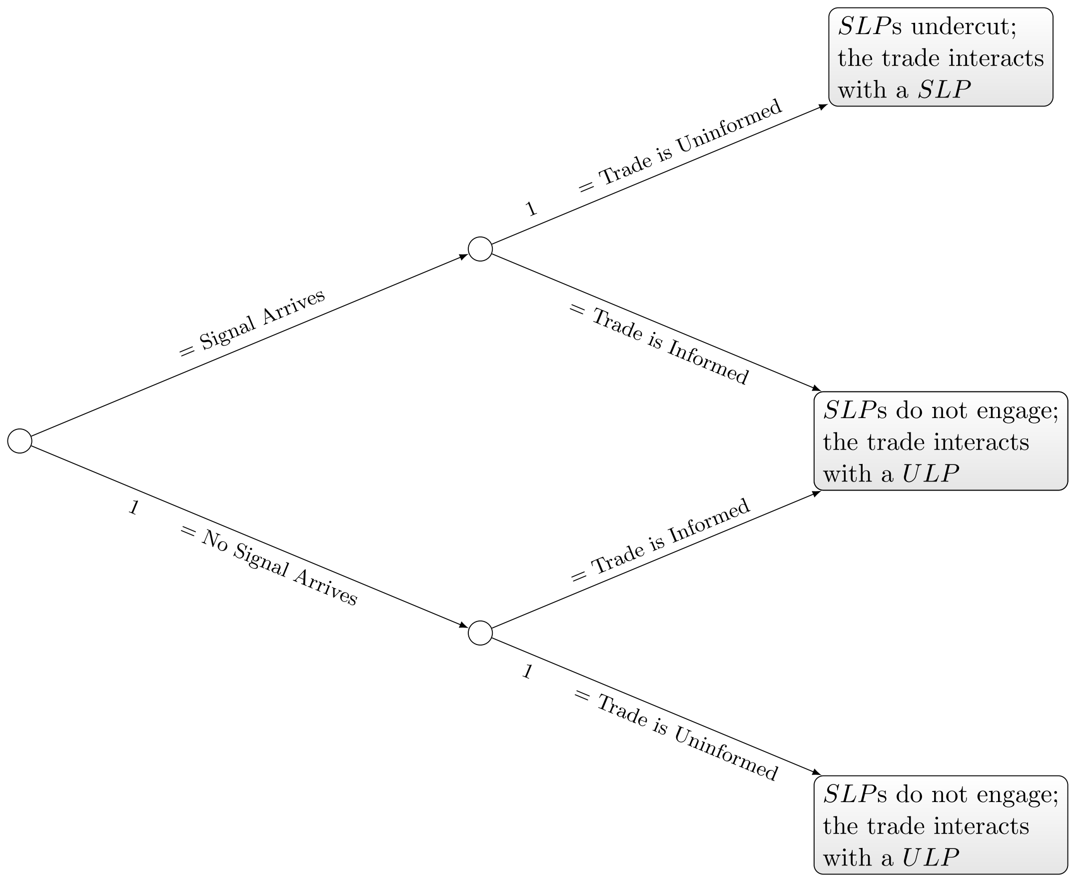
If a SLP receives a signal that an upcoming trade is informed, the SLP will simply sit out and not post any quotes allowing the ULPs to interact with the incoming informed trade. If no SLP receives a signal then all $SLPs$ sit out again allowing the ULPs to interact with the upcoming trade. However, if an SLP receives a signal that the upcoming trade is uninformed, they will undercut the existing quote on that side of the market. The other $SLPs$, whether they receive a signal or not, will observe this quote improvement and will infer that a signal has been received and will submit their own undercutting orders and an undercutting run will ensue.[8] This behavior is outlined in the outcome tree in Figure 1.

From the outcome tree it is straightforward to see that the probability of a run, which can be thought of as the prevalence of runs in that market, is the probability that at least one SLP receives a signal, ϕ, multiplied by the probability that the next trade to arrive is uninformed, $(1 \quad \pi)$ as shown in 1.

[7]The cost c can be thought of as the cost of investing in the capacity to process, analyze, and respond quickly to information based in order flow. As discussed in greater detail in A.1 the exact arrival time of the next trade to arrive is random and follows an exponential distribution with arrival rate parameter .

[8]The assumption that all $SLPs$ can infer the signals of others via monitoring quote updates could be relaxed such that only those $SLPs$ receiving a signal engage in the undercutting run without changing any of the key inference. In this case could be redefined to be the probability that at least two $SLPs$ receive a signal $= 1 \quad (1 \quad)^m \quad m \, (1 \quad)^{m \quad 1}$, and all inference remains exactly the same since in both cases is increasing in both m and . Additionally the profit to undercutting is random, since the arrival of the uninformed trade is random and so it is unclear exactly when during the run the uninformed trade will arrive. However, given that SLPs know the arrival rate of trades, they can compute the expected time during an undercutting run a trade will arrive and so can compute the expected profit of a run that is earned by the winning quote provider, which we denote Π. The likelihood that a given SLP wins the undercutting run is $\frac{1}{m}$, so expected profits to undercutting are $\frac{}{m}$. For this market to be in equilibrium it must be the case that $c = \frac{}{m}$ which implies that the number of SLPs is $m = \frac{}{c}$.

Figure 1. Informed Trading Signal Arrivals and SLPs' Undercutting Choices.



$$P(UndercuttingRun) = \phi(1 \quad \pi). \tag{1}$$

The derivative of equation 1 with respect to informed trading risk is $\frac{P(UndercuttingRun)}{} = \quad \phi$, which is always less than zero. Thus, when informed trading risk is higher - i.e. π is larger - undercutting diminishes confirming the inverse link between the prevalence of undercutting and informed trading. We present the equilibrium outcomes and comparative statics for the model in Appendix A.1, showing that liquidity gets worse as undercutting risk increases. Our analysis is consistent with empirical findings in the literature (e.g.,Foley et al. (2021) and Foley et al. (2022)). This occurs because undercutting exacerbates adverse selection by preventing uninformed trades from interacting with ULPs who set posted quotes.

3 Related Literature

In this section, we link our paper's contributions to the existing literature. We contribute by developing an informed trading risk measure that is computed using aggregate frequencies of best 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 order 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-driven techniques as in Bogousslavsky et al. (2023); and (4) do not require significant 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 using quote improvements minus 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 also provide new evidence relevant for the debate about asset pricing implications of informed trading as $QIDRes$ from two prior quarters predicts monthly returns. 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. Because, our measure of informed trading risk is unrelated to stock liquidity and does not constitute a persistent stock characteristic, we attribute its return predictability to limits to arbitrage such as short sale constraints.

4 Data and Methodology

4.1 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.

We construct national best bid and ask prices (NBBOs), from 09:45am to 3:45pm each day, 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. Our daily undercutting measure, 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 such NBBO updates for stock j on day t. We flag a quote deterioration as trade-driven if it occurs no later than 10 milliseconds after a trade.

$$QID_{jt} = \frac{\#Impr_{jt} \quad \#DeterTrade_{jt}}{\#Impr_{jt} + \#DeterTrade_{jt}} \tag{2}$$

Panel A in Table 1 contains summary statistics of the national best quoted ask and bid updates. The mean and median daily national best bid (NBB) improvements are 1,046.5 and 579, respectively; the analogous mean and median for national best ask (NBO) are 1,052.9 and 584, respectively. Consistent with the prevalence of undercutting activity, the mean and median of daily trade-driven NBB deteriorations are 279.05 and 110, respectively, with ask-side analogues of 277.7 and 109. More compelling evidence for the prevalence of undercutting obtain from the fractions single-tick quote updates. Liquidity providers are expected to undercut the existing best price by the minimum amount of possible price improvement, i.e., one tick. Hence, best quote improvements are most likely to occur at single-tick updates. By contrast, trade-driven quote deteriorations ending undercutting runs more likely reflect multiple-tick updates as marketable orders may consume the depth available beyond the top of the order book. Consistent with this, on a typical stock-day, near 90% of quote improvements reflect single-tick updates. This is significantly higher than the analogous 61% ratio for trade-driven quote deteriorations. Consequently, and reflecting the larger frequency of best quote improvements than deteriorations, the thirteenth row in Table 1 shows that over 99% of QID observations are positive. Specifically, only 3,222 of observations (only 0.05% of the sample) correspond to a negative QID quantity, even though QID can be as small as 1.

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.[9] 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 Raven-

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

pack "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.[10]

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 $RET^{m-12}_{j;m-2}$, 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 book value divided by market capitalization at the end of month $m-1$.[11] 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.[12]

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);[13] 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

[10] 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.

[11] 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.

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

[13] 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.

outstanding using observations from month $m \quad 2$.

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).

4.2 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 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 2 documents a strong positive association between our measure of undercutting, QID, and percent bid-ask spread that yields a R^2 of 58.44%.

To operationalize our intuition that informed trading risk discourages 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

$$QID^q_{jt} = a^q_j + b^q_j \ln(PQSP)^q_{jt} + u^q_{jt}, \tag{3}$$

where QID^q_{jt} measures undercutting activity in stock j on day t of quarter q; $\ln(PQSP)^q_{jt}$ 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., a^{q-1}_j and b^{q-1}_j, respectively, to construct daily estimates of unexpected undercutting activity in the current quarter. Finally, we scale unexpected undercutting by the standard deviation of daily QID^q_{jt} observations, $S(QID)^q_j$, to account for cross-sectional differences in the variability of undercutting activity. Such variability reflect factors like the more tightly bounded undercutting in stocks with binding minimum tick sizes, which in turn reduces the variation in QID in these stocks.[14] Thus, abnormal undercutting activity for stock j on day t of quarter q is given by:

$$QIDRes^q_{jt} = \frac{QID^q_{jt} - \left(a^{q-1}_j + b^{q-1}_j \ln(PQSP)^q_{jt}\right)}{S(QID)^{q-1}_j}. \tag{4}$$

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

Reflecting the construction of $QIDRes$, we expect it to possess the following two properties: (1) is it distributed with a mean and a standard deviation that are close to 0 and 1, respectively;[15] (2) it should not be correlated with other relevant microstructure and liquidity outcomes. We find strong support of this in the data. The last two rows in Table 1's Panel A report the summary statistics for $QIDRes_{jt}$, indicating that the measure is tightly distributed around zero, with the mean of 0.07 and the standard deviation of 1.53. Panel B in Table 1 contains the correlation coefficients between daily abnormal undercutting activity, $QIDRes^q_{jt}$ and contemporaneous micrsotructure outcomes defined earlier, including quoted, effective, and realized spreads; price impact; realized volatility;

[14] 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.

[15] Despite the standardization of unexpected undercutting by equation (4), we do not expect $QIDRes$ to exhibit a mean of exactly 0 and a standard deviation of exactly 1. This is because in each quarter both the conditional mean and the standard deviation used to standardize undercutting are estimated based preceding quarter's data and will differ from the current quarter's realized mean and standard deviation.

absolute daily return; and trading volume. None of these correlation coefficient exceed 0.06 in absolute value, suggesting that $QIDRes$ is orthogonal to these outcomes.

In Appendix A.3, we examine the qualitative robustness of our findings to two modified constructions of $QIDRes$. The first modification, denoted $QIDResInt$, controls for the variation in the unconditional average of undercutting. Our qualitative findings extend if, instead of $S(QID)_j^{q-1}$, we use a_j^{q-1} to normalize unexpected undercutting. The second alternative, denoted $QIDResV$, augments equations (3) and (A.10) with the volatility of 1-minute returns based on quote midpoints. This approach ensures that our measures do not conflate informed trading risk with 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.

5 Results

5.1 The 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.[16] 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

[16]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.

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.[17]

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 TSP 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 four bins according to their quoted spreads in May and June of 2016:[18] : 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¢,[19] bin 3 (intermediate spread) greater than 10¢ but less than 15¢, and bin 4 (wide

[17] 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.

[18] 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.

[19] 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

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

$$QID_t^j = \alpha_0 + \alpha_p Pilot^j + \alpha_e Event_t^j + \beta \left(Pilot^j \times Event_t^j \right) + u_t + \varepsilon_t^j, \quad (5)$$

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 (5) 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.[20]

Table 2 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 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.

less than 5¢, thus the threshold for tick constrained stocks is 5.5¢ for the conclusion of the TSP.

