Dark Trading at the Midpoint:
Pricing Rules, Order Flow and Price Discovery

Robert P. Bartlett, III
University of California, Berkeley
School of Law

Justin McCrary
University of California, Berkeley
School of Law

February 12, 2015

JEL codes: G10, G15, G18, G23, G28, K22
Keywords: Tick size; high-frequency trading; dark pool, internalization; fragmentation; market quality

Dark Trading at the Midpoint:
Pricing Rules, Order Flow and Price Discovery

Abstract:

We exploit a regulatory discontinuity in the minimum pricing increment, or tick size, to explore the mechanism by which tick size regulations benefit "dark" trading venues at the expense of traditional stock exchanges and the associated effects on displayed liquidity and price discovery. Under Rule 612 of Regulation NMS, orders priced at or above $1.00 per share are required to be priced in penny increments, while all other orders can be quoted in subpennies. Using all trading data from 2011-2013, our regression discontinuity analysis of this $1.00 cut-off confirms that the penny tick size facilitates "queue-jumping" of exchanges by dark pools; however, we show that the most significant portion of such queue-jumping occurs by means of trades at the midpoint of the NBBO—a form of trading that is generally considered to have significant benefits for institutional investors. Further, while queue-jumping in dark venues is associated with a significant decrease in the quantity of displayed liquidity on public exchanges, we show that the largest effect is on liquidity providers utilizing high frequency trading (HFT) algorithms, thus revealing an important inverse relation between dark trading and HFT. Finally, notwithstanding the drop in displayed liquidity associated with dark trading, a tick size rule that encourages queue-jumping of exchanges' limit orders books is associated with an overall grater rate of actual trading. In light of recent concerns about HFT, these findings suggest that the deterioration of liquidity on public exchanges facilitated by wider tick sizes may be less detrimental for market quality than previously believed. They also emphasize the difficulty of stemming the flow of trades away from stock exchanges without prohibiting midpoint trading within non-exchange venues.

1. Introduction

Recent years have witnessed a dramatic transformation of U.S. equity trading from a market dominated by just a handful of conventional stock exchanges to one involving dozens of dispersed trading venues. Urged on by Regulation National Market System (Reg NMS) in 2005, a host of non-exchange "dark" venues now exist that compete successfully with exchanges for trades in U.S. listed equities. As of the end of 2014, these venues accounted for more than 30% of all consolidated trading volume, leaving approximately one dozen stock exchanges to compete fiercely for the remaining volume. The fate of the once storied New York Stock Exchange provides but one example of the disruptive forces at play. Accounting for nearly 83% of the consolidated volume of NYSE-listed stocks as recently as 2000, the "big board" accounted for just over 20% of such volume as of the end of 2014—its very survival now in the hands of the Intercontinental Exchange, a rival futures exchange that acquired the NYSE in 2012.

Within this environment, recent theoretical and empirical work has established a somewhat surprising link between the fragmentation of U.S. equity markets and the minimum price variation (MPV) for quoting equity securities (see, e.g., Kwan, Masulis, and McInish, 2014; Bartlett and McCrary, 2013; Buti, Rindi, and Werner, 2011). The reason arises from Rule 612 of Reg NMS which requires all orders submitted to exchanges and priced at or above $1.00 per share be priced in penny increments but which allows *trading* to occur in subpenny increments. In principle, a liquidity provider that faces a long queue of limit orders priced at the national best bid or offer (NBBO) on a public exchange can therefore turn to so-called "dark" trading venues to submit more aggressive, non-displayed orders to be executed in subpenny increments against incoming market orders. In this fashion, current MPV rules help contribute to the increasingly fragmented U.S. trading market, but at the potential cost of impairing price discovery. In particular, by facilitating queue jumping of exchanges' limit order books, existing MPV rules may discourage traders from providing publicly displayed liquidity used to infer market prices (Buti, Rindi, and Werner, 2011.) A pending proposal by the Securities and Exchange Commission (SEC) to increase the MPV to $0.05 only heightens these concerns given the potential such a change might have for increasing the incidence of queue-jumping in dark pools.

In this study, we investigate whether the prevailing MPV rule harms the incentive of liquidity providers to display trading interest by facilitating queue-jumping in dark pools. We offer three central empirical findings. First, to the extent the MPV rule facilitates queue-jumping in dark pools, it does so primarily by facilitating trades at the midpoint of the NBBO—a form of subpenny trading that is generally considered to have significant benefits for institutional investors. Second, queue-jumping in dark venues is associated with a significant decrease in the quantity of displayed liquidity; however, our empirical analysis indicates that the largest effect is likely to be on liquidity providers associated with high frequency trading (HFT). Finally, we demonstrate that notwithstanding the drop in displayed liquidity associated with queue-jumping, an MPV rule that encourages queue-jumping of exchanges' limit orders books is associated with an overall greater rate of actual trading. In combination, these findings suggest that the deterioration of liquidity on public exchanges facilitated by wider tick sizes may be less detrimental for overall market quality than previously believed.

In our empirical tests, we consider a trade as potentially queue-jumping exchanges' displayed order books because of the MPV rule where the trade both executes away from an exchange in a subpenny increment and the execution price rests within the prevailing NBBO.[1] While any trade meeting these conditions might reflect queue-jumping because of the MPV, we empirically measure the incidence of two distinct forms of subpenny queue-jumping that differ in the offsetting benefit they provide to investors through price improvement. In permitting subpenny trades in Reg NMS, the SEC justified the practice to enable broker-dealers to provide price improvement over the NBBO for incoming marketable orders.[2] As emphasized in Buti, Rindi, and Werner (2011) and Dick (2010), however, the fact that Reg NMS does not require a minimum amount of price improvement provides no assurance that a marketable order will receive a meaningful improvement over the NBBO in a dark venue to compensate for any adverse effect such trading might have for liquidity provision on public exchanges. Consistent with this concern, Delassus & Tyc (2011) document a sharp increase over time in the incidence of subpenny trades that offer only de minimis price improvement (e.g., \$0.0001 per share), suggesting subpenny trades can be motivated simply to "step ahead" of exchanges' limit order books.

In contrast to such trading (which, in keeping with industry practice, we refer to as "stepping ahead"), a separate class of subpenny trades often execute at the midpoint of the NBBO. This latter form of subpenny trading, which we refer to as "subpenny midpoint trading," typically arises from nondisplayed orders from investors with instructions to execute at the NBBO midpoint against marketable orders submitted to a trading venue. As modeled in Buti and Rindi (2012), such orders can be used by aggressive traders to undercut depth at the top of exchange limit order books given that these orders provide superior pricing to an incoming marketable order. However, the fact that these trades maximize price improvement for both sides of the transaction by splitting the spread between buyer and seller has historically diminished concerns that they might harm price discovery without any offsetting benefits. Indeed, the SEC adopted this position in devising its controversial "trade-at" rule included in the agency's pending pilot study to increase the MPV to \$0.05.[3] In light of the agency's concern about the effect of increasing the MPV on queue jumping and price discovery, the rule prohibits a venue not displaying the NBBO from filling an incoming order unless it can provide price improvement of at least \$0.05 per share. The rule specifically exempts, however, any trade executed at the midpoint of the NBBO, limiting its reach to subpenny stepping-ahead.

Notwithstanding widespread concern about the relation between MPV rules and stepping-ahead, the means by which a wider MPV facilitates queue-jumping remains an open question in light of the recent findings of Kwan, Masulis, and McInish (2014) (KMM). Using a proprietary dataset that classifies the trading venue for each reported trade in a dark facility, KMM examine the effect of the prevailing MPV rule on dark trading by exploiting an exception to Rule 612 for stocks priced under \$1.00 where subpenny quotations are permissible. They find that the effect of increasing the MPV on a dark venue's market share was localized entirely among trading venues classified as dark electronic communication networks

[1] Queue-jumping can also occur without subpenny pricing where a trade occurs on a non-exchange venue at the NBBO. Because such trades occur at the same (penny) price increment as the NBBO, the minimum pricing increment should have no effect on their incidence. As discussed below, such trades can be problematic for brokers in light of their best execution obligations.
[2] A trader seeking immediate execution will ordinarily obtain it by submitting a marketable order to be executed at the NBBO against resting limit orders.
[3] Spurred by Congress to explore whether decimalization of stock prices harmed the liquidity of small capitalization firms, the SEC proposed a tick size pilot in 2014 to increase the MPV from a penny to \$0.05 for a test group of securities. The agency's longstanding concern with queue-jumping and price discovery, however, led it to create a subset of securities that would also be subject to a trade-at requirement.

("dark ECNs"), which they describe as venues that maintain nondisplayed limit order books.[4] Examination of recently released Forms ATS for these types of venues, however, indicates that their subpenny trades are commonly priced at the midpoint of the NBBO, casting doubt on the hypothesis that the effect of the prevailing MPV rule on queue-jumping is necessarily a story about subpenny stepping-ahead.

To isolate the mechanism by which the prevailing MPV rule affects the incidence of dark trading and the supply of displayed liquidity, we turn to a regression discontinuity (RD) analysis of subpenny trading at the $1.00 cut-off established by Rule 612. We hypothesize that a primary channel through which the MPV rule affects the incidence of dark trading is through its effect on midpoint trading rather than through stepping-ahead. Specifically, we conjecture that for small orders investors generally prefer to buy (sell) securities through posting limit orders at the bid (ask) on exchanges' displayed limit order books rather than taking liquidity through placing marketable orders and, consequently, paying the spread.[5] As emphasized by KMM, however, where spreads are constrained by the penny MPV, investors seeking to place aggressive buy (sell) orders will be forced to join long queues at the national best bid (ask) given the price-time priority rules by which limit orders are filled on exchanges.[6] It is in this environment that dark ECNs can compete with exchanges by offering midpoint pricing that enables investors to post nondisplayed midpoint orders that execute against marketable orders in the venue. Although turning to dark ECNs raises execution risk, the significant price improvement offered by midpoint pricing combined with the nontrivial execution risk of posting orders to an exchange should result in smart-order routing protocols first checking dark ECNs for midpoint liquidity. In contrast, when quotes are no longer constrained by a penny MPV, the finer pricing increments will more accurately reflect heterogeneous pricing among investors on public exchanges, which both lowers spreads and shortens quote queues at any single price point, including the NBBO. Both effects should reduce the attractiveness of using nondisplayed midpoint orders relative to placing displayed orders on exchanges.

Using TAQ data from 2011-2013, we provide evidence that the current MPV rule has precisely this effect on the incidence of midpoint trading in dark venues. In particular, we show that the probability of midpoint trading in a dark pool at the $1.00 cut-off reveals an approximate 11% discontinuous increase as

[4] We interpret KKM's definition of "dark ECN" to include venues that match market orders to posted limit orders and that match market orders with other market orders. As such, we treat dark ECNs as equivalent to so-called "crossing networks" managed by broker-dealers. KMM's concept of a "dark ECN" should also be distinguished from a conventional ECN (such as Lavaflow or Bloomberg Tradebook). Conventional ECNs display to the public the best available quotes of their subscribers either directly (e.g., through Finra's Alternative Display Facility) or indirectly through incorporation into a formal stock exchanges' displayed liquidity. Through 2014, for instance, Lavaflow published its quotations as part of the quotation feed of the National Stock Exchange. Dark ECNs, in contrast, do not display to the public the price or depth of any limit orders submitted by their subscribers.

[5] We ignore for the moment the effect of exchange access fees and rebates, which can also induce investors to trade by means of posted limit orders on exchanges. We address these fees and rebates in Section 5. Additionally, we focus here on small orders to put to the side concerns about price impact. Where an investor trades a large order (e.g., a block trade), concern about revealing its trading interest before consummation of the full order may induce an investor to use a dark pool for reasons having nothing to do with the size of the MPV. Because large orders by institutional investors are commonly broken apart into smaller "child" orders for execution, small orders can originate from both institutional and retail investors.

[6] This effect is also driven by the ban on locked markets contained in Rule 610(d) of Reg NMS. For instance, if the NBBO is at 10.01 x 10.02, and a trader attempts to price improve the national best bid by submitting a buy order at the next available penny increment of 10.02, the order would be submitted at the price of the national best ask, thereby locking the market. Because Rule 611(d) prohibits market centers from accepting orders that lock the NBBO, the venue would therefore reject the order (if the best ask sits on an away market) or convert it to a marketable order for execution at the best ask of 10.02 (if the best ask sits on the receiving venue). A trader wanting to post to an exchange an aggressively priced bid would therefore be forced to join all other buy orders resting at the national best bid of 10.01.

the MPV moves from subpenny increments to penny increments for those securities most likely to be constrained by the penny MPV. In contrast, the probability of a trade executing with de minimis price improvement shows a discontinuous 2% *decline* as prices moved above the $1.00 cut-off. We attribute this latter result to the fact that subpenny stepping-ahead is widely-believed to be conducted by broker-dealer internalizers (Dick, 2010). These venues typically fill retail marketable orders obtained directly from affiliated retail brokerage houses or through negotiated payment-for-order flow agreements, thereby ensuring a supply of inbound marketable orders regardless of the MPV. Although a broker's best execution obligation may incentive an internalizer to provide de minimus price improvement for such trades, this obligation is constant across penny and subpenny MPV regimes. However, because market orders for stocks trading for less than $1.00 per share are filled less by midpoint trades, a greater portion of such orders flow to internalizers under the $1.00 cut-off, resulting in stepping ahead being more (rather than less) prominent in the subpenny trading environment. These results are consistent with KMM's finding that the MPV affects the market share of dark ECNs but not necessarily broker-dealer internalizers. In contrast to KMM, however, these results suggest that the mechanism by which a wider MPV facilitates queue-jumping in dark venues is not through stepping-ahead but through investors' pursuit of midpoint trading.

Turning to the effect of queue-jumping on liquidity provision, we similarly conduct RD analysis of the $1.00 cutoff to examine how the rise in midpoint trading above $1.00 affects the provision of displayed liquidity across public exchanges. Consistent with concerns that queue-jumping affects the incentive to provide displayed liquidity, our analysis of quote activity across exchanges reveals a dramatic drop in quote updating at the $1.00 cut-off. This finding suggests a significant decline in quote competition caused by constraining the MPV to a penny and its associated increase in queue-jumping in dark pools.

While this evidence is consistent with queue-jumping harming the incentive to provide displayed liquidity, analysis of intra-millisecond quote activity indicates that the effect is driven largely by liquidity providers engaged in HFT. Given recent concerns with liquidity provision offered by HFT firms (Egginton, Van Ness and Van Ness (2012); Zhang (2010); Cartera and Penalva (2011)), assessing the overall welfare effects associated with the loss of liquidity provision due to a wider MPV is therefore likely to be more complicated than simply evaluating the quantity of liquidity provision. While such an undertaking is beyond the scope of this paper, we nevertheless conduct an RD analysis of the overall rate of trading at the $1.00 cut-off and find a discontinuous increase in trading above the cut-off notwithstanding the diminished quote competition in the penny-priced trading environment. This counterintuitive finding underscores the possibility that even if a wider MPV diminishes liquidity provision because of greater queue-jumping in dark pools, overall market quality may nevertheless reflect a more liquid trading environment.

Our study is most closely related to theoretical and empirical literature examining the effect of the prevailing MPV rule on competition between exchanges and dark pools. Buti, Rindi and Werner (2011) model the competition between an exchange limit order book and a dark pool, demonstrating that for a liquid stock with high limit order book depth, traders will choose to trade in a dark pool where they can undercut the existing price in the exchange limit order book. Buti, Rindi, Wen, and Werner (2011) model the intermarket competition between a public limit order book and an internalization pool dark venue, which also offers a smaller tick size. Market orders sent to the public limit order book are intercepted by the internalization pool in which better prices are available, resulting in a decline in liquidity demanded from the limit order book. KMM provide empirical support that a wider MPV causes a reduction in the market share of the liquidity provided by exchanges and an increase in that of dark venues. All of these papers speculate that subpenny stepping-ahead is the mechanism through which the MPV rule gives dark venues a competitive advantage over exchanges. Our empirical results, however, show that the primary mechanism by which this occurs is through the more innocuous form of midpoint trading.

Our study also speaks to a burgeoning literature examining the consequence of market fragmentation on overall market quality. The dispersion of trading away from public exchanges to an increasing number of non-exchange venues has prompted considerable concern that this development might have adverse effects on price discovery and trading costs. Most of these concerns are rooted in the potential harm non-exchange trading poses for liquidity on exchanges' public limit order books which determine transaction prices for both displayed and non-displayed venues. For instance, a long-standing concern with broker-dealer internalization is the potential for dealers to engage in "cream skimming" in which broker-dealers internalize uninformed orders (such as those submitted by retail traders) causing exchanges to receive a disproportionate share of informed trades (Harris, 1995; Easley, Keifer, and O'Hara, 1996; Bessembinder and Kaufman, 1997). To the extent this occurs, internalization should result in wider spreads and reduced depth in the public "lit" market to compensate for the increased percentage of informed traders in the public order flow (Chakravarty and Sarkar, 2002). Theoretical models have also extended this concern to the emergence of dark venues designed to absorb institutional order flow, although their predictions for market quality are highly dependent on parameter assumptions (See Hendershott and Mehndelson, 2000; Buti Rindi, and Werner, 2011; Ye, 2011; Zhu, 2014). Empirical studies investigating this issue have similarly drawn conflicting conclusions depending on the data set and empirical methodology utilized (see O'Hara and Ye, 2011; Weaver, 2011; Gresse, 2012; Nimalendran and Ray, 2014).