[20] 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.

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.

We reinforce the link between undercutting costs and QID by exploiting the relevance of ***relative*** tick sizes for undercutting costs. Following O'Hara, Saar, and Zhong (2019), we focus on stock splits and reverse splits as events that raise and reduce undercutting costs, respectively, by changing relative tick sizes. With a fixed minimum tick size, i.e., 1¢, the share price decline due to a stock split raises relative tick size, while the share price rise due to a reverse split reduces it. For example, to improve the best ask price of $10 a liquidity provider must quote a round-lot or larger ask order at $9.99, incurring a relative cost of 1bps. With a 2-for-1 split, the best ask should shift to $5, leading to a $4.99 reflecting the next better ask price offered by an undercutting algorithm; and this corresponds to a 2bps relative cost—twice as large as the pre-split cost. As such, we expect undercutting activity, and hence QID, to fall following stock splits and to rise following stock reverse splits. This is exactly what we find.

Our analysis of QID around stock splits also addresses the generalizability of findings using the TSP experiment that focuses on small-cap firms. Specifically, when forming event windows that span 30 days around a stock (reverse) split, we exclude any stock featuring a closing share price of $5 or less over the even window. We identify 476 split and 27 reverse-split events that fit this criteria. The average and median market-capitalization of stocks with split events are $8.36 billion and $2.25 billions, respectively. The average and median market-capitalization of stocks with reverse-split events are $12.66 billion and $2.51 billion, respectively.

Panel A in Figure 3 shows that average QID drops from over 0.6 to below 0.5 following stock splits; in contrast, it rises from below 0.2 to over 0.3 following stock revers splits. Importantly, this cannot be attributed to the corresponding variation in relative quoted spreads, e.g., it cannot reflect the positive association documented in Figure 2. In fact, Panel B in Figure 3 shows no significant variation in relative quoted spreads around stock split events. This leads us to attribute the observed changes in QID around these events to changes in undercutting costs.

Our collective findings establish the impact of 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.

5.2 $QIDRes$ 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-trading-day windows around each information event occuring on day t, with pre-event 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,[21] as well as three versions of PIN, discussed by Duarte et al. (2020), the $OWRPIN$ measure of Odders-White and

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

Ready (2008),[22] and MIA measures of Johnson and So (2018).[23] Figure 4 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.

Turning to $QIDRes$ in Figure 5 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 afterward. 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 2—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 $QIDRes$ 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, as our next analysis indicates, the behavior of $QIDRes$ appears to be distinct from that of volatility around information events. Figure A.2 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

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

[23] 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.

$QIDRes$, we control for other observables that, according to Figure 5, 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 $z \in \{\mathrm{PR}, \mathrm{NA}\}$ 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 5 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.[24] We fit the models once only using $QIDRes$ and once using $QIDRes$ and all other controls. Tables 3 and 4 show that a 5-day change in $QIDRes$ positively predicts the 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, we find that $QIDRes$ possesses significant incremental predictive power for imminent information events relative other liquidity and information variables.

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

5.3 *QIDRes* and Information Content of Trades

We next relate the spikes in $QIDRes$ around information arrivals, discussed in Section 5.2, 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.[25] 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 6 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 $QIDRes$ 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 risk by decomposing the transaction cost associated with each trade, as captured by effective spread, into permanent and

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

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 5 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 6 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.

5.4 *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.[26] 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.

[26] 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.

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 risk to high $QIDRes$.[27] 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 risk measures, 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 risk measures in each bi-weekly period.[28] We finally plot the time-series means of these averages against change-in-short-interest portfolios.

Figure 7 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 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 \-shaped pattern. Panels D and E in Figure 7 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

[27] Nonetheless, Appendix A.3 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 $QIDRes$.

[28] 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 $QIDRes$.

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,[29] 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.[30] 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 5 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 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.

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

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

5.5 $QIDRes$ 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. Panel A in Table 6 shows that the high $QIDRes$ portfolio, containing stock-days with abnormally low undercutting activity, have coefficients significantly 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 $QIDRes$ tercile to its high tercile. 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.

Panel B in Table 6 documents the extent of price reversals conditional on both $QIDRes$ and realized volatility of 1-minute returns based on midpoint prices, $qvol$. This analysis addresses the possible link between volatility and undercutting activity, reflecting reduced liquidity provision, and hence undercutting, when volatility is high. We sort each cross-section ***independently*** into terciles of $QIDRes$ and realized volatility, before estimating the extent of price reversals conditional on both. First, we observe that roughly equal number of observations fall in the nine $QIDRes$-volatility categories, indicating a near-zero correlation between abnormal undercutting and realized volatility—in fact, the correlation coefficient in the full sample is -0.0011 (see Panel B in Table 1). Second, the finding that highest $QIDRes$ tercile is associated with weakest subsequent reversals extends across different levels of realized volatility.

5.6 Intraday Analysis of $QIDRes$

In this section, we analyze the relationship between $QIDRes$ and informed trading risk by examining this link at different times of the trading day. Our analysis is motivated by the premise that information asymmetry, and the liquidity providers' risk of trading with informed investors, declines over the course of the trading day (see, e.g., Madhavan, Richardson, and Roomans (1997)).[31] We construct three "intraday" versions of $QIDRes$ that reflect undercutting activity at three time-of-day segments of the trading day. First, we inspect the correlation between $QIDRes_{jt}$ and each of these intraday versions. Second, we compare the behaviors of intraday $QIDRes$ measures around information events.

To construct intraday $QIDRes$, we divide each trading day into three segments: 9:45am–11:45am (morning, am), 11:45am–1:45pm (mid-day, md), and 1:45pm–3:45pm (afternoon, pm), which allows us to construct the three respective intraday undercutting activity measures $QID(\tau)^q_{jt}$, with $\tau \in \{am, md, pm\}$. Quarter $q-1$ quantities of these intraday undercutting activity measures are then entered, in turn, on the left hand side of equation (3).[32] The resulting parameter estimates as well as standard deviations of intraday QID measures enter equation (4) to produce $QIDRes(am)^q_{jt}$, $QIDRes(md)^q_{jt}$, and $QIDRes(pm)^q_{jt}$. This process decomposes $QIDRes$ on each stock day into its intraday components.

If $QIDRes$ captures informed trading risk and if such risk is higher in earlier trading hours of the

[31] Also see Admati and Pfleiderer (1988) and Wood, McInish, and Ord (1985), among others.

[32] We use the same right-hand-side variable in equation (3) when constructing different intraday versions of QID. This allows us to attribute any differences in the resulting $QIDRes$ measures to time-of-day effects in undercutting rather than those in quoted spreads.

trading day then we expect our baseline $QIDRes_{jt}^{q}$ to be more strongly correlated with its morning component, $QIDRes(am)_{jt}^{q}$, than with the other two components. We find strong evidence of this. Figure 8 exhibits empirical distributions of R^2 statistics obtained from regressing $QIDRes$ on each of its intraday components. These estimates are carried out at the stock-quarter level, capturing the association between $QIDRes$ and the intraday component *only* using time-series variations. Consistent with a declining informed trading risk over the course of the trading day, the association between $QIDRes$ and $QIDRes(am)$ is strongest and that between $QIDRes$ and $QIDRes(pm)$ is the weakest. Importantly, the clearly distinguishable locations of R^2 empirical distributions given different τ's is evidence of statistical dominance, which strongly speaks to the statistical and economic significance of our findings. More concretely, the mean (median) stock-quarter-specific R^2's are 65.8% (69.5%), 57.2% (59.9%), and 47.1% (47.7%) when variation in $QIDRes$ is examined against that in the underlying component from morning, mid-day, and evening, respectively. In sum, a much larger portion of the variation in $QIDRes$ is attributable to abnormal undercutting activity in earlier trading hours rather than later windows.

We provide additional evidence using the intraday variation in the intensity of informed trading by examining $QID(\tau)_{jt}^{q}$'s behavior around unscheduled press releases.[33] With higher intensity of informed trading earlier in the day, we expect $QIDRes(am)_{jt}^{q}$ to display greater spikes around information events than do other intraday versions of $QIDRed$. Figure 9 documents exatly this.

5.7 Asset Pricing Implications of $QIDRes$

The literature has documented that informed trading risk measures predict stock returns: higher past informed trading probability/intensity is associated with higher expected returns. However, there is no theoretical or empirical consensus regarding what drives this return predictability. For example, Easley and O'Hara (2004) argue that informed trading should be priced since the risk driven by information asymmetry is non-diversifiable; hence, investors holding a stock with more private information, and hence informed trading, demand a premium as compensation for this exposure. Consistent with this prediction, Easley et al. (2002) show that PIN is priced in the cross-section.[34] Duarte and Young (2009) propose an alternative explanation for return

[33] Qualitative similar conclusions obtain around earnings announcements and other news arrivals.

[34] 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

predictability of informed trading intensity/probability measures by showing that PIN's cross-sectional return predictability primarily reflects liquidity premia. They argue that since informed trading intensity is correlated with liquidity, PIN's return predictability conflates the effects of information asymmetry with those of priced illiquidity (Amihud and Mendelson (1980)). Following this literature, we also show that $QIDRes$ predicts stock returns. However, we attribute this return predictability to limits to arbitrage, reflecting the unique features of $QIDRes$.

In contrast to prior measures of informed trading, we do not expect any return predictability demonstrated by $QIDRes$ to be associated with compensation for bearing the risk associated with a stock characteristic or the premium demanded to hold less liquid stocks. In fact, we provide strong evidence that $QIDRes$ fits neither of these notions. Table 7 presents the correlations between $QIDRes$, ITI and PIN based information trading measures as well as common liquidity measures: quoted spread, effective spread, lambda, Amihud, and ILM. Panel A presents the correlations for 2010-2019 (omitting PIN measures where we only have data for 2010-2012) and Panel B presents all measures for the 2010-2012 period. This table shows virtually zero cross-sectional correlation between monthly averages of $QIDRes$ and various measures of liquidity, and only minimal correlation with other measures of informed trading. Panel A in Table A.1 presents evidence that $QIDRes$ is very weakly correlated with a host of stock characteristics. Finally, in Panel B of Table A.1, we document evidence of slight mean-reversion in $QIDRes$, indicating that is does not constitute a persistent stock characteristic.

Importnatly, the lack of correlation with liquidity is not true for other measures of informed trading where different versions of ITI and PIN appear to be positively related to liquidity. For example, Panel A shows that the average of the absolute correlation coefficients obtained between different versions of ITI and various stock illiquidity measures is about 0.15, with the highest pairwise absolute correlation of 0.37. Similarly, the average absolute correlation between different versions of PIN and stock illiquidity measures is around 0.15, with a high pairwise absolute correlation of 0.26. These collective facts clearly distinguish $QIDRes$ from existing measures, strongly suggesting that it cannot predict returns in the context of existing theories on return predictability of informed trading risk. We next investigate whether $QIDRes$ predicts returns.

priced, with Armstrong et al. (2011) providing empirical evidence supportive of this prediction.

We begin this analysis using simple portfolio sorts.[35] Table 8 shows that stocks with higher $QIDRes$ feature higher expected returns. For example, we find that average three-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 $q \quad 1$, is 30bps 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 $q \quad 1$. These quantitative findings extend when we form test portfolios using $QIDRes$ in quarter $q \quad 2$. Bogousslavsky et al. (2023) document next-month return predictability using ITIs; hence, complementary to their results, our finding that $QIDRes$ predicts monthly returns two quarters forward indicates that $QIDRes$ can predict future returns over longer horizons.

We next fit cross-sectional regressions to examine return predictability of $QIDRes$ while controling for key stock characteristics. Our regression analysis estimates

$$RetRf_{j;q;m} = \gamma^0 + \gamma^1 \left(QIDRes_{j;q \quad 1}\right) + \gamma^2 \left(QIDRes_{j;q \quad 2}\right) + \Lambda^{>} \mathrm{Control}_{j;q;m \quad 1} + u_{j;q;m}, \tag{6}$$

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; $QIDRes_{j;q \quad 1}$ and $QIDRes_{j;q \quad 2}$ denotes abnormal undercutting activity in quarters $q \quad 1$ and $q \quad 2$, respectively, for stock j; $\mathrm{Control}_{j;q;m \quad 1}$ 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 $m \quad 2$.