We provide compelling evidence that the increase in dark trading afforded by a wider ticker size does indeed diminish the competition of liquidity providers on public exchanges. However, we also demonstrate that the nature of this diminished liquidity reflects a diminished amount of high-speed, algorithmic liquidity provision. By highlighting how markets can oscillate between dark trading and HFT, we therefore underscore the need for future research to disentangle the ways in which both forms of trading may have simultaneous and offsetting effects on overall market quality.

Finally, our findings have immediate policy implications for the ongoing debate over the optimal tick size for emerging growth companies. The goal of the SEC's pending pilot tick size study is to encourage liquidity and market coverage for emerging growth companies by enhancing the profitability of market-making. Notwithstanding this goal, our findings suggest that some of the primary beneficiaries of a wider MPV are likely to be dark ECNs who can be expected to accumulate a greater share of market trading but who may or may not choose to cross-subsidize sell-side research. Likewise with respect to the proposed trade-at rule, our findings highlight the likelihood that off-exchange trading will continue to persist during the trade-at pilot given the rule's exemption for midpoint trades and their importance in facilitating off-exchange trading.

2. Institutional Details

In this section, we describe how "dark" and "lit" trading venues compete for different forms of trading interest. We also describe how the MPV rule applies differently to these various forms of trading interest, which results in the MPV rule having different effects on dark and lit trading venues.

2.1. Trading Venues and Order Types

Historically, a central objective of U.S. trading venues has been to facilitate the interaction of two forms of trading interest, often referred to as "passive" and "active" liquidity. Generally taking the form of a dealer or specialist quote or a trader's limit order, passive liquidity represents a standing commitment to buy or sell a security at a specified price. Like an option, this commitment lasts until cancelled or accepted by a contra-side, "active" liquidity trader seeking immediate execution by means of an executable market order.

For exchanges and electronic communication networks (ECNs), a variety of factors have led these venues to focus on competition among passive liquidity providers as the fundamental building block for attracting marketable order flow. Central among these factors has been a broker's duty of best execution which has long required brokers in possession of a customer's market order to obtain the best price reasonably available for it (Macey and O'Hara, 1997). With the implementation of the Intermarket Trading System Plan in the 1980s, an exchange or ECN could potentially draw marketable order flow to the venue by attracting passive liquidity providers to post displayed orders that compete for price priority in the venue's limit order book, thus inducing brokers to route market orders to the venue. This focus on attracting passive liquidity was further encouraged by the SEC's Order Handling Rules in 1997 and the Order Protection Rule in Rule 611 of Reg NMS.[7] In combination, these rules enabled a customer submitting a limit order to establish a trading venue's best offer or best bid, while inducing a broker or trading venue holding a market order to route the order to the venue having the best bid or offer across all exchanges (i.e, to the venue holding the NBBO).

While these rules protect liquidity providers who establish the NBBO, the SEC's desire to induce competition among passive liquidity providers has often been tempered by a somewhat inconsistent desire to provide superior pricing to active liquidity traders. This latter policy has enabled non-exchange venues to attract active liquidity without necessarily offering passive liquidity. For instance, the SEC has long endorsed the practice of broker-dealer internalization, whereby broker-dealers fill incoming market orders from retail investors either as an agent matching their customers' buy and sell orders or as a principal taking the other side of those orders.[8] Although the Order Protection Rule requires a broker-dealer to

[7] Among other things, the Order Handling Rules (17 CFR §242.604) require market makers and specialists to display publicly the limit orders they receive from customers when such orders are better than the market maker or specialist's quote. In effect, the rule ensures that the general public can compete directly with market makers in the quote-setting process. The SEC's Order Protection Rule (17 CFR §242.611) requires (such to several exceptions) trading centers to establish and enforce procedures designed to prevent "trade-throughs"—trade executions at prices inferior to the best-priced quotes displayed by automated trading centers.

[8] To be sure, internalization is potentially problematic for a broker-dealer seeking to internalize an order if that broker also holds a customer limit order on the same side of the market for the same security. Under Finra Rule 5320 (the "Manning Rule"), a broker-dealer holding such a limit order "is prohibited from trading that security on the same side of the market for its own

execute an incoming market order at a price that is no worse than the NBBO, the fact that the national best bid (NBB) should always be lower than the national best offer (NBO) leaves open the possibility that a broker could execute an incoming market order to buy (sell) a security at a price that is better than what the order would receive if routed to the NBO (NBB). For instance, if the market for a security stands at 10.00 x 10.05, a broker receiving a customer's market order to buy the security could route the order to the venue offering to sell at 10.05, but the broker could also comply with the Order Protection Rule by simply selling the customer the same stock at a price that is 10.05 or lower (e.g., 10.04). Indeed, a broker choosing to internalize such an order will often provide such price improvement over the NBBO to comply with their best execution obligations, providing a common justification for the practice.[9]

Likewise for larger institutional orders, the SEC has similarly permitted non-exchange trading venues to emerge that allow for the direct interaction of active liquidity to facilitate trading within the NBBO. Generally done on an agency basis, this form of trading has roots in the "upstairs" market of the NYSE. In contrast to the continuous auction run on the "downstairs" floor of the exchange, brokers working in the upstairs market facilitated large-block trades by locating counterparties to the transaction with prices determined through negotiation (Madhavan & Cheng, 1997). Today, this type of trading is most commonly done in any of the dozens of dark pools that operate as registered Alternative Trading Systems (ATS). While these venues differ in how they match trading interest, their business models generally rely on the ability to match marketable orders from institutional investors against one another with pricing determined by reference to the NBBO. For instance, a common order discussed in these venues' Forms ATS involves traders submitting nondispayed orders to sell or buy a security at the midpoint of the NBBO.[10] In the event of an incoming contra-side market order, such an order will be executed at the NBBO midpoint, allowing both parties to avoid paying any spread.[11] According to Tabb Group (2015), prominent dark pools such as Barclays DirectEx, IEX, and BIDs report more than seventy percent of their

---

account at a price that would satisfy the customer order." To avoid this restriction, brokers often sell market orders (but not limit orders) to dedicated broker-dealer internalizers under payment-for-order flow arrangements.

[9] As discussed in Ferrell (2001), the pressure to provide price improvement over the NBBO arose in large part due to the Third Circuit's decision in Newton v. Merrill, Lynch, Pierce, Fenner& Smith, Inc., 135 F.3d 266 (3d Cir. 1998), where the Third Circuit found that a broker-dealer that automatically executed customer trades at the NBBO may not be in compliance with its best execution obligations. Additionally, the manner in which Reg. NMS discussed the desirability of brokers' providing price improvement for their customers has also created a perception within the industry that best execution may require a broker to seek out opportunities for customer price improvement. In a comment letter to the SEC outlining how internalizers can often be subject to significant market risk when trading with their customers, TD Ameritrade (2010) articulated this perception: "One could certainly suggest that the [market-maker] simply avoid the price improvement opportunity and that the market maker or broker should have simply sent the order to fill at the NBBO. In such case, however, the broker would run the risk of being accused of violating its best execution obligation, as Regulation NMS elevated price improvement above all else." Finally, the incentive for offering price improvement over the NBBO is also encouraged by Rule 605 of Reg. NMS, which requires that broker-dealers publicly disclose their rate of price improvement over the NBBO as a core measure of execution quality.

[10] For instance, Credit Suisse's Crossfinder, the largest ATS by trading volume, notes in its Form ATS that "[p]articipants have the option on Orders to specify, relative to the National Best Bid or Offer ('NBBO'), a peg to the midpoint, a peg to the bid, a peg to the offer, or in penny increments from the bid or offer, and a minimum quantity." UBS, which runs a similarly large ATS, notes in its Form ATS that eligible orders can include limit orders, market orders, and orders that are pegged "to the near, midpoint, or far side of the NBBO." Notably, nine pages of the UBS Form ATS are dedicated to providing hypothetical "crossing scenarios," almost all of which involve midpoint pegged orders.

[11] In recent years exchanges have also begun to permit orders pegged to the NBBO midpoint. Such orders are part of a broader category of nondisplayed orders that exchanges have long accepted. As discussed in Buti and Rindi (2011), a trader can choose at the time of order submission whether an order will be non-displayed, partially displayed (often referred to as a "reserve order"), or fully displayed to the public and included in a venue's quoted depth. Hidden and the undisclosed portion of reserve orders execute against incoming marketable orders only after all displayed orders priced at or better than the undisclosed orders have been filled. According to the SEC (2013), the volume of stock trades that result from hidden liquidity on exchanges is typically between 11% and 14% of all volume.

trades are done at the NBBO midpoint. While sending an order to such a venue raises obvious execution risks compared to the certainty of accessing an exchanges' displayed liquidity, the possibility of this form of price improvement provides a potentially offsetting benefit for traders seeking immediate execution.

A cursory analysis of the Rule 605 execution reports submitted by exchanges, broker-dealer internalizers, and dark pools highlights the manner in which different trading venues focus on these different order types . Table 1, for instance, summarizes the stark difference in order types received in October 2014 for a prominent exchange (Nasdaq) compared to a prominent dark pool (Credit Suisse Cross Finder) and a large broker-dealer internalizer (G1 Execution Services).[12]

[Insert Table 1]

2.2. The MPV Rule and Order Types

In light of these divergent business models, the MPV rule contained in Rule 612 has a number of implications for how exchanges and non-exchange venues compete with one another for order flow. Rule 612 applies broadly across any venue, reaching any "national securities, national securities association, alternative trading system, vendor, or broker dealer," thus applying to exchanges and non-exchange venues alike. The rule is also sufficiently broad to capture displayed and non-displayed orders as it prohibits the "display, rank or accept[ance]" of "a bid or offer, an order, or an indication of interest in any NMS stock priced in an increment smaller than $.01" if such trading interest is priced at $1.00 per share or more. Any venue posting passive liquidity (whether limit orders or quotes) must accordingly abide by the rule.[13]

Rule 612 does, however, disproportionately affect exchanges and conventional ECNs in light of their disproportionate emphasis on using passive liquidity (e.g., dealer quotes and customer limit orders) to compete for order flow. Specifically, customers and dealers attempting to set the NBBO on an exchange or conventional ECN must ensure that all quotes and orders are made in penny increments, except for orders priced less than $1.00 per share which can be made in subpenny increments. While this rule also applies to any limit orders posted to a dark pool, the fact dark pools seek to facilitate the interaction of marketable orders within the NBBO has allowed these venues to avoid Rule 612's restriction on order pricing by using order types that are technically unpriced but still tied to the price of the prevailing NBBO. For instance, to facilitate midpoint trading when the midpoint of the NBBO is a fraction of a penny, the SEC has explicitly endorsed the use of "Midpoint Peg Orders."[14] Even if the NBBO spread is a penny, these order types allow a trader to submit an immediately executable order that will execute in a subpenny increment against any incoming marketable orders. For stocks trading with narrow spreads, the

[12] G1 Execution Services was formally the market-making division of E*Trade until its spin-off in February 2014. E*Trade's Rule 606 Filing for the Fourth Quarter of 2014 indicated that 70% of all E*trade market orders were routed to G1 Execution Services, making it the primary recipient of E*Trade's significant volume of retail market orders. As noted above, CrossFinder represents the largest dark pool in terms of its volume of trading.

[13] Within the academic literature, application of Rule 612 across trading venues has often been a source of confusion. For instance, papers modeling the role of tick size on market competition (see, e.g., Buti, S.,Rindi,B.,Wen,Y.,Werner,I.,2011) have often assumed Rule 612 does not apply to non-exchange venues. As noted in the text, however, the rule expressly applies to both exchange and non-exchange venues. Equally important, recent SEC enforcement actions (see infra note 15) have made clear the SEC's willingness to enforce the rule against both exchange and non-exchange venues alike.

[14] *See* Exchange Act Release No. 51808, at 231 ("Rule 612 will not prohibit a sub-penny execution resulting from a midpoint or volume-weighted algorithm or from price improvement, so long as the execution did not result from an impermissible sub-penny order or quotation.").

ability to place an order that is effectively priced in subpennies at the NBBO midpoint might therefore offset the greater execution uncertainty of trading in a dark venue.[15]

As with dark pools, Rule 612 is also less constraining for broker-dealer internalizers in light of the manner in which they interact with incoming market orders. Here, the reason arises from the fact that trading generally occurs when an internalizer chooses to execute against an incoming marketable order using its proprietary capital rather than by means of crossing pegged orders. Because the internalizer does not display or rank orders or quotes, the SEC permits internalizers to fill incoming buy and sell orders at prices that improve the NBBO in subpenny increments.[16]

2.3. Order Routing of Marketable Orders

The differentiated application of Rule 612 across different trading venues can ultimately have significant implications on order flow given the widespread use of smarter order routers to manage activity liquidity. This especially true for market orders from retail brokerage firms that are commonly sold to broker-dealer internalizers in payment for order flow agreements. These arrangements assure internalizers a constant supply of market orders, providing such venues with an opportunity to trade ahead of exchanges at a price that is at or better than the NBBO. Moreover, as documented in Bright Trading (2010), even where an internalizer chooses not to fill the order, the possibility that price-improving liquidity exists within a dark pool leads most smart order routers to check these venues for a trade within the NBBO before routing an order to the exchange holding the best displayed price.[17] In combination, these order routing practices of marketable orders and the existence of subpenny liquidity in dark pools allows for the possibility that the MPV rule can facilitate queue-jumping of exchanges' displayed limit order books.

3. Data and Empirical Design

3.1 Sample Construction

To analyze the effect of the MPV rule on queue-jumping and displayed liquidity, we used the consolidated quote and trade data contained in the NYSE Euronext's daily Trade and Quote (TAQ) database. The TAQ database provides intraday trade and quote data time-stamped to the millisecond for

[15] To be sure, the legal distinction between orders "priced" at the NBBO midpoint and orders "pegged" at the NBBO midpoint has occasionally been lost on trading venues. A recent disciplinary proceeding against the dark pool managed by UBS, for example, arose in large part because of a "technical problem" by which immediate or cancel (IOC) orders priced at the midpoint of the NBBO were submitted by the UBS smarter order router rather than IOC orders "pegged" to the midpoint. As summarized by the SEC, "[w]hen seeking to place an order in UBS ATS at the NBBO midpoint, UBS's smart order router would send an immediate-or-cancel limit order that was explicitly denominated at the price the router had calculated to be the midpoint of the NBBO, rather than sending an order with a price that was *pegged* to the midpoint of the NBBO." While functionally equivalent orders, the fact that the UBS orders were technically priced in subpenny increments violated Rule 612 given that the rule "does not permit an ATS to accept and rank an order that is explicitly denominated in a sub-penny price (even if that sub-penny price is equal to the midpoint of the NBBO)."

[16] As discussed below, broker-dealer internalizers may also be structured as a hybrid whereby some retail orders are executed using a broker-dealer's proprietary capital and some are routed to an affiliated dark pool where they interact with orders provided by third-party subscribers who might include institutional investors and high-frequency trading firms. The large retail market making business of UBS , for instance, operates in this fashion.

[17] Interactive Brokers (IB), for instance, notes in its Rule 605 report for September 2014 that its smart order routing system "continually scans competing market centers and automatically seeks to route orders to the best market, taking into account factors such as quote size, quote price, exchange or ATS transaction fees or rebates and the availability of price improvement…." The report also notes that IB maintains its own dark pool to which it routed 30% of non-directed market orders it received for NYSE and Nasdaq-listed securities.

all transactions reported to the Securities Industry Automation Corporation (SIAC). The TAQ data are comprised of two files, one corresponding to trades and one corresponding to quotes.

Pursuant to the Consolidated Tape Association (CTA) Plan and the Unlisted Trading Privileges (UTP) Plan, all U.S. exchanges and FINRA are obligated to collect and report to SIAC for dissemination on the Consolidated Tape last sale data in securities listed on the NYSE, Nasdaq, the Amex, and all regional exchanges. These reported transactions are the recorded in the TAQ daily Consolidated Trade File. Although the Consolidated Tape does not directly record the identity of non-exchange participants reporting a trade, the SEC has required since March 2007 that all off-exchange transactions be reported to a formal FINRA-managed Trade Reporting Facility (TRF) established at certain stock exchanges which report directly to the SIAC. As described by O'Hara and Ye (2011), this requirement means that off-exchange trades made through a broker-dealer internalizer or in dark pool (both of which were historically reported to an exchange and then consolidated with the exchanges' own trades when reported to the Consolidated Tape) are now effectively segregated and reported to SIAC as having been executed at a FINRA TRF.

In addition to the Consolidated Trade file, TAQ also includes a daily Consolidated Quote File which records historical quotation data reported to SIAC. As with their trade reporting obligations, all exchanges and FINRA are required to report to SIAC for publication in the Consolidated Quotation System (CQS) any change in the best bid and best offer (including aggregate quotation sizes) currently available on each trading venue. The CQS thus provides for any moment of the trading day a snapshot of the total, consolidated trading interest at the best bid and offer ("Consolidated BBO") available at each exchange and through a FINRA member. We use TAQ's Consolidated Quote File to calculate the NBBO over the course of each trading day for every security in our sample. Some of our analyses rely exclusively on the Consolidated Trade File, others rely exclusively on the Consolidated Quote File, and yet others require that we interleave the two files (i.e., align them in chronological order for the same security). For some of our analyses, it is additionally necessary to classify trades as having been buy- or sell-side initiated. We follow much of the literature and use the Lee and Ready (1991) algorithm. A challenge with the analyses involving interleaving, including analyses using the Lee-Ready algorithm, is that the timestamps in the two files are not perfectly synchronized, as is widely recognized in the literature. Since our analyses are from recent years, we follow the recent literature in assuming that trades occur contemporaneously with quotes (e.g., Bessembinder and Venkataraman 2010).