Table 9 summarizes our findings when we fit fixed-effect panel regressions based on equation (6): we find a statistically significant positive association between $QIDRes$ and expected stock 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 be-

[35] 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 5.1). 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. 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.

tween 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 10 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 (6) 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, ITIs from quarters q 1 and q 2 may, by construction, contain information about future returns. In sharp contrast, average $QIDRes$ from quarters q 1 and q 2 are not conditional on any future trading or pricing outcome.

As discussed earlier, we may interpret the robust return predictability of $QIDRes$ neither in the context of Easley and O'Hara (2004)'s "stock characteristic" story, nor in the context of Duarte and Young (2009)'s "illiquidity premia" story. This leads us to attribute the return predictability of $QIDRes$ to limits to arbitrage. Specifically, $QIDRes$ does not differentiate between positive and negative information, so if informed traders acting on positive and negative information is equally likely, then we would not expect $QIDRes$ to have any association with future returns. However, reflecting the well-documented selling constraints (e.g. Saffi and Sigurdsson (2011), and Dixon (2021)), it must be more difficult for investors to trade on negative information. As a result, high $QIDRes$ is more likely to capture informed trading motivated by positive, rather than negative, signals; and thus should positively predict returns.[36] Specifically, stocks with higher $QIDRes$ in a given quarter (1) experienced more information events than is normal in those quarters, and (2) due to short selling constraints, these information events were, on average, positive.

We conclude by showing that return predictability of $QIDRes$ is concentrated among stocks with tighter short sale constraints. We do so by splitting the sample based on observed equilibrium

[36]See Bogousslavsky et al. (2023) for a similar discussion.

lending fees in the securities lending markets. We examine $QISRes$'s return predictability conditional on the level of lending fees, with higher such fees reflecting tighter 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 quarter into terciles of this average security lending fee. Table 11 shows that $QIDRes$ predicts expected returns more strongly among stocks with high lending fees.

6 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 $QIDRes$. 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 indirectly measures informed trading risk.

We contrast $QIDRes$ with existing measures of informed trading intensity/probability whose constructions are computationally demanding, require proprietary data, or are applicable to only a subset of stock-days. 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 following days when $QIDRes$ indicates higher informed trading risk; (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 $QIDRes$ from the preceding two quarters predicts monthly stocks returns. 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

Admati, A. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies 1*, 3–40.

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.

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., Z. R. Huszar, and B. D. Jordan (2010). The good news in short interest. *Journal of Financial Economics 96* (1), 80–97.

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.

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. (2021). Why do short selling bans increase adverse selection and decrease price efficiency? *The review of asset pricing studies 11*(1), 122–168.

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.

Dixon, P. N. and E. K. Kelley (2022). Business cycle variation in short selling strategies: Picking during expansions and timing during recessions. *Journal of Financial and Quantitative Analysis 57*(8), 3018–3047.

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 nancial 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.

Madhavan, A., M. Richardson, and M. Roomans (1997). Why do security prices change? a transaction-level analysis of nyse stocks. *The Review of Financial Studies 10*(4), 1035–1064.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of nancial Markets 16*(4), 712–740.

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

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

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

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.

Saffi, P. A. and K. Sigurdsson (2011). Price efficiency and short selling. *The Review of Financial Studies 24*(3), 821–852.

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.

Wood, R., T. McInish, and J. K. Ord (1985). An investigation of transactions data nyse stocks. *Journal of Finance 40*, 723–739.

Figures and Tables

Figure 2. 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, excluding stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment.

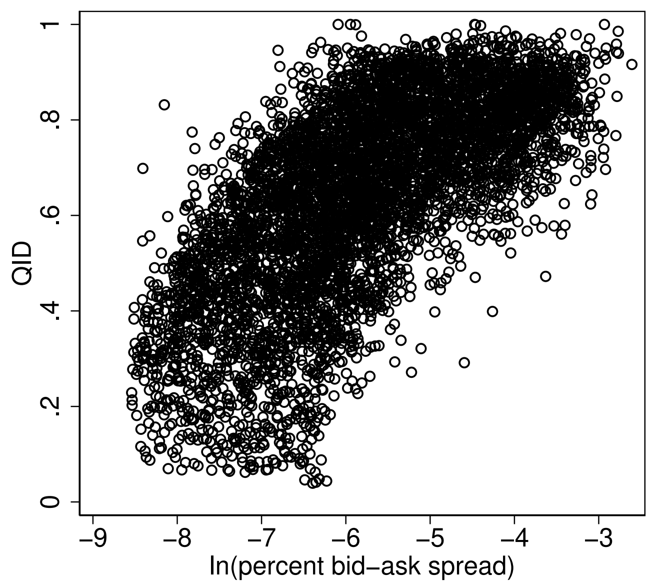
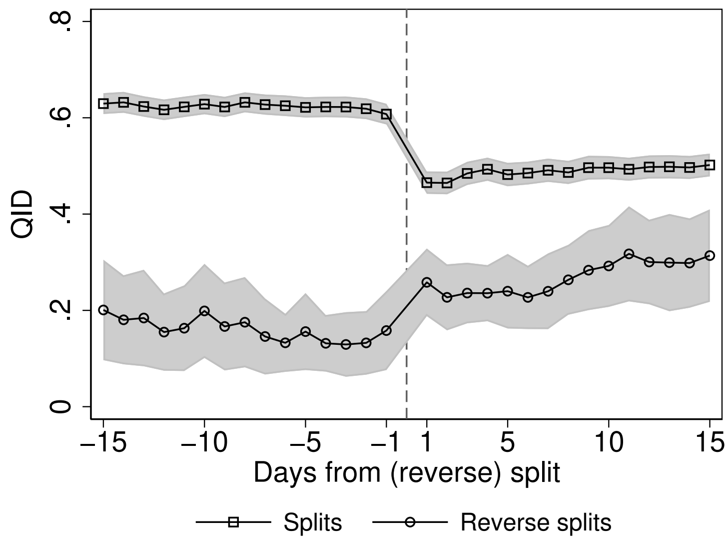


Figure 3. Relative Tick Size and Undercutting activity: Stock Splits and Reverse Splits. The figure presents average QID around stock splits. Stock split and revere-split dates are obtained from CRSP, with event windows covering 15 days prior to a split date and 15 day as of the split date. Averages and 95% confidence intervals of QID (Panel A) and relative quoted spread (Panel B), both winsorized at the 1^{st} and 99^{th} percentiles of each day if the main sample, are plotted against days from the event. The sample includes stock-days of NMS-listed common shares between Jan 01, 2010 through Dec 31, 2019 that coincide with stock-split event windows. Included stocks must minimum a daily closing price of $5 and must feature non-missing observations over the event window. Stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment are excluded.

Panel A: Undercutting activity measure



Panel B: Relative quoted spread

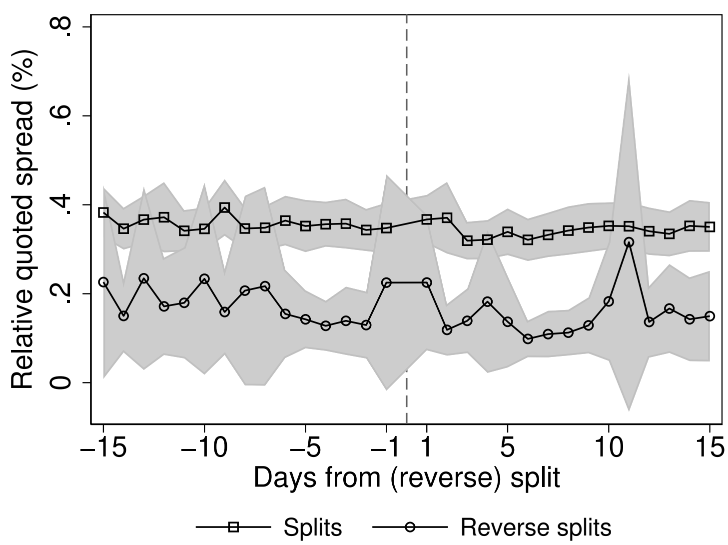


Figure 4. Existing Measures of Informed Trading around Unscheduled Corporate Announcements.
The figure presents medians of ITI, PIN, and MIA around around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Five versions of ITI 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. Stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment are excluded. Earnings announcement dates are obtained from COMPUSTAT; unscheduled press release dates and news arrivals not associated with any identified event are obtained from Ravenpack.

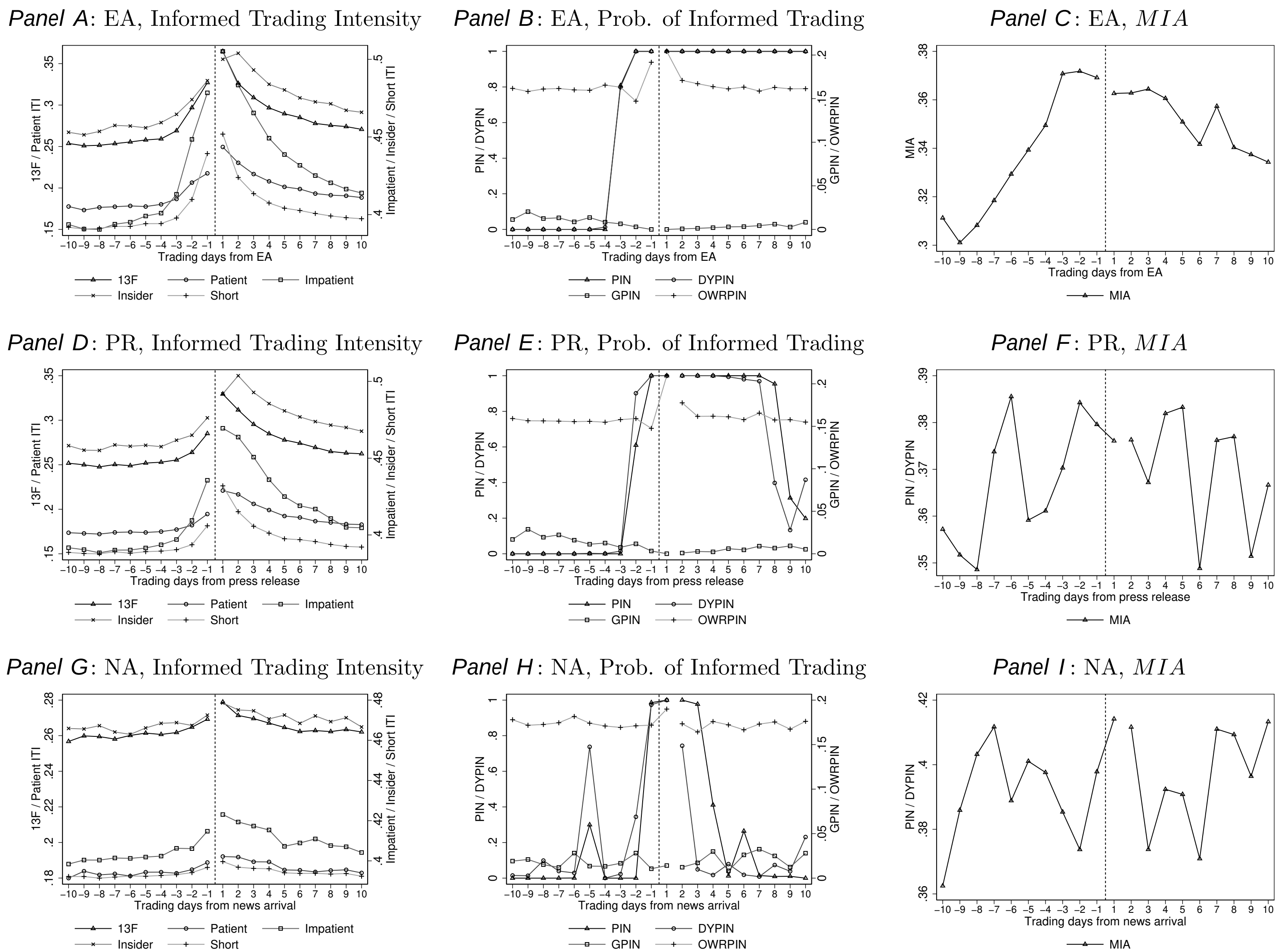


Figure 5. Undercutting Activity, Liquidity, and Information Asymmetry around Scheduled and Unscheduled Corporate Announcements.
The figure presents abnormal undercutting activity, dollar bid-ask spread, abnormal trading volume, and abnormal daily absolute return around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Daily abnormal undercutting values are calculated based on equation (4). 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.