Because we are interested in how the change in MPV rule at $1.00 affects queue-jumping, we limit our sample to trades and quotes that are priced at less than $4.50 per share during the three year period spanning 2011-2013.[18] We also filter the TAQ data to exclude quotes that are marked as cancelled or corrected (CQS cancel code of "B" or "C"), and we filter to ensure that all quotes and trades occur during the trading day after the opening cross and before the closing auction (i.e., between 9:45:00.000 and

[18] To efficiently access those records, and in particular to take advantage of the index structure of the TAQ files as stored on Wharton Research Data Services (WRDS) where we performed the bulk of our computations, we accomplish this subsetting in a multi-step procedure. We first identify the subset of CRSP universe securities that had a closing price of below $4.50 at some point during 2011-2013, using the CRSP dsf file. We then used the CRSP dsenames file to identify the corresponding ticker symbols on a day-by-day basis. We then constructed time series plots of each identified security over time, and verified that the CRSP data on closing price was in tight agreement with TAQ end-of-day prices (or bid-ask midpoint when trade prices were missing, consistent with CRSP measurement protocols). We then pulled extracts of all trades and quotes for those securities from the full TAQ data, taking advantage of the index structure using key merging. Finally, we restrict our attention to prices and NBBO values that are below $4.50.

15:35:00.000). Finally, since the only identifier for securities in the TAQ data is ticker symbol, which does not uniquely identify securities due to the retirement and recycling of ticker symbols, and also because TAQ contains some ticker symbols that do not correspond to actual securities, we further limit our sample to those ticker symbols that could be matched to a CRSP record on a day-by-day basis. With these restrictions imposed on the TAQ data, the full sample contains 676,680,160 trades and 8,123,048,164 updates to venues' Consolidated BBO. Given that many of these securities trade at spreads wider than the penny MPV, we focus our analysis on a subsample of securities where the penny MPV is most likely to be a binding constraint. We construct this subsample by partitioning the overall sample into three liquidity bins based on average trading volume.[19] As shown in Table 2, average quoted (midpoint) spreads for the highest liquidity group of securities were approximately \$0.013 (\$0.008), indicating that the penny MPV binds traders seeking to display liquidity for these securities.[20] In contrast, the higher average quoted (midpoint) spreads of \$0.28 (\$0.02) for medium and \$0.344 (\$0.055) for low liquidity securities indicate that traders can generally improve one side of the displayed market for these securities using penny priced orders.[21] As such, the group of 683 high liquidity securities constitutes our core sample in our empirical analyses.

[Insert Table 2]

3.2 Regression Discontinuity: Overview

To assess how the MPV rule affects queue-jumping and liquidity provision on exchanges, we follow KMM in using an RD framework to leverage the change in MPV for orders priced at or above \$1.00 per share. As noted by Hahn, Todd and van der Klaauw (2001), the "regression discontinuity data design is a quasi-experimental data design with the defining characteristic that the probability of receiving treatment changes discontinuously as a function of one or more underlying variables" (p. 1). The MPV rule fits nicely within this data design on account of the sharp regulatory distinction involving the MPV created by Rule 612 of Reg. NMS. Under this rule, the MPV regulation that applies to any given trading order varies sharply: an order is allowed be posted in below penny increments if and only if the order is less than \$1.

Using this discontinuous treatment of MPV regime, we develop the following baseline model to evaluate the effect of changing the MPV on the trading environment by measuring directly the conditional expectations of market measures given two-decimal prices, or $E[Market\ Measure_i |\ Price_i]$, where *Market Measure*$_i$ is an outcome for security-time *i* and *Price*$_i$ is the running variable, or price truncated to two decimals (e.g., \$0.98, \$.0.99, \$1.00, \$1.01, etc.). We adapt this estimation strategy and variable definitions to the constraints of the data. For example, when focusing on data from CRSP, our notion of

[19] Partitions for the three liquidity bins were determined based on identifying a level of average daily trading volume within the sample that would divide the sample into thirds. Issuers for which averaged daily trading volume was less than 50,000 shares were assigned to the Low liquidity category; issuers for which averaged daily trading volume as greater than 50,000 shares but less than 250,000 shares were assigned to the Medium liquidity category; and issuers for which average daily trading volume exceeded 250,000 were assigned to the High liquidity category.

[20] That is, a trader seeking to submit a competitive buy (sell) order will be required to submit the order at the national best bid (ask) as any price that is more aggressive than the national best bid will cross the market, causing the order to be rejected or converted into a marketable buy (sell) order.

[21] The large average quoted spreads for medium and low liquidity securities reflect the influence of outlier quotations arising from the absence of meaningful quotation activity that occurs periodically in the TAQ data. As such, midpoint spreads (quoted spread / quote midpoint) may provide a better assessment of the relative spreads of these two liquidity categories during normal trading periods.

"security-time" is a given day for one of our sample securities. As another example, in analyses involving the Consolidated Quote File, security-time reflects a security's NBB as of the end of millisecond.[22] We specify below in our empirical results the particular form taken for security-time. As emphasized by Lee (2008), the core underlying assumption of the RD design is "smoothness," or continuity of potential confounders given the running variable, and the plausibility of this assumption can be evaluated by examining the continuity of pre-determined characteristics given the running variable. Consequently, our analyses pertain both to assessing the magnitude of the discontinuity in an outcome given price and to assessing the magnitude of any discontinuity in a pre-determined characteristic given price.

3.3 Regression Discontinuity: Delisting and Smoothness Concerns

While the sharp cut-off in the MPV at $1.00 per share makes Rule 612 a potential candidate for an RD research design, two potential issues relating to the trading environment around the $1.00 cut-off could potentially produce biased estimates. First, all major U.S. exchanges impose a $1.00 minimum bid price requirement for continued listing. On Nasdaq and the NYSE, for instance, a firm that trades for thirty consecutive trading days with a closing bid below $1.00 risks triggering a review of its continued listing eligibility. While the existence of this rule potentially increases the probability of a security's delisting to the extent it trades at less than $1.00 per share, the rule can compromise the use of our RD framework only if this risk changes discontinuously at the cut-off. Two factors suggest this is not the case. First, an exchange's decision of whether to initiate a delisting proceeding is discretionary. Second, even if a proceeding is commenced, firms are entitled to a lengthy compliance period of 180 days to increase their stock price (e.g., through a reverse stock split).

Notwithstanding these institutional rules, given the enhanced delisting risk for stocks trading below $1.00, we examined empirically whether this risk changed discontinuously at the $1.00 cut-off. Using CRSP's daily closing prices of the sample securities, we calculated the probability that a security would still be trading 180 days in the future given the closing price for a given trading day. Figure 1 plots this probability as a function of a security's two-decimal closing price. This figure contains a number of distinctive features that are specific to the RD context and deserve explanation. First, the probabilities for each two-decimal closing price are presented as open circles of different colors. Dollars are displayed in red, quarters are displayed in blue, nickels and dimes are displayed in light gray, and all other price points are displayed in dark gray. This color scheme relates to the potential distinctiveness of these price points, as we discuss below. Second, a black vertical line is superimposed at the actual cutoff of $1.00, and green vertical lines are superimposed at "placebo" cutoffs of $2.00, $3.00, and $4.00. These vertical lines can be used to assess visually whether the evidence of a discontinuity at $1.00 stands in contrast to the pattern around other dollar points where no policy changes discontinuously.

We present our formal RD estimates of the discontinuity in the probability of still trading 180 days later in Table 3. The first column presents a basic local linear regression estimate of the discontinuous change in this probability at the $1.00 cut-off, indicating a smooth, continuous decrease in this measure as the price declines below $1.00 per share. We use the triangle kernel since that is known to be boundary optimal (McCrary 2008) and a bandwidth of twenty-five cents. Column 2 presents an alternative approach that eliminates dollar, quarter, dime, and nickel price increments to address price clustering (described below). The models in Columns 3 and 4 are the same as those in Column 1, but assess

[22] We use the NBB in our empirical analyses for ease of exposition. Our results, however, are robust to using the NBO.

robustness to alternative bandwidths.[23] For all four models, standard errors are reported in parentheses below point estimates.[24] All four specifications confirm that the delisting risk, while heightened for firms trading below $1.00 per share, is not discontinuously so.