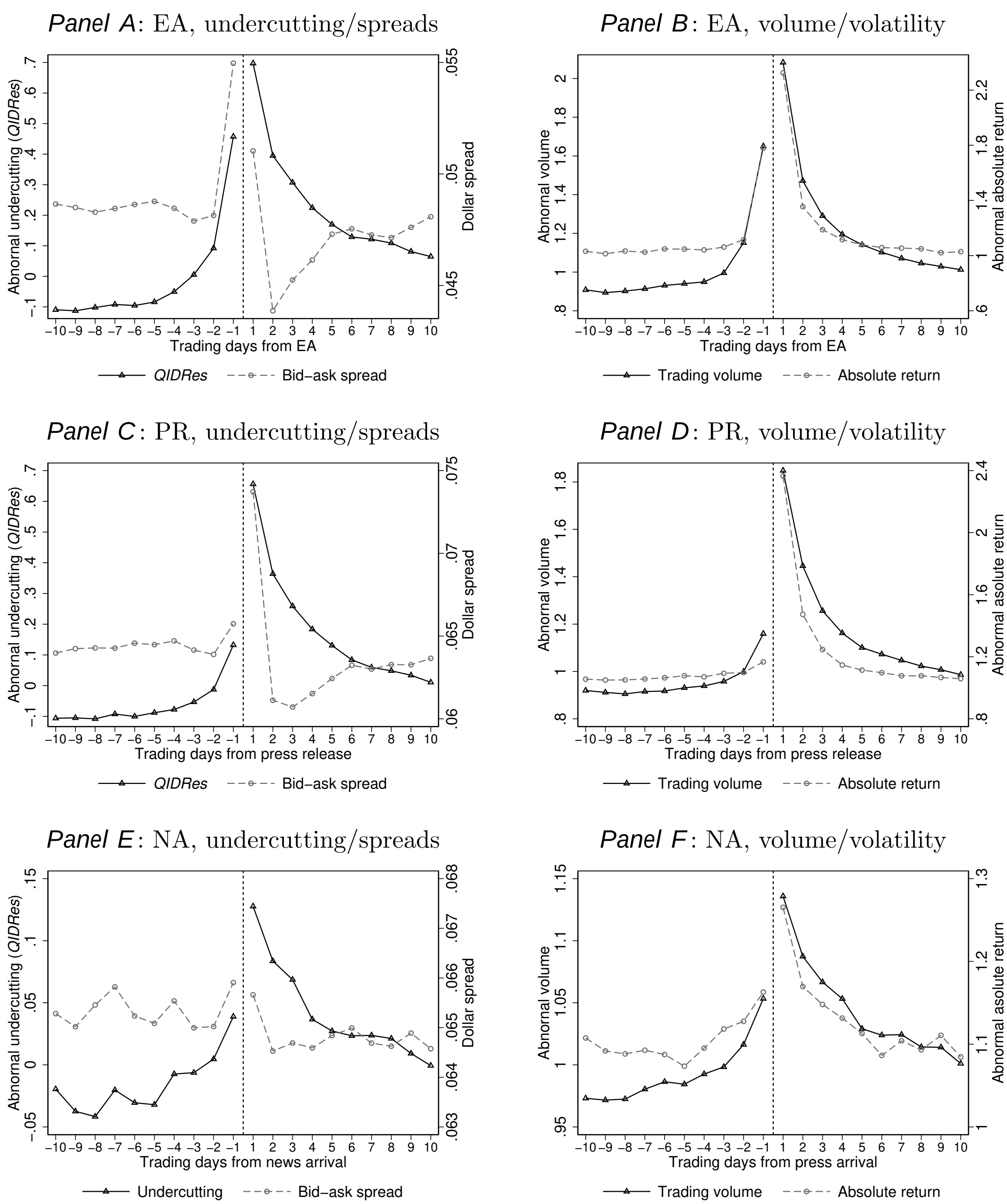


Figure 6. Undercutting Activity and Information Content of Trades, Events, and News.
Panels A through C present median abnormal undercutting activity around earnings announcements (EA), unscheduled press releases (PR), and news arrivals not associated with any identified event (NA). Earnings announcements 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 (4). 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, excluding stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. 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.

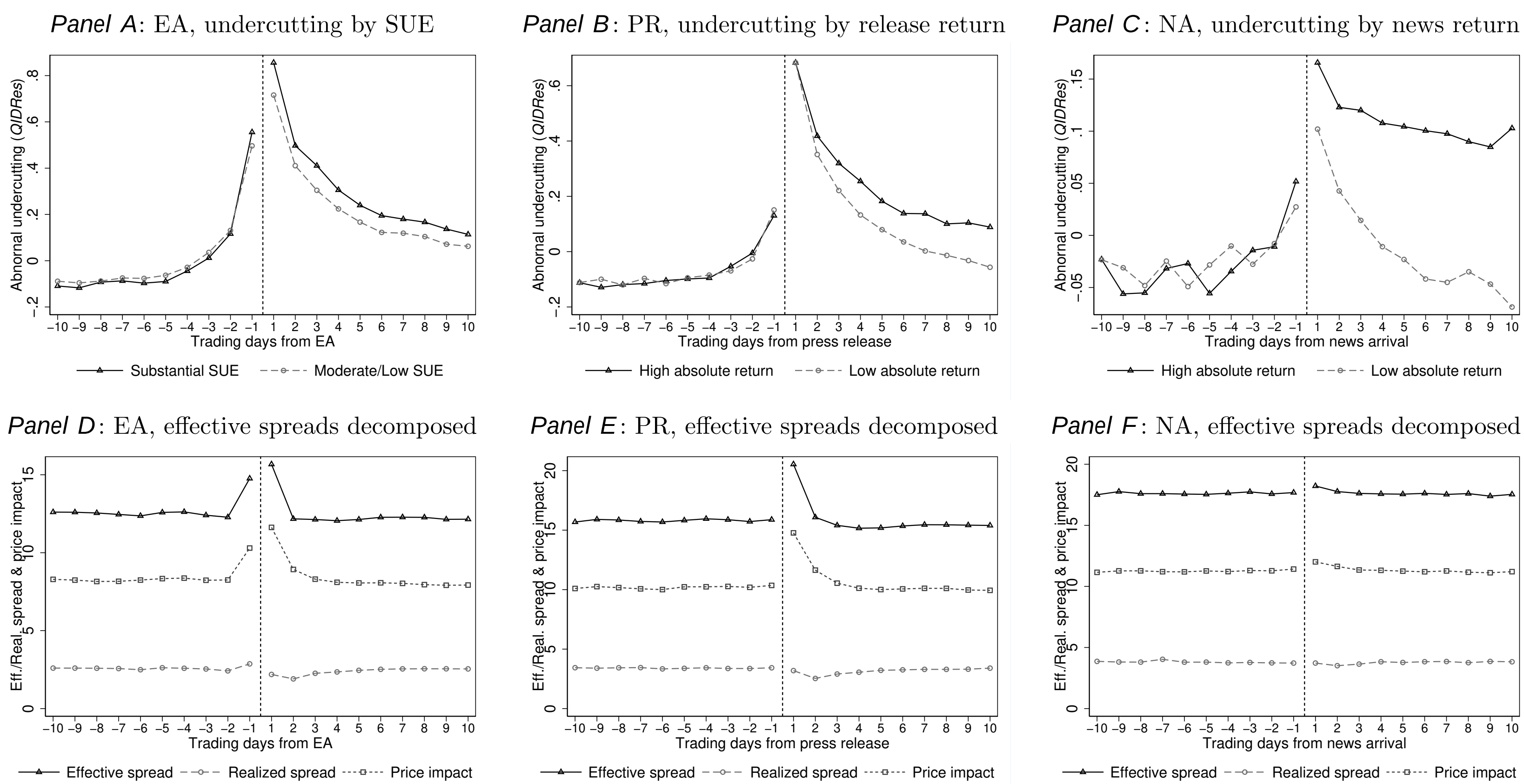


Figure 7. Informed Trading measures and Short Selling Activity.
The figure presents averages of various informed trading measures across levels of short selling activity. For averages of daily informed trading measures are calculated over bi-weekly intervals and matched with corresponding percentage 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. Daily $QIDRes$ observations are adjusted relative to the respective cross-stock average. 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, excluding stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment.

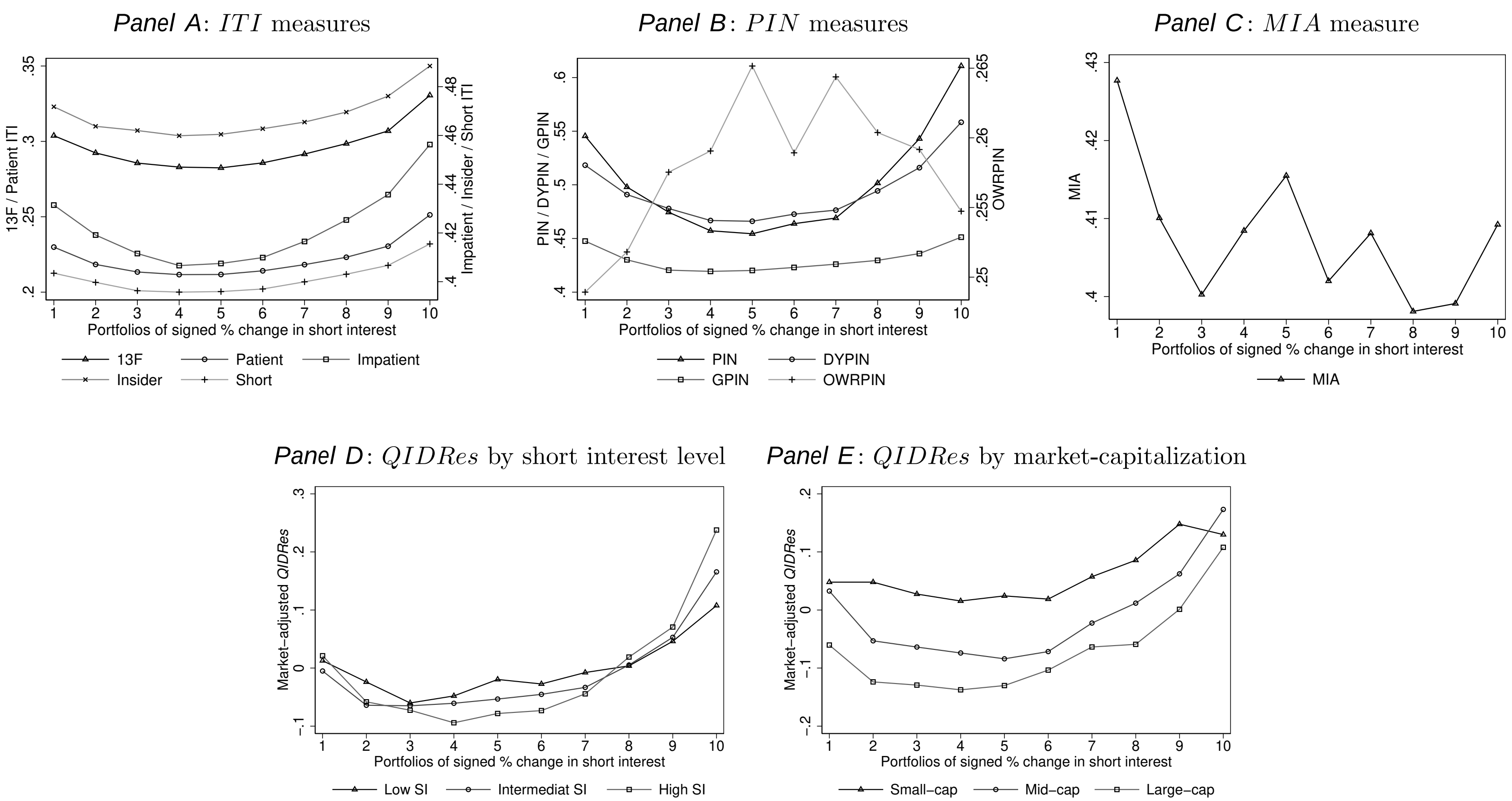


Figure 8. Intraday Sources of Variation in $QIDRes$.

The figure presents the decomposition of the variation in $QIDRes$ into into intraday components. In each quarter q and for each stock j, $QIDRes^q_{jt}$ is regressed on intraday component $QIDRes(\tau)^q_{jt}$, with τ 2 fam, md, pmg. The R^2 statistic from each regression for time-of-day τ is stored. The figure plots kernel densities for empirical distributions of R^2's across stock-quarters by τ. 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, excluding stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment.

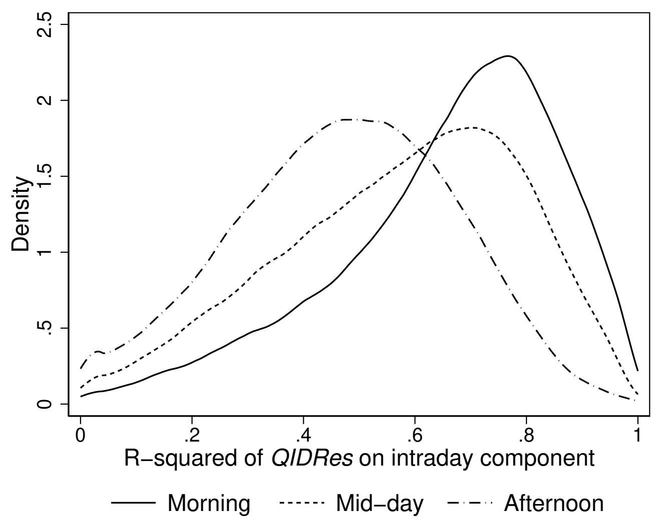


Figure 9. Undercutting Activity and Information Asymmetry around Unscheduled Corporate Announcements by Dime of Day.
The figure presents abnormal undercutting activity at different time-of-day windows around unscheduled press releases (PR). Intraday abnormal undercutting values are calculated based on equation (4) with $QID(\tau)$ refelecting undercutting activity in time-of-day window τ 2 fam, md, pmg. 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, excluding stocks-dates for firms designated as treatment or control stocks during the SEC's Tick Size Pilot experiment. Unscheduled press release dates are obtained from Ravenpack.

Panel A: Median $QIDRes$ by time of day

Panel B: Mean $QIDRes$ by time of day

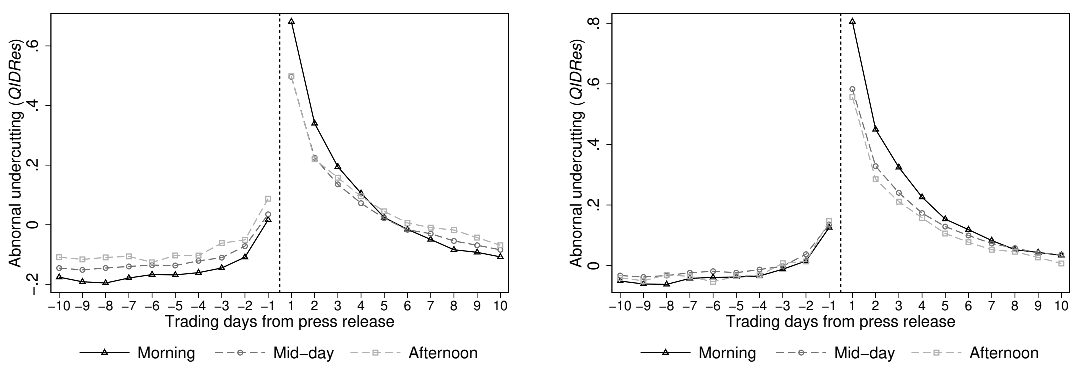


Table 1. Summary Statistics: Quote Revisions, QID, and $QIDRes$.

Panel A reports summary statistics of NBBO revisions as well as undercutting and abnormal undercutting measures. For each stock j on day t, NBBO improvements and deteriorations are counted separately for the bid (NBB) and ask (NBO) sides of the market. Trade-driven best quote deteriorations reflecting quote updates recorded no later that 10 milliseconds after a trade are constructed separately. For both categories, the share of single-tick updates devides the number of single-tick quote updates by all quote updates in the respective category. All quote improvements, $\#Impr_{jt}$, reflect the sum of the corresponding best bid and ask side improvements. Trade-drive quote deteriorations, $\#DeterTrade_{jt}$, reflect the sum of corresponding trade-driven best bid and ask deteriorations. The undercutting activity measure, QID, is constructed according to equation (2). Abnormal undercutting, $QIDRes$ is constructed according to equation (4). $QIDRes$ summary statistics are provided both before and after winsorizing each daily cross-section at the 1^{st} and 99^{th} percentiles. 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 B reports correlation coefficients between daily $QIDRes_{jt}$ and contempranous measures of quoted, effective, and realized spreads; price impacts, volatility, and trading volume.

Panel A: Summary statistics

Variable	Observations	Mean	S.D.	Skew	Percentiles 1st	5th	25th	50th	75th	95th	99th
All NBB revisions	6,662,352	1916.45	3167.51	29.51	2	17	337	1056	2394	6442	13066
NBB improvements	**6,662,352**	**1046.47**	**1717.63**	**29.90**	**1**	**9**	**183**	**579**	**1320**	**3511**	**7046**
Share of single-tick	6,662,352	*0.79*	0.28	1.96	0.00	0.00	0.77	*0.89*	0.96	1.00	1.00
NBB deteriorations	6,662,352	869.98	1479.88	32.02	0	7	151	469	1068	2955	6108
Trade-driven NBB deteriorations	**6,662,352**	**279.05**	**523.87**	**9.09**	**0**	**0**	**25**	**110**	**323**	**1096**	**2268**
Share of single-tick	6,662,352	*0.58*	0.34	0.37	0.00	0.00	0.33	*0.61*	0.91	1.00	1.00
All NBO revisions	6,662,352	1924.87	3146.58	16.76	1	17	341	1066	2403	6464	13091
NBO improvements	**6,662,352**	**1052.88**	**1715.27**	**17.70**	**1**	**9**	**184**	**584**	**1326**	**3530**	**7081**
Share of single-tick	6,662,352	*0.79*	0.28	1.95	0.00	0.00	0.76	*0.89*	0.96	1.00	1.00
NBO deteriorations	6,662,352	872.00	1464.30	18.03	0	7	153	473	1070	2958	6100
Trade-driven NBO deteriorations	**6,662,352**	**277.70**	**521.51**	**8.95**	**0**	**0**	**24**	**109**	**322**	**1094**	**2265**
Share of single-tick	6,662,352	*0.57*	0.34	0.35	0.00	0.00	0.32	*0.61*	0.91	1.00	1.00
QID	**6,662,352**	**0.61**	**0.27**	**0.42**	**0.03**	**0.12**	**0.41**	**0.64**	**0.84**	**0.99**	**1.00**
$QIDRes$	**6,662,352**	**0.07**	**1.52**	**1.53**	**3.46**	**1.85**	**0.70**	**0.01**	**0.77**	**2.20**	**4.23**
$QIDRes$ (winsorized at $1^{st}/99^{th}$ percentile)	**6,662,352**	**0.07**	**1.38**	**1.46**	**3.12**	**1.85**	**0.70**	**0.01**	**0.77**	**2.20**	**3.77**

Panel B: Correlation coefficients between daily $QIDRes$ and contemporanous microstructure outcomes

Microstructure outcome	Quoted spread Dollar	Quoted spread Relative	Effective spread Dollar	Effective spread Relative	Realized spread Dollar	Realized spread Relative	Piece Impact Dollar	Piece Impact Relative	Volatility Realized	Volatility jDaily returnj	Trading volume
Correlation coefficient	0.0465	0.059	0.0001	0.005	0.0001	0.0032	0.0025	0.0052	0.0011	0.0009	0.0045

Table 2. 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 (5) 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)
$Pilot$ $Event$	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
$Impr$								
$Pilot$ $Event$	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]
$DeterTade$								
$Pilot$ $Event$	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)
$Pilot$ $Event$	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
$Impr$								
$Pilot$ $Event$	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]
$DeterTade$								
$Pilot$ $Event$	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 3. 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 jrj as well as arrivals of information events, including earnings announcements (EA); press releases (PR); or news arrivals (NA) over days t 5 through t 1, specified using indicator variables $I(Inf)_{t\ \ 1}$ through $I(Inf)_{t\ \ 5}$. 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
$\Delta QIDRes_{t-1}$	**0.049***	**0.052***	**0.043***	**0.033***	**0.040***	**0.042***	**0.062***	**0.030***	**0.052***	**0.079***
	[9.50]	**[12.27]**	**[9.65]**	**[7.79]**	**[10.51]**	**[10.39]**	**[13.21]**	**[5.93]**	**[9.52]**	**[18.80]**
Observations	285,847	408,344	402,579	434,403	482,783	502,762	447,180	226,810	260,723	575,013

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
$\Delta QIDRes_{t-1}$	**0.037***	**0.043***	**0.028***	**0.026***	**0.028***	**0.033***	**0.048***	**0.026***	**0.045***	**0.057***
	[6.52]	**[9.53]**	**[5.87]**	**[5.57]**	**[6.60]**	**[7.50]**	**[9.62]**	**[4.78]**	**[7.80]**	**[12.65]**
$\Delta qsp_{t\ \ 1}$	0.33***	0.29***	0.26***	0.092	0.071	0.16**	0.11	0.094	0.34***	0.031
	[2.59]	[2.98]	[3.68]	[1.44]	[1.37]	[2.51]	[1.28]	[0.99]	[4.51]	[0.63]
$\Delta tv_{t\ \ 1}$	0.039***	0.034***	0.065***	0.045***	0.052***	0.050***	0.055***	0.060***	0.045***	0.062***
	[9.96]	[10.30]	[15.96]	[10.94]	[13.20]	[12.77]	[13.19]	[12.24]	[9.54]	[13.66]
$\Delta jrj_{t\ \ 1}$	0.018***	0.021***	0.014***	0.0050	0.0041	0.0035	0.011***	0.020***	0.013***	0.0030
	[5.31]	[7.29]	[3.55]	[1.33]	[1.23]	[1.11]	[3.14]	[3.64]	[2.99]	[0.98]
$I[Inf]_{t\ \ 1}$	0.73***	0.63***	0.85***	0.64***	0.71***	0.92***	0.95***	0.56***	0.67***	1.03***
	[34.68]	[37.60]	[44.27]	[37.78]	[46.58]	[58.20]	[53.12]	[26.59]	[32.06]	[66.36]
$I[Inf]_{t\ \ 2}$	0.093***	0.035*	0.0026	0.0036	0.028*	0.071***	0.062***	0.040*	0.14***	0.12***
	[4.06]	[1.95]	[0.12]	[0.20]	[1.67]	[4.00]	[3.10]	[1.81]	[6.55]	[6.87]
$I[Inf]_{t\ \ 3}$	0.042*	0.047***	0.012	0.0066	0.035**	0.053***	0.083***	0.080***	0.11***	0.11***
	[1.81]	[2.59]	[0.56]	[0.36]	[2.09]	[2.95]	[4.09]	[3.60]	[5.20]	[6.44]
$I[Inf]_{t\ \ 4}$	0.041*	0.055***	0.0020	0.045**	0.0065	0.041**	0.046**	0.070***	0.072***	0.16***
	[1.75]	[3.00]	[0.09]	[2.41]	[0.39]	[2.30]	[2.27]	[3.13]	[3.25]	[9.11]
$I[Inf]_{t\ \ 5}$	0.10***	0.17***	0.035	0.055***	0.070***	0.059***	0.062***	0.059***	0.12***	0.19***
	[4.34]	[9.67]	[1.62]	[3.01]	[4.17]	[3.29]	[3.03]	[2.67]	[5.56]	[11.09]
Observations	275,157	395,593	387,235	417,403	465,960	486,121	433,618	221,401	254,999	558,273