[Insert Table 3]

A second potential challenge for designing an RD analysis based on Rule 612 arises from the phenomenon of price clustering of trades at increments of five and ten cents (nickels and dimes) described in Ikenberry & Weston (2008). As noted by Barreca, Lindo, and Waddell (2011), price clustering or "heaping" at particular price points has the potential to undermine the smoothness assumptions that are at the heart of the RD approach. Our visual display that we described above is designed to facilitate "pulling out" potential heap points from the other averages, and in some of our figures, the phenomenon of heaping will be readily apparent. Heaping in the running variable means that the running variable density function is not continuous. One of us has written on how a discontinuous density function may indicate a violation of smoothness. On the other hand, McCrary (2008) also emphasizes that a continuous density function is not necessary for the RD approach to be valid and gives an example of an application in which the running variable density function is discontinuous and yet the RD approach is valid. We argue here that price heaping does not invalidate the use of the RD design in financial markets and that smoothness is likely to be satisfied.

Our first set of arguments are theoretical and rooted in the efficient markets hypothesis. If smoothness were not correct—that is, if a price to the right of $1 were discontinuously predictive of the underlying "latent" value of a security—then a profitable trading strategy exists that would take advantage of that discontinuity. A "no arbitrage" assumption thus points to the validity of the RD approach in this context. Moreover, since a security may trade below $1 in one millisecond and above $1 in another millisecond, any invalidity of the RD approach here must point to within-day mechanisms that are consistent with arbitrage possibilities.

Our second set of arguments are empirical. There are a variety of pre-determined characteristics of securities that we can measure, and we can show that these pre-determined characteristics are in fact smooth functions of price. This is precisely the form of testing recommended by Lee (2008), Imbens and Lemieux (2008), and Lee and Lemieux (2010).

In Figure 2, we plot as a function of a security's closing price the average implied volatility for all outstanding call option contracts on that security for that trading day.[25] For purposes of this analysis,

[23] Column 3 uses a narrow bandwidth of twenty cents, and Column 4 uses the data-driven bandwidth of Imbens and Kalyanaraman (2012). The tension between the models in Columns 1 and 3 is the same as the familiar tension between bias and variance: a narrower bandwidth is closer to the regression discontinuity ideal of "in the limit" and potentially is less biased, yet a narrower bandwidth yields are more noisy estimate. A further wrinkle is that the accuracy of standard errors becomes more suspect with bandwidths that are too small. There is also tension between the models in Columns 1 and 4. In these data, twenty-five cents is a generally robust choice of bandwidth and has the important result of being comparable across outcomes and transparent. However, this transparency is not without costs. We examine curves of a variety of shapes and variances, and for some of these a bandwidth of twenty-five cents is probably too wide and for others it is probably too narrow. Column 4 seeks to replace transparency with a reliable data-driven method. While this method generally performs well, it also occasionally yields bandwidth choices and point estimates which human judgment squarely rejects. Overall, we have a mild preference for the model in Column 1, but we think it is important to consider what each of the four models tells us about the data.

[24] The standard errors we use in this context are the so-called "HC3" standard errors that adjust for heteroskedasticity and exhibit better small-sample performance than White standard errors. See MacKinnon and White (1985) and Hausman and Palmer (2011).

information concerning an option contract's strike price and expiration date, as well as the closing stock price, was obtained from OptionMetrics. For simplicity, we calculated implied volatility using Black-Scholes with a risk free rate of 0.5%.[26] The plot reveals a smooth, downward slopping curve, reflecting a decline in implied volatility as stock prices increase. Table 3 provides point estimates using the same four specifications used in examining delisting. None indicate evidence of any discontinuity at the cut-off.

Figure 3 presents a second smoothness test, this time conducted using cumulative monthly and quarterly returns based on data from CRSP. For each security on trading day t, we compute monthly (quarterly) returns as the conditional expectation, given the security's closing price on day t, of the cumulative returns over the past 20 (60) trading days, respectively. Cumulative returns are calculated using CRSP daily closing prices as $-1 + \exp\left(\sum_{t=1}^{T} \ln\left(1 + R_{r-t}\right)\right)$, where T is either 20 or 60 and R_r is the daily return. Panel A plots monthly returns as a function of the daily closing price truncated to two decimal places, and Panel B plots quarterly returns. Neither shows evidence of any discontinuity at the \$1.00 cut-off. Point estimates are provided in Table 4 and confirm this conclusion.

4. Empirical Results

In this section we first re-examine how the penny tick size facilitates queue-jumping of exchanges' displayed liquidity, and we demonstrate the importance of subpenny midpoint trading for this practice. We then document the adverse effect of queue-jumping on price competition among liquidity providers, particularly HFT firms. Finally, given these findings, we examine the welfare effects of a tick size rule that facilitates queue-jumping relative to one that facilitates HFT.

4.1 Tick Size and Dark Trading

As noted previously, wider tick size can affect queue jumping by forcing traders who post limit orders to join long queue-lines at the NBBO. For an investor looking to avoid paying the spread, this heightened risk of non-execution can increase the attractiveness of midpoint trading within a dark venue. To examine whether a wider tick size results in longer queue lines at the NBBO, we calculated the quoted depth at the NBB for each security as of the end of each trading millisecond. Figure 4 plots the average log quoted depth at the end of a millisecond as a function of the two-decimal price of the NBB for that millisecond.[27] As the figure shows, quoted depth at the NBB increases discontinuously above the \$1.00 cut-off. Table 4 provides point estimates of this discontinuity using the four different specifications discussed previously. Our preferred estimate indicates that quoted depth at the NBB increases by approximately 115 round lots when priced at \$1.00 relative to when the two-decimal NBB is priced at \$0.99. Considering that quoted bid depth just below \$1.00 was approximately 7 round lots, this point estimate suggests that quoted bid depth at the \$1.00 cut-off is over 15 times greater than that available just below it.

---

[25] We focus on call contracts because the low average price of our sample securities has the effect of greatly diminishing the demand for put contracts. As such, there are substantially fewer put contracts outstanding than call contracts, adding noise to empirical analysis of the put contract data. In unreported estimates using put contracts, however, we corroborate the qualitative conclusion from the call contracts that implied volatility is a smooth function of the closing price in the equities market.

[26] While this approach does not necessarily reflect volatility for each option contract given that they are generally American options, it provides a rough estimate of market-based volatility that is consistently calculated across issuers. In any event, regression of this measure of volatility against the volatility measure calculated by OptionMetrics for an issuer's outstanding option contracts yields an R-squared of .99 with a constant (slope) of approximately 0 (1).

[27] We utilize log quoted depth because of the long right-hand tail of the quoted depth distribution.

[Insert Figure 4]

[Insert Table 4]

Given these longer queue lines in the penny quoting regime, we next analyze whether the penny MPV was also associated with a greater incidence of non-exchange trading as measured by trade executions reported to a FINRA TRF (exchange code=”D”). This analysis requires us to examine the venue of a trade execution (tracked in the Consolidated Trade File) as a function of the quoting regime (tracked in the Consolidated Quote File); therefore, we used the interleaved trade and quote data. A practical challenge in utilizing the interleaved data, however, is the need to assign each trade to the appropriate quoting regime. We address this challenge by assuming that the NBB in effect for a security at the beginning of a trading second (a “security-second”) represents a good estimate of the NBB in effect for the duration of that security-second. To examine the reasonableness of this proxy, Figure 5 plots the NBB that prevailed for any given millisecond of a security’s trading day against the NBB as of the beginning of that second. As the figure reveals, the NBB as of the beginning of a security-second was virtually a perfect predictor of any randomly drawn NBB for that security-second. Accordingly, we use the NBB as of the beginning of a security-second as our proxy for the NBB that applied to all trades executed over the course of that second.

[Insert Figure 5]

Using this approach, we calculated for each security-second the number of trades reported to a FINRA TRF as a function of the two-decimal NBB prevailing at the beginning of that security-second. Figure 6 provides a scatterplot of the results. As the figure shows, trades occurring when the NBB was at or above $1.00 per share revealed a sharp increase in TRF-reported executions. Table 4 provides point estimates of this discontinuity, which confirm a significant jump in the rate of FINRA trades just above the cut-off. Overall, these results are consistent with the findings of KMM concerning the effect of a wider tick size on the incidence of dark trading.