Table 4. 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 jrj as well as arrivals of information events, including earnings announcements (EA); press releases (PR); or news arrivals (NA) over days t 5 through t 1, specified using indicator variables $I(Inf)_{t\ 1}$ through $I(Inf)_{t\ 5}$. 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
$\Delta QIDRes_{t-1}$	**0.0096****	**0.018*****	**0.014*****	**0.0095*****	**0.019*****	**0.014*****	**0.019*****	**0.0042**	**0.0097*****	**0.020*****
	[2.04]	**[4.85]**	**[3.99]**	**[2.82]**	**[5.61]**	**[4.00]**	**[5.18]**	**[1.11]**	**[2.61]**	**[6.77]**
Observations	264,162	392,899	390,291	434,513	469,206	486,223	424,563	223,100	260,629	584,506

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
$\Delta QIDRes_{t-1}$	**0.0054**	**0.013*****	**0.0079****	**0.0057**	**0.0083****	**0.0079****	**0.014*****	**−0.00014**	**0.0036**	**0.0060***
	[1.07]	**[3.34]**	**[2.14]**	**[1.60]**	**[2.27]**	**[2.07]**	**[3.43]**	**[−0.03]**	**[0.93]**	**[1.95]**
$\Delta qsp_{t\ 1}$	0.26**	0.13	0.26***	0.16***	0.060	0.11*	0.099	0.066	0.21***	0.057*
	[1.98]	[1.38]	[4.60]	[3.22]	[1.25]	[1.82]	[1.37]	[0.93]	[4.13]	[1.77]
$\Delta tv_{t\ 1}$	0.021***	0.012***	0.025***	0.026***	0.040***	0.029***	0.035***	0.035***	0.026***	0.031***
	[6.39]	[4.24]	[7.44]	[7.57]	[11.47]	[8.88]	[10.66]	[9.53]	[7.49]	[9.40]
$\Delta jrj_{t\ 1}$	0.012***	0.014***	0.011***	0.021***	0.017***	0.018***	0.0058**	0.0030	0.00069	0.0100***
	[3.69]	[5.48]	[3.28]	[7.16]	[5.47]	[6.42]	[1.99]	[0.69]	[0.24]	[4.75]
$I[Inf]_t$ 1	0.17***	0.21***	0.24***	0.21***	0.27***	0.26***	0.23***	0.21***	0.29***	0.38***
	[9.19]	[15.06]	[16.79]	[15.86]	[19.99]	[19.69]	[16.23]	[13.84]	[21.66]	[36.12]
$I[Inf]_t$ 2	0.16***	0.13***	0.19***	0.14***	0.24***	0.18***	0.13***	0.19***	0.19***	0.23***
	[8.66]	[9.02]	[12.96]	[10.99]	[17.76]	[13.10]	[8.74]	[12.68]	[14.16]	[21.82]
$I[Inf]_t$ 3	0.052***	0.074***	0.15***	0.15***	0.23***	0.12***	0.13***	0.099***	0.16***	0.16***
	[2.74]	[5.12]	[9.97]	[11.65]	[17.52]	[8.66]	[9.13]	[6.53]	[11.84]	[14.57]
$I[Inf]_t$ 4	0.053***	0.099***	0.10***	0.050***	0.11***	0.11***	0.081***	0.11***	0.16***	0.15***
	[2.75]	[6.86]	[6.79]	[3.77]	[8.19]	[8.16]	[5.50]	[7.19]	[11.96]	[14.25]
$I[Inf]_t$ 5	0.13***	0.16***	0.15***	0.12***	0.11***	0.16***	0.15***	0.13***	0.17***	0.18***
	[6.95]	[11.28]	[10.49]	[9.07]	[8.26]	[11.58]	[10.09]	[8.43]	[12.17]	[17.06]
Observations	254,568	380,563	375,202	418,031	450,407	468,964	411,449	218,048	254,752	566,665

Table 5. 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: $Y_t^j = \eta_0 + \eta_i \quad I(t \quad i, t+i)_t^j + \epsilon_t^j$, where $I(t \quad i, t+i)_t^j$ is an indicator function that equals 1 in the i 2 $f0, 1, 2g$ 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 trade window	Informed trading measure $QIDRes$	ITI_{13D}	$ITI_{patient}$	$ITI_{impatient}$	$ITI_{insider}$	ITI_{short}	PIN	$DYPIN$	$GPIN$	$OWRPIN$	MIA
t	**0.083*** **	0.014***	0.0058***	0.015***	0.0054***	0.0074***	0.022***	0.031***	0.0043	0.0054**	0.0035
	[5.46]	[7.73]	[3.49]	[10.45]	[3.07]	[10.60]	[3.75]	[4.79]	[0.68]	[2.15]	[0.52]
$[t \quad 1, t+1]$	**0.071*** **	0.0072***	0.0028**	0.0084***	0.0030**	0.0045***	0.016***	0.019***	0.0014	0.0067**	0.00097
	[3.94]	[4.73]	[2.03]	[6.72]	[2.44]	[8.03]	[3.03]	[3.82]	[0.31]	[1.97]	[0.25]
$[t \quad 2, t+2]$	**0.066*** **	0.0056***	0.0015	0.0068***	0.0025**	0.0036***	0.011**	0.017***	0.0015	0.0067**	0.0040
	[3.94]	[4.17]	[1.22]	[5.92]	[2.40]	[7.23]	[2.31]	[3.92]	[0.39]	[2.03]	[1.17]
Sample mean	0.0112	0.3041	0.2225	0.4395	0.4401	0.4252	0.5679	0.5317	0.4337	0.2712	0.3128

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

INSFIT trade window	Informed trading measure $QIDRes$	ITI_{13D}	$ITI_{patient}$	$ITI_{impatient}$	$ITI_{insider}$	ITI_{short}	PIN	$DYPIN$	$GPIN$	$OWRPIN$	MIA
t	**0.068*** **	0.013***	0.0080***	0.013***	0.0011	0.0054***	0.040***	0.033***	0.011	0.0061*	0.0095
	[3.42]	[6.08]	[3.79]	[7.38]	[0.50]	[5.91]	[5.96]	[3.99]	[1.50]	[1.82]	[1.18]
$[t \quad 1, t+1]$	**0.056*** **	0.0097***	0.0048***	0.0091***	0.00056	0.0042***	0.022***	0.015***	0.0034	0.0035	0.0045
	[3.42]	[6.31]	[3.62]	[7.26]	[0.45]	[6.59]	[4.19]	[2.79]	[0.71]	[1.32]	[0.94]
$[t \quad 2, t+2]$	**0.050*** **	0.0075***	0.0045***	0.0077***	0.00061	0.0032***	0.017***	0.015***	0.0034	0.0019	0.0047
	[3.31]	[5.50]	[3.69]	[6.74]	[0.58]	[5.86]	[3.52]	[3.20]	[0.86]	[0.74]	[1.16]
Sample mean	0.0365	0.3061	0.2278	0.4401	0.4353	0.4243	0.5875	0.5415	0.4565	0.2713	0.3322

Table 6. Price Reversals by Abnormal Undercutting Activity and Realized Volatility.
This table reports the extent of price reversal over the next 10 trading days conditional on day t abnormal undercutting activity and realized volatility. For Panel A results, each daily cross-section is sorted into terciles of $QIDRes$. For each such tercile panel regressions of compound returns over the next $n \in \{1, 2, \ldots, 10\}$ days from day t's close, denoted $CR^j_{t;t+n}$, on day t returns, denoted R^j_t, are estimated. For Panel B results, each daily cross-section is sorted *independently* into terciles of $QIDRes$ and realized volatility, $qvol$. For each of the nine categories, panel regressions of compound returns over 1 day forward ($CR^j_{t;t+1}$), 5 days forward ($CR^j_{t;t+5}$), and 10 days forward ($CR^j_{t;t+10}$), on day t 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 $QIDRes$ tercile and n (Panel A) or by $QIDRes$ and volatility terciles (Panel B). 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 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: Price reversals up to 10 days forward by day-t $QIDRes$ tercile

	Dependent Variable									
$QIDRes$ tercile	$CR_{t;t+1}$	$CR_{t;t+2}$	$CR_{t;t+3}$	$CR_{t;t+4}$	$CR_{t;t+5}$	$CR_{t;t+6}$	$CR_{t;t+7}$	$CR_{t;t+8}$	$CR_{t;t+9}$	$CR_{t;t+10}$
Low	0.053***	0.057***	0.070***	0.078***	0.084***	0.088***	0.089***	0.093***	0.10***	0.097***
	[4.20]	[4.33]	[4.97]	[5.54]	[6.06]	[6.28]	[6.39]	[6.76]	[7.70]	[7.24]
Medium	0.051***	0.056***	0.066***	0.074***	0.076***	0.081***	0.083***	0.087***	0.092***	0.090***
	[3.53]	[3.81]	[4.32]	[4.83]	[5.02]	[5.33]	[5.62]	[5.87]	[6.62]	[6.32]
High	0.025***	0.030***	0.038***	0.044***	0.047***	0.050***	0.054***	0.054***	0.059***	0.057***
	[3.22]	[3.63]	[4.42]	[5.03]	[5.34]	[5.51]	[5.96]	[5.96]	[6.55]	[6.21]

Panel B: Price reversals 1, 5, and 10 days forward by day-t $QIDRes$ and realized volatility terciles

		Dependent variable								
		$CR_{t;t+1}$			$CR_{t;t+5}$			$CR_{t;t+10}$		
		Realized volatility tercile			Realized volatility tercile			Realized volatility tercile		
$QIDRes$ tercile		Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	Slope	0.056***	0.055***	0.046***	0.092***	0.084***	0.076***	0.099***	0.10***	0.090***
		[3.73]	[4.00]	[5.25]	[5.55]	[5.56]	[6.59]	[6.17]	[6.58]	[7.45]
	Observations	536,727	539,164	538,302	536,727	539,164	538,302	536,727	539164	538302
Medium	Slope	0.047***	0.051***	0.054***	0.070***	0.073***	0.085***	0.080***	0.092***	0.098***
		[4.59]	[2.76]	[3.96]	[5.96]	[3.81]	[5.64]	[6.49]	[5.08]	[6.69]
	Observations	538,961	538,924	540,782	538,961	538,924	540,782	538,961	538,924	540,782
High	Slope	0.024***	0.023***	0.026***	0.045***	0.044***	0.052***	0.053***	0.058***	0.060***
		[3.48]	[2.87]	[3.30]	[5.15]	[4.70]	[5.42]	[5.38]	[5.66]	[5.98]
	Observations	540,265	539,962	540,681	540,265	539,962	540,681	540,265	539,962	540,681

Table 7. 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}	0.10									
3	$ITI_{patient}$	0.12	0.79								
4	$ITI_{impatient}$	0.11	0.73	0.56							
5	$ITI_{insider}$	0.01	0.35	0.39	0.38						
6	ITI_{short}	0.14	0.44	0.36	0.63	0.17					
7	QSP	0.01	0.11	0.08	0.17	0.06	0.23				
8	$EFSP$	0.00	0.13	0.09	0.19	0.05	0.25	0.97			
9	$Lambda$	0.03	0.12	0.03	0.28	0.09	0.31	0.24	0.29		
10	AM	0.01	0.10	0.05	0.22	0.01	0.21	0.27	0.31	0.57	
11	ILM	0.03	0.18	0.05	0.35	0.05	0.37	0.37	0.42	0.60	0.47

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}	0.09													
3	$ITI_{patient}$	0.08	0.80												
4	$ITI_{impatient}$	0.11	0.75	0.59											
5	$ITI_{insider}$	0.02	0.34	0.31	0.41										
6	ITI_{short}	0.08	0.51	0.49	0.67	0.28									
7	PIN	0.04	0.33	0.33	0.38	0.13	0.53								
8	$DYPIN$	0.04	0.31	0.30	0.34	0.16	0.43	0.61							
9	$GPIN$	0.05	0.01	0.00	0.02	0.07	0.09	0.03	0.02						
10	$OWRPIN$	0.01	0.01	0.01	0.03	0.01	0.06	0.05	0.01	0.02					
11	QSP	0.02	0.04	0.08	0.04	0.06	0.21	0.17	0.11	0.17	0.07				
12	$EFSP$	0.02	0.04	0.08	0.04	0.06	0.23	0.18	0.12	0.17	0.09	0.93			
13	$Lambda$	0.00	0.05	0.05	0.13	0.16	0.24	0.22	0.12	0.21	0.21	0.28	0.28		
14	AM	0.00	0.03	0.02	0.10	0.02	0.17	0.14	0.09	0.11	0.13	0.14	0.15	0.66	
15	ILM	0.02	0.02	0.01	0.06	0.18	0.27	0.26	0.15	0.22	0.10	0.44	0.42	0.63	0.41

Table 8. 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 1 (Panel A) or from quarter q 2 (Panel B), 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.