[Insert Figure 6]

While Figure 4 and Figure 6 are consistent with claims that wider tick sizes facilitate dark trading through queue-jumping, neither speaks to whether such queue-jumping represents stepping ahead or subpenny midpoint trading. To examine this question, we examine the incidence of three types of trade executions that occurred away from exchanges during the sample period. First, for non-exchange trades we estimate the incidence of pricing exactly at the midpoint of the NBBO. We assume that such trades reflect orders such as midpoint peg orders that are designed to be filled at the midpoint of the NBBO. Second, for non-exchange trades we estimate the incidence of pricing with exactly $0.0001 of price improvement over the NBBO. To do so, we classify orders as buy or sell initiated using the Lee and Ready (1991) algorithm.[28] We focused on this level of price improvement for two reasons. First, this fixed dollar amount of price improvement represents roughly the same economic value of a trade for transactions priced at $1.00 per share as those priced at $0.99, making the measure robust to changes in the tick size at $1.00. Second, commentary on subpenny trades that offer only de minimis price improvement have frequently focused on the use of off-exchange trades that offer just $0.0001 of

[28] Where quoted spreads were $0.0002, we classified a trade with $0.0001 of price improvement as a midpoint trade.

improvement over the NBBO.[29] Finally, for off-exchange trades we estimate the incidence pricing at exactly the NBBO (again, based on the direction of the trade). Given well known limitations of the Lee and Ready (1991) algorithm with regard to mismatching trades to the appropriate NBBO (Holden & Jacobsen, 2011), we limited our analysis in all three cases to trades that occurred at or within the prevailing NBBO and to trades where the NBBO was neither locked nor crossed.

Figure 7 plots the rate of each form of off-exchange trading as a function of the two-decimal NBB that applied to each trade. Figure 7A presents our core finding regarding the frequency of midpoint trading at the $1.00 cut-off. As the figure reveals, midpoint trading demonstrates a sharp, discontinuous increase as the NBB crosses above the $1.00 threshold, highlighting a dramatic change in the incidence of this form of trading in the penny and subpenny quoting environments. Table 4 reports point estimates of this change at the $1.00 cut-off; these indicate a discontinuous increase of 11-12 percentage points in the frequency of midpoint trading. In combination with the sharp change in quoted depth presented in Figure 6, Figure 7A is consistent with traders in the subpenny quoting environment opting to route orders to exchanges rather than to seek midpoint executions given the lower queue lines at the NBBO for orders priced less than $1.00 per share.

[Insert Figure 7]

In contrast, Figures 7B and 7C reveal the opposite result with respect to the frequency of subpenny stepping ahead and trades priced at exactly the NBBO within non-exchange venues. Point estimates in Table 4 reveal a discontinuous drop just above the $1.00 cut-off in both forms of trades, particularly with respect to the measure for stepping ahead. This latter finding is in direct conflict with the conventional wisdom that the penny MPV favors dark trading venues because of the ability of such venues (but not exchanges) to engage in subpenny stepping ahead. To the extent such trades arise primarily from broker-dealer internalizers, however, the finding is entirely consistent with such firms' best execution obligation. This obligation requires that upon receiving a market order, an internalizer seeking to trade against it must ensure that the price reflects the best price reasonably available for the customer. As noted previously, a broker-dealer that routinely fills such orders at the NBBO potentially risks violating this obligation, thereby providing a significant incentive to provide subpenny price improvement for any internalized trade regardless of price. Critically, because subpenny price executions are permissible across the $1.00 cut-off, there is therefore no *a priori* reason why the size of the MPV should necessarily affect an internalizer's decision to provide subpenny price improvement for any given trade.

At the same time, the organizational structure of broker-dealer internalizers can easily lead to a greater portion of retail orders flowing to an internalizer's market making desk in light of the significant drop in midpoint orders below the $1.00 cut-off. For instance, internalizers purchasing retail market orders commonly route a portion of these orders to affiliated dark pools where they interact with contra-side institutional orders, principal and agency trades by the internalizer's market-making desk, and trading interest form other broker-dealers.[30] Because these venues typically observe price-time priority

[29] Our results are also robust to using any amount of price improvement up to $0.005 per share (or one-half of a penny spread).
[30] For instance, the Retail Market Making Desk of UBS (one of the largest internalizers) routes retail market orders acquired from retail brokers to its dark pool where the orders can interact with contra-side institutional orders, UBS principal trades, and orders from HFT firms. Moreover, the Form ATS for UBS makes clear that this structure allows retail market orders to interact with midpoint peg orders received from subscribers to the dark pool. In addition, a January 2015 settlement by UBS with the SEC highlights how retail market orders were also filled by HFT firms who subscribed to the dark pool with the intention of trading

rules, retail orders will therefore be filled by any available midpoint liquidity before being filled by a liquidity provider at a price that is closer to or at the NBBO. To the extent retail orders flow to these venues, the sharp drop in the availability of midpoint orders should therefore be expected to result in a greater share of retail market orders being filled with the residual liquidity offered by a market-making desk (or other dark pool subscriber) with little or no price improvement. While the coarse nature of the TAQ data prohibits empirical testing of this hypothesis, the increase just below the $1.00 cut-off in both stepping ahead (Figure 7B) and in dark trades priced at exactly the NBBO (Figure 7C) is consistent with this institutional structure.

4.2 Liquidity provision

To examine how queue-jumping affects the supply of displayed liquidity, we analyzed price competition among providers of displayed liquidity across and within trading venues. We leveraged the reporting obligations imposed by Rule 602(b) of Reg. NMS to construct our measure of displayed price competition. As noted previously, all exchanges and FINRA are parties to a Consolidated Quotation Plan that requires the reporting to SIAC for publication in the CQS any change in its Consolidated BBO (including aggregate quotation sizes). Moreover, under Rule 602(b), each broker dealer is obligated to honor its bids and offers submitted to a venue for inclusion in the venue's Consolidated BBO, thereby creating strong incentives for a broker-dealer to revise and update as promptly as possible its posted orders as the broker's buying or selling interest changes.[31]

For any given stock the result is a steady stream of changing BBOs across exchanges throughout the trading day. Table 5, for example, shows the data from the Consolidated Quote File for the company StemCells, Inc. on June 1, 2011 for the ten second interval following 10:56:41.000. As reflected in the table, activity at each reporting exchange can easily be inferred from changes in either order price or size (reported in round lots). The first two rows, for instance, reveal that on Exchange "P" (the NYSE Arca) a broker that had previously had a sell order posted at $0.6551 at 10:56:41.157 had by 10:56:42.867 reduced its sell order size by 4 round lots (or 400 shares). As shown at 10:56:45.540, updates occurring within a single millisecond are reported as having the same time entry; however, TAQ preserves the order in which the SAIC received the BBO update (Hasbrouck, 2010). As such, the TAQ data provides time series information on how frequently quotes comprising exchanges' Consolidated BBOs are changing across exchanges. We use the rate of quote updates per second as our initial measure for price competition among display liquidity. Our reasoning is that a larger supply of displayed liquidity should, all other things being equal, yield a greater rate of quote updates of exchanges' Consolidated BBOs as brokers compete to optimize pricing and depth at the top of a venue's order book.

---

with retail orders at or near the NBBO in the same fashion as a conventional broker-dealer internalizer. In particular, HFT subscribers used an order type referred to as "PrimaryPegPlus" (PPP) to submit orders pegged to the NBB or NBO, plus or minus a subscriber-entered percentage of the quoted spread which was often as little as one percent. Because these orders were priced superior to the NBBO, they received execution priority over orders resting at the NBB or NBO given the dark pool's time-price priority rules. For this reason, a retail trade that failed to secure a midpoint execution could thus be executed with de minimis price improvement within the dark pool. As summarized by the SEC, "[w]hen a resting PPP order executed in UBS ATS, the order that executed against it – such as one from a retail broker-dealer – received a slightly better execution price than if the trade had occurred at the bid or offer." UBS eliminated the PPP order in 2012; however, it continues to permit subscribers to offer de minimis subpenny price improvement for orders priced less than $1.00 per share.

[31] Rule 602(b)(3) specifically encourages such updating by broker-dealers by discharging a broker-dealer from its obligation to honor a previously posted bid or offer so long as "prior to the presentation of an order for the purchase or sale of a subject security, a responsible broker or dealer has communicated to its exchange or association … a revised quotation size."

[Insert Table 5]

Figure 8 plots the rate of quote updates per second as a function of the two-decimal price of the NBB at the beginning of each second, revealing a sharp discontinuous drop in quote updates at exactly $1.00 per share. Table 6 presents point estimates. As shown in the figure, quote updates were approximately one half as frequent when the two-decimal NBB was priced at $1.00 per share compared to when it was priced at $0.99 per share. This finding is consistent with a drop in liquidity provision at the $1.00 cut-off; however, the possibility remains that a portion of this finding could relate to the mechanical effect of the finer pricing grid for orders priced less than $1.00 per share. In particular, as suggested by Harris (1999) in the context of decimalization, order and cancellation messages should increase with smaller pricing increments as traders use the finer price increments to seek greater precision in pricing orders.