Panel A: Monthy returns to portfolios $QIDRes$ from quarter q 1

	$QIDRes$ quintile					
Monthly portfolio return	Low	2	3	4	High	High Low
Excess return	1.00**	1.10**	1.15***	1.35***	1.18***	0.18
	[2.41]	[2.60]	[2.68]	[3.16]	[3.12]	[1.57]
3-factor alpha	0.20**	0.024	0.085	0.30***	0.095	0.30**
	[2.26]	[0.29]	[0.93]	[4.32]	[1.09]	[2.24]
4-factor alpha	0.17**	0.033	0.097	0.30***	0.11	0.28***
	[2.32]	[0.36]	[1.15]	[4.33]	[1.37]	[2.69]
6-factor alpha	0.21***	0.040	0.11	0.29***	0.11	0.32***
	[2.75]	[0.44]	[1.37]	[3.98]	[1.49]	[3.10]

Panel B: Monthy returns to portfolios $QIDRes$ from quarter q 2

	$QIDRes$ quintile					
Monthly portfolio return	Low	2	3	4	High	High Low
Excess return	1.04***	1.07***	1.21***	1.21***	1.28***	0.24*
	[2.69]	[2.93]	[3.08]	[3.03]	[3.08]	[1.72]
3-factor alpha	0.16	0.069	0.089	0.13	0.19***	0.35**
	[1.29]	[0.73]	[1.36]	[1.47]	[3.01]	[2.13]
4-factor alpha	0.19	0.072	0.073	0.13	0.17**	0.37*
	[1.46]	[0.78]	[1.28]	[1.55]	[2.13]	[1.98]
6-factor alpha	0.17	0.042	0.10**	0.14*	0.17**	0.34*
	[1.36]	[0.41]	[2.00]	[1.70]	[2.09]	[1.93]

Table 9. 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 (6) 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		
$QIDRes_{q\ 1}$	**0.17*** **	**0.21*** **	**0.23*** **	**0.21*** **	**0.21*** **	**0.21*** **	**0.21*** **	**0.21*** **	**0.21*** **	**0.23*** **
	[2.75]	**[3.17]**	**[3.52]**	**[3.23]**	**[3.22]**	**[3.17]**	**[3.16]**	**[3.21]**	**[3.27]**	**[3.60]**
$QIDRes_{q\ 2}$	**0.047**	**0.084**	**0.091***	**0.088**	**0.087**	**0.084**	**0.083**	**0.084**	**0.087**	**0.095***
	[0.84]	**[1.54]**	**[1.68]**	**[1.62]**	**[1.61]**	**[1.54]**	**[1.53]**	**[1.56]**	**[1.62]**	**[1.76]**
Illiquidity				1.13**	2.38***	0.0096	0.36	0.23		
				[2.39]	[2.91]	[0.08]	[1.02]	[0.46]		
β^{mkt}	0.13	0.26	0.29	0.26	0.25	0.26	0.25	0.26	0.26	0.28
	[0.45]	[1.28]	[1.45]	[1.28]	[1.26]	[1.28]	[1.26]	[1.31]	[1.30]	[1.44]
β^{hml}	0.18	0.15	0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.15
	[1.10]	[0.97]	[0.93]	[0.98]	[0.99]	[0.98]	[0.97]	[0.97]	[0.99]	[0.95]
β^{smb}	0.064	0.076	0.089	0.075	0.075	0.074	0.072	0.077	0.072	0.085
	[0.44]	[0.51]	[0.60]	[0.51]	[0.51]	[0.50]	[0.49]	[0.52]	[0.49]	[0.58]
BM	0.27**	1.02***	1.09***	1.00***	1.00***	1.02***	1.04***	1.02***	1.00***	1.07***
	[2.15]	[3.03]	[3.31]	[2.99]	[2.98]	[3.00]	[3.07]	[3.03]	[2.96]	[3.25]
ln(Mcap)	0.0098	2.45***	2.43***	2.41***	2.41***	2.45***	2.46***	2.44***	2.39***	2.39***
	[0.23]	[10.07]	[9.99]	[10.10]	[10.05]	[10.03]	[10.07]	[9.98]	[9.94]	[9.96]
DYD	0.38	0.42	0.16	0.69	0.70	0.43	0.44	0.39	0.67	0.47
	[0.19]	[0.21]	[0.08]	[0.35]	[0.35]	[0.21]	[0.22]	[0.20]	[0.34]	[0.24]
Id. Vol.	0.21**	0.072	0.058	0.067	0.065	0.073	0.070	0.070	0.064	0.052
	[2.37]	[1.00]	[0.83]	[0.93]	[0.90]	[1.01]	[0.97]	[0.98]	[0.88]	[0.72]
$RET_{\ 1}$	1.09	4.45***	4.46***	4.48***	4.48***	4.44***	4.44***	4.45***	4.48***	4.48***
	[1.01]	[4.09]	[4.09]	[4.11]	[4.11]	[4.09]	[4.10]	[4.09]	[4.12]	[4.12]
$RET_{(\ 12;\ 2)}$	0.36	1.77***	1.74***	1.73***	1.72***	1.77***	1.77***	1.76***	1.70***	1.69***
	[1.29]	[5.98]	[5.76]	[5.92]	[5.89]	[5.94]	[5.97]	[5.80]	[5.64]	[5.58]
$IOShr$	0.43***	0.94***	1.39***	0.97***	0.98***	0.93***	0.95***	0.93***	0.96***	1.41***
	[2.78]	[3.18]	[4.21]	[3.29]	[3.34]	[3.17]	[3.23]	[3.17]	[3.29]	[4.25]
$IOShr_H HI$			1.75***							1.69***
			[3.44]							[3.26]
TO			31.0**							33.1***
			[2.46]							[2.64]
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 10. The Cross-Section of Expected Returns and Informed Trading: Horse Race Regressions. This table reports on the relation informed trading measures and the cross-section of expected returns. equation (6) 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 9. 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			
$\boldsymbol{QIDRes_{q-1}}$	**0.22*** **	**0.20** **	**0.21** **	**0.23**	**0.20**	**0.21**	**0.25**
	[3.07]	**[2.39]**	**[2.50]**	**[1.53]**	**[1.37]**	**[1.34]**	**[1.15]**
$\boldsymbol{QIDRes_{q-2}}$	**0.092***	**0.10**	**0.087**	**0.28** **	**0.26** **	**0.27** **	**0.41*** **
	[1.67]	**[1.50]**	**[1.29]**	**[2.41]**	**[2.35]**	**[2.24]**	**[3.01]**
$ITI_{13D;q\ 1}$	0.35		0.47		1.16	1.05	3.96
	[0.29]		[0.30]		[0.34]	[0.31]	[1.12]
$ITI_{13D;q\ 2}$	1.94		2.00		4.41	4.33	2.41
	[1.53]		[1.31]		[1.65]	[1.62]	[0.73]
$ITI_{patient;q\ 1}$	1.05		1.19		**5.20***	**5.36***	**6.53***
	[0.80]		[0.65]		**[1.71]**	**[1.74]**	**[1.82]**
$ITI_{patient;q\ 2}$	0.24		1.31		0.55	0.63	0.86
	[0.18]		[0.76]		[0.20]	[0.23]	[0.25]
$ITI_{impatient;q\ 1}$	0.88		2.22		3.64	3.01	4.20
	[0.59]		[1.23]		[0.88]	[0.76]	[0.80]
$ITI_{impatient;q\ 2}$	1.11		**3.26***		0.39	0.67	3.98
	[0.81]		**[1.93]**		[0.13]	[0.24]	[1.30]
$ITI_{insider;q\ 1}$	1.57		**2.68***		0.060	0.46	2.33
	[1.41]		**[1.85]**		[0.02]	[0.13]	[0.51]
$ITI_{insider;q\ 2}$	**2.43*** **		**2.52** **		**5.18***	**4.87***	3.68
	[2.66]		**[2.12]**		**[2.00]**	**[1.84]**	[1.01]
$ITI_{short;q\ 1}$	1.14		0.81		2.56	4.16	7.20
	[0.41]		[0.22]		[0.34]	[0.53]	[0.68]
$ITI_{short;q\ 2}$	1.28		2.24		10.5	10.3	11.9
	[0.48]		[0.65]		[1.68]	[1.62]	[1.58]
$MIA_{q\ 1}$		**1.67*** **	**1.54*** **				0.88
		[3.13]	**[2.95]**				[0.55]
$MIA_{q\ 2}$		0.22	0.22				0.42
		[0.45]	[0.47]				[0.34]
$PIN_{q\ 1}$				0.13		0.25	0.18
				[0.21]		[0.37]	[0.22]
$PIN_{q\ 2}$				0.13		0.028	0.13
				[0.26]		[0.05]	[0.21]
$DYPIN_{q\ 1}$				0.60		0.68	0.12
				[0.98]		[1.09]	[0.17]
$DYPIN_{q\ 2}$				0.56		0.48	0.60
				[0.92]		[0.76]	[0.76]
$GPIN_{q\ 1}$				0.44		0.42	0.37
				[0.96]		[0.90]	[0.57]
$GPIN_{q\ 2}$				0.99*		1.00*	1.11
				[1.93]		[1.83]	[1.47]
$OWRPIN_{q\ 1}$				0.76		0.78	0.52
				[1.47]		[1.33]	[1.22]
$OWRPIN_{q\ 2}$				**0.83***		**0.83***	**0.74***
				[1.74]		**[1.74]**	**[1.90]**
Observations	216,077	119,098	118,113	25,045	25,045	25,045	16,065

Table 11. 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. Equation (6) is estimated within terciles of quarter q 3's average security lending fees obtained from FIS database from 2010 through 2018. The sample includes NMS common shares from January 2010 to December 2018, 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 set of stock characteristics is identical to that used in Table 9. 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.

Independent Variable	Tercile of security lending fee: Low		Intermediate		High	
$QIDRes_{q\ 1}$	0.16**	0.15**	0.23***	0.21***	0.39***	0.38***
	[2.24]	[2.19]	[3.00]	[2.87]	[3.31]	[3.24]
$QIDRes_{q\ 2}$	0.038	0.036	0.15*	0.14*	0.19	0.18
	[0.60]	[0.57]	[1.86]	[1.83]	[1.61]	[1.55]
Stock characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Liquidity controls	Yes	No	Yes	No	Yes	No
Observations	72,747	72,913	70,888	71,012	67,101	67,227

A Appendix

A.1 A Simple Framework of Undercutting and Informed Trading

Consider a simple one period rational expectation equilibrium model that builds off of Glosten and Milgrom (1985). An asset takes the equally likely value of 0 or 1. The fraction π of liquidity demanders are informed and know the true value of the asset only buying when the the value equals 1 and only selling when the value equals 0. The remaining $1 - \pi$ fraction of liquidity demanders are uninformed and buy and sell with equal probability. The exact arrival time of the next trade to arrive is random and follows an exponential distribution with arrival rate parameter λ. Liquidity providers come in two types: sophisticated and unsophisticated. Unsophisticated liquidity providers, denoted $ULPs$, are passive, competitive, and thus set prices equal to the conditional expected value of the asset. Sophisticated liquidity providers, denoted $SLPs$, can pay a cost c which will, with probability ρ inform them about whether the next trade to arrive is informed or uninformed and on which side of the market the trade will arrive. It does not inform them about the arrival time of the upcoming trade.[37] There are m SLPs where the value m is determined in equilibrium such that the expected profit associated with being an SLP is equal to the cost c, and so SLPs are competitive. The likelihood that at least one of the m $SLPs$ receives a signal is $\phi = 1 - (1 - \rho)^m$.

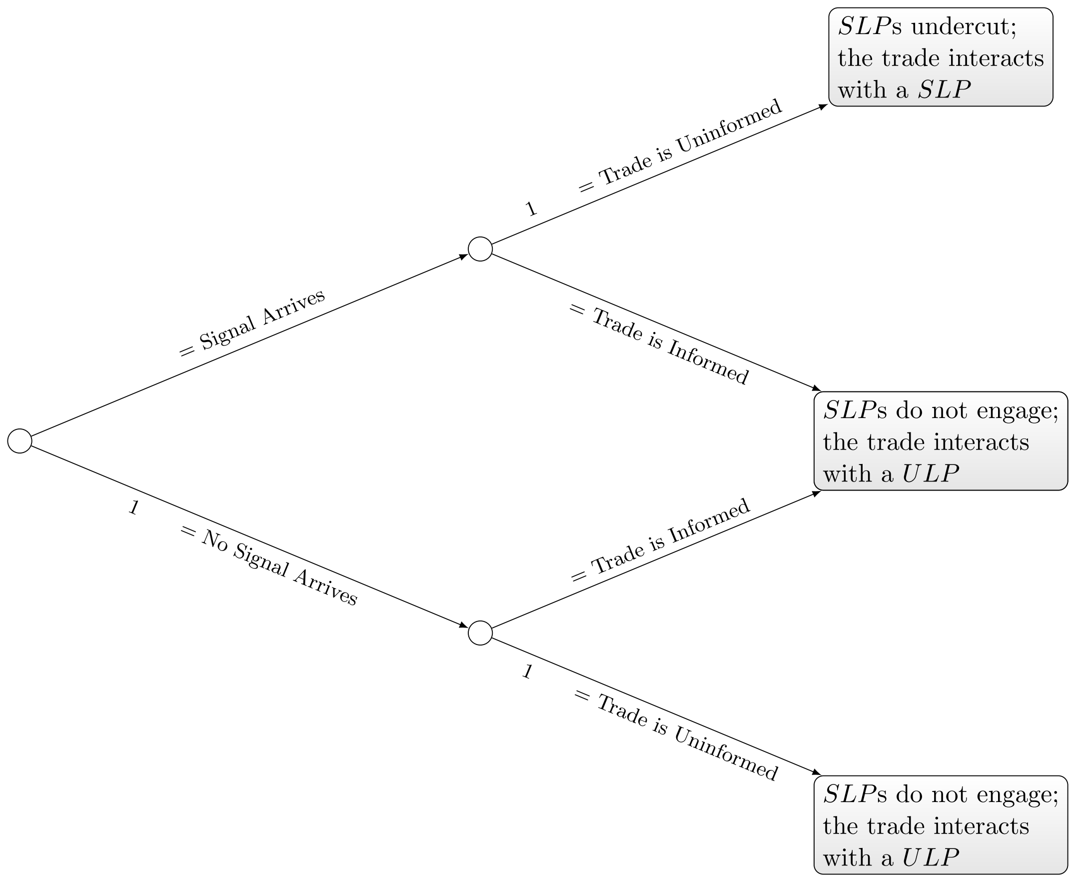
If a SLP receives a signal that an upcoming trade is informed, the SLP will simply sit out and not post any quotes allowing the ULPs to interact with the incoming informed trade. If no SLP receives a signal then all $SLPs$ sit out. However, if an SLP receives a signal that the upcoming trade is uninformed, they will undercut the existing quote on that side of the market. The other $SLPs$, whether they receive a signal or not, will observe this quote improvement and will infer that a signal has been received and will submit their own undercutting orders and an undercutting run will ensue.[38] The outcome tree in Figure A.1 illustrated this setup.

In this setup all informed trades interact with ULPs, and some uninformed trades interact with ULPs and some with SLPs. The probability that a ULP interacts with an informed trade is the probability of an informed trade arriving (π) divided by the probability that a trade interacts with a ULP, which is $1 - \phi(1 - \pi)$. The probability that a ULP interacts with an uninformed trade is simply the compliment as shown in equations A.1 and A.2,

[37]The cost c can be thought of as the cost of investing in the capacity to process, analyze, and respond quickly to information based in order flow.

[38]The assumption that all $SLPs$ can infer the signals of others via monitoring quote updates could be relaxed such that only those $SLPs$ receiving a signal engage in the undercutting run without changing any of the key inference. In this case could be redefined to be the probability that at least two $SLPs$ receive a signal $= 1 - (1 -)^m - m(1 -)^{m-1}$, and all inference remains exactly the same since in both cases is increasing in both m and . Additionally the profit to undercutting is random, since the arrival of the uninformed trade is random and so it is unclear exactly when during the run the uninformed trade will arrive. However, given that SLPs know the arrival rate of trades, they can compute the expected time during an undercutting run a trade will arrive and so can compute the expected profit of a run that is earned by the winning quote provider, which we denote Π. The likelihood that a given SLP wins the undercutting run is $\frac{1}{m}$, so expected profits to undercutting are $\frac{\ }{m}$. For this market to be in equilibrium it must be the case that $c = \frac{\ }{m}$ which implies that the number of SLPs is $m = \frac{\ }{c}$.

Figure A.1. Informed Trading Signal Arrivals and SLPs' Undercutting Choices.



$$P(I) = \frac{\pi}{1 \quad \phi(1 \quad \pi)}, \tag{A.1}$$

$$P(U) = \frac{(1 \quad \phi)(1 \quad \pi)}{1 \quad \phi(1 \quad \pi)}. \tag{A.2}$$

The bid and the ask prices are set by the ULPs equal to the expected value of the asset conditional on the trade occurring as shown in equations A.3 and A.4,

$$Ask = 1 \quad P(I) + \frac{1}{2}P(U), \tag{A.3}$$

$$Bid = 0 \quad P(I) + \frac{1}{2}P(U). \tag{A.4}$$

Inserting A.1 and A.2 into A.3 and A.4 renders,

$$Ask \;\; = \frac{1 + \pi \quad \phi(1 \quad \pi)}{2(1 \quad \phi(1 \quad \pi))}, \tag{A.5}$$

$$Bid \;\; = \frac{1 \quad \pi \quad \phi(1 \quad \pi)}{2(1 \quad \phi(1 \quad \pi))}. \tag{A.6}$$

$$Spread \;\; = \frac{\pi}{1 \quad \phi(1 \quad \pi)}. \tag{A.7}$$

With this framework it is straightforward to show that undercutting behavior is inversely related to informed trading risk. To see this, consider that the probability of an undercutting run is the probability that at least one SLP receives a signal,ϕ, multiplied by the likelihood that the upcoming trade is uninformed,$(1 \quad \pi)$,

$$P(UndercuttingRun) = \phi(1 \quad \pi). \tag{A.8}$$

The derivative of this value with respect to informed trading risk is $\frac{P(UndercuttingRun)}{} = \quad \phi$, which is always less than zero, confirming the inverse relation between undercutting activity and informed trading risk - i.e. when the risk of informed trading goes up, the likelihood of undercutting runs diminishes.

The model also produces an additional prediction: that the spread will be increasing in undercutting risk - i.e. liquidity gets worse as undercutting risk increases a result documented empirically(Foley et al. (2021), Foley et al. (2022)). To see this, consider that the spread from A.7 can be rewritten as,

$$Spread \;\; = \frac{\pi}{1 \quad P(UndercuttingRun)}. \tag{A.9}$$

$P(UndercuttingRun)$ is bounded by 0 and $1 \quad \pi$ it is straightforward to see that as $P(UndercuttingRun)$ increases, so too does the bid ask spread. The spread is bounded on the top by the value of 1. Thus, the prevalence of undercutting runs can cause markets to fail if SLPs interact with too many of the uninformed trades.[39]

A.2 Is $QIDRes$ a Stock Characteristic?

This section present evidence that $QIDRes$ is not persistent stock/firm characteristic. Panel A in Figure A.1 presents pairwise correlation coefficients between $QIDRes_{q \quad 1}$, $QIDRes_{q \quad 2}$ 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 $QIDRes_{q \quad 1}$ and $QIDRes_{q \quad 2}$ 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.

[39] In the case of no undercutting runs, i.e. $\;= 0$, $P(UndercuttingRun) = 0$, the spread will equal , its minimum given the prevalence of informed traders in the market. In the other extreme case where $\;= 1$, implying that all uninformed trades trigger undercutting runs, $P(UndercuttingRun) = 1$, and the spread goes to 1 indicating a failed market.

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, ($\mathrm{BM}_{j;m\ \ 1}$), 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 ($RET_{j;(m\ \ 12;m\ \ 2)}$), 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	$QIDRes_{q\ \ 2}$	0.056												
3	β^{mkt}	0.009	0.003											
4	β^{hml}	0.007	0.005	0.03										
5	β^{smb}	0.030	0.025	0.12	0.15									
6	BM	0.059	0.048	0.09	0.33	0.05								
7	ln(Mcap)	0.044	0.046	0.26	0.10	0.40	0.27							
8	DYD	0.016	0.015	0.13	0.10	0.16	0.10	0.10						
9	Id. Vol.	0.057	0.027	0.14	0.07	0.32	0.06	0.31	0.15					
10	$RET_{\ 1}$	0.007	0.013	0.00	0.00	0.00	0.09	0.02	0.01	0.03				
11	$RET_{(\ 12;\ 2)}$	0.117	0.118	0.01	0.09	0.04	0.25	0.07	0.08	0.07	0.03			
12	$IOShr$	0.016	0.034	0.29	0.03	0.01	0.20	0.41	0.12	0.08	0.00	0.03		
13	$IOShr_HHI$	0.019	0.028	0.18	0.02	0.06	0.18	0.35	0.01	0.14	0.01	0.01	0.60	
14	TO	0.06	0.02	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	0.085***	0.085**	0.087***	0.090***
	33.5	2.41	37.24	40.37
$QIDRes_{q\ \ 1}$	0.064***	0.064***	0.078***	0.11***
	[19.16]	[3.01]	[3.70]	[5.15]
$QIDRes_{q\ \ 2}$	0.093***	0.093**	0.11***	0.14***
	[27.93]	[2.54]	[4.27]	[5.78]
Quarter FE	No	No	Yes	Yes
Stock FE	No	No	No	Yes
Clustered Errors	N/A	Quarter & Stock	Quarter & Stock	Quarter & Stock
Observations	75,017	75,017	75,017	74,792

A.3 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 (3) to fit parameters from the previous quarter, but it defines $QIDResInt$ as follows

$$QIDResInt_{jt}^{q} = \frac{QID_{jt}^{q} - \left(\widehat{a_j^{q-1}} + \widehat{b_j^{q-1}} \ln(PQSP)_{jt}^{q}\right)}{\widehat{a_j^{q-1}}} \tag{A.10}$$

wher $\widehat{a_j^{q-1}}$ is obtained from equation (3). This modification accounts from the potential cross-sectional variation in unconditional average undercutting. 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

$$QID_{jt}^{q} = \alpha_j^{q} + \beta_j^{q} \ln(PQSP)_{jt}^{q} + \gamma_j^{q} qvol_{jt}^{q} + v_{jt}^{q}, \tag{A.11}$$

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:

$$QIDResV_{jt}^{q} = \frac{QID_{jt}^{q} - \left(\widehat{\alpha_j^{q-1}} + \widehat{\beta_j^{q-1}} \ln(PQSP)_{jt}^{q} + \widehat{\gamma_j^{q-1}} qvol_{jt}^{q}\right)}{\widehat{a_j^{q-1}}}. \tag{A.12}$$

Figure A.2 shows that $QIDResSD$ and $QIDResV$ behave qualitatively very similarly to the basking $QIDRes$ around major information events.

Figure A.2. 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.