[Insert Figure 8]

[Insert Table 6]

Accordingly, to examine more precisely the effect of the change in tick size on the provision of displayed liquidity we also analyzed the intra-second volatility of the NBB across the $1.00 cut-off. This analysis required us to measure for each security-second the number of times the NBB changed within the second. To eliminate volatility arising simply from the ability to quote in subpenny prices, we estimated this outcome measure by calculating the price change of the NBB as actually reported (e.g., $1.00, $0.9901, $0.981, etc.) as well as if rounded to two-decimal places (e.g., $1.00, $0.99, $0.98, etc.). Figure 9A plots the volatility of the unadjusted NBB as a function of the two-decimal NBB price as of the beginning of the security-second. As the figure reveals, the rate of within-second changes to the NBB drops dramatically as the two-decimal NBB crosses above the $1.00 cut-off, consistent with a drop in price competition at the NBB. Formal point estimates are provided in Table 6.

Remarkably, as Table 6 and Figure 9B reveal, this discontinuity persists *even when changes to the NBB are measured using the NBB rounded to two-decimal places*. In effect, this latter finding suggests that even if we eliminate the existence of subpenny pricing increments below $1.00 (i.e., by rounding the NBB to two-decimal places), the NBB remains systematically less volatile at $1.00 than at $0.99. We attribute this discontinuous change in two-decimal NBB volatility across the $1.00 cut-off to a drop in the price competition of displayed liquidity above $1.00.

[Insert Figure 9]

4.3 Welfare Effects

While the foregoing results are supportive of the allegation that queue-jumping diminishes the provision of displayed liquidity, assessing the welfare implications of this finding are complicated by several factors. For one, the fact that a significant portion of queue-jumping arises from midpoint trading naturally raises the question of whether the lower trading costs to investors afforded by such trading offsets any impairment to price discovery due to lower price competition on public exchanges. Moreover, evaluating the ultimate effect of queue-jumping on price discovery is made difficult by the emergence of HFT firms as important providers of displayed liquidity. As noted by the SEC in its 2010 Concept Release on Market Structure, passive market making represents a central trading strategy for many HFT firms, with such firms now representing a core component of liquidity provision on public exchanges.

For instance, while market-making on the NYSE was once the province of human specialists, O'Hara, Saar, and Zhong (2014) document how NYSE market-making is now largely carried out by HFT firms who function as Designated Market Makers (DMMs) and Supplemental Liquidity Providers (SLPs).[32] Although several studies have suggested HFT market-making can enhance price discovery (see, e.g., Brogaard, 2010; Hendershott and Riordan, 2011; Jovanovic and Menkveld, 2010; Groth, 2011), regulators and academics have increasingly focused on the possibility that HFT might also contribute to market volatility and overall market instability (see, e.g., Kirilenko, Kyle, Samadi, and Tuzun, 2011; Zhang, 2010, Cratera and Penalva, 2011; Egginton, Van Ness and Van Ness, 2012). The decision in 2014 of Virtu Financial—a prominent HFT firm and one of the NYSE's six DMMs—to pull its initial public offering in light of negative publicity about HFT following the release of Michael Lewis' *Flash Boys* provided a stark illustration of the depth of this growing skepticism about these "new market makers" (Menkveld, 2013).

In light of these concerns, we therefore re-examined price competition across the $1.00 cut-off at the level of the security-millisecond to determine whether the decrease in liquidity provision at the cut-off was likely the result of HFT algorithms providing less liquidity. Although the sharp change in liquidity provision in Figures 8 and 9 (where we used security-seconds) is highly suggestive of this result, we reasoned that the persistence of these results at the level of the security-millisecond would reveal the influence of HFT liquidity providers given that only computer algorithms could respond to prices that changed by the millisecond. Moreover, we also examined the incidence of HFT around the $1.00 cut-off by looking for evidence of "strategic runs" within the quotation data. As described by Hasbrouck and Saar (2010), proprietary algorithms utilized by HFT firms (as distinct from agency algorithms used by institutional investors to minimize trading costs) typically operate in a millisecond environment in which they periodically send a battery of order and cancellation messages within a single millisecond either to trigger or respond to market events. Accordingly, in addition to examining the rate of quote updates per millisecond, we also examined the incidence of security-milliseconds that experienced at least five quote updates over the course of the millisecond. As with our use of security-seconds, we assumed that the two-decimal NBB at the beginning of a security-millisecond represented a good proxy for NBB prevailing over the duration of the millisecond, which is confirmed by Figure 5.

Figure 10 presents the results of this approach. Figure 10A and 10B replicate Figures 8 and 9B using the security-millisecond rather than the security-second as the unit of analysis. As the figures indicate, evidence of a discontinuous change at the $1.00 cut-off in quote updates and in NBB volatility (rounded to 2-decimal places) persists even when the unit of analysis is the security-millisecond. These results are confirmed in Table 7 which provides point estimates for the discontinuities. With regard to the presence of strategic runs, Figure 10C similarly confirms the importance of HFT firms in the sharp drop in liquidity provision just above the $1.00 cut-off. As suggested by the scale of the y-axis in Figure 10C, the incidence of security-milliseconds with more than five updates is generally an uncommon event; however, the figure nevertheless reveals a large discontinuous drop in the frequency of these security-milliseconds as the NBB crosses above the $1.00 cut-off. In particular, analysis of Figure 10A indicates

[32] Like traditional specialists on the NYSE, the NYSE's DMMs are required to maintain a fair and orderly market in designated stocks and must also quote at the NBBO a certain percentage of the time. NYSE's SLPs are not required to maintain a fair and orderly market, but they are obligated to maintain a bid and ask at the NBBO in each of their designated securities for at least 10% of the trading day.

that security-milliseconds having a 2-decimal NBB priced at \$0.99 are three times as likely to experience five quote updates as those priced at \$1.00.

[Insert Figure 10]

[Insert Table 7]

Because these latter tests indicate HFT algorithms likely account for the change in liquidity provision at the \$1.00 cut-off, our final analysis sought to provide a comparative assessment of what are fundamentally two different trading environments created by the change in tick size at \$1.00. That is, through increasing the incidence of queue-jumping, the wider MPV at \$1.00 effectively transitions the trading environment from a market dominated by intense HFT price-competition on exchanges' displayed order books to one characterized by significant dark trading due to subpenny queue-jumping.

An obvious challenge in undertaking this comparison, however, is the fact that conventional market quality metrics are likely to be affected by the mechanical aspects of the change in MPV at \$1.00. For instance, in unreported RD analyses quoted spreads revealed a discontinuous increase of one-third of a penny (s.e.=0.0005) decrease as prices moved from the penny to subpenny quoting environment, a finding that is almost certainly the result of the MPV changing at \$1.00 given that the penny tick size constrains quoted spreads for our sample securities. Likewise, our analysis in Figure 7B revealed that price-improving orders are commonly priced just inside the NBB or NBO, indicating that effective spreads would similarly be lower under the \$1.00 cut-off simply by virtue of a narrowing quoted spread. While one might conclude that lower quoted and effective spreads in a subpenny quoting environment are by themselves sufficient to indicate a superior trading environment, the considerable drop in quoted depth under \$1.00 (Figure 4) undermines such a conclusion. For instance, in light of the lower quoted depth under the cut-off, more orders will have to "walk up the book" and execute against limit orders with less attractive prices.[33]

Accordingly, to provide a preliminary assessment of the two trading environments, we turn to a more general inquiry focused on how the sudden change in trading environment occasioned by the change in MPV affects the overall rate of trading. Our general intuition was that the net effect of using a wider tick size to transition between these two trading environments should ultimately be reflected in how well each environment facilitates actual trading activity, which we assume reflects an efficient allocation of assets. We therefore examine for each security-second the number of trades reported to the consolidated tape. As above, we further assume that the NBB as of the beginning of a security-second provides a reasonable estimate of the NBB that prevailed for trades occurring within that security-second to estimate how the rate of trading changes over the \$1.00 cut-off.

Figure 11 plots both the rate of trading per security-second as a function of a security-second's two-decimal NBB. As the figure reveals, trades per second show a discontinuous increase just above the

[33] Consistent with this concern, examination of Intermarket Sweep Orders (ISO) at the \$1.00 cut-off reveals a discontinuous decrease as the price moves from the subpenny to penny quoting environment (unreported). In general, under Reg NMS, an order marketed as an ISO is exempt from the Order Protection Rule. As such, a trading venue receiving an inbound liquidity-taking ISO can fill it without checking other venues for superior prices. However, the broker sending the ISO is responsible for sending simultaneous orders that sweep all venues holding superior prices. As such, ISO orders allow a trader to sweep through multiple levels of a venue's order book, thereby providing a proxy for trading interest that is too large to be filled using available depth at the NBBO.

$1.00 cut-off, which is confirmed formally in Table 7. We interpret this findings as indicating that, while wider tick sizes can harm displayed liquidity, they may improve the probability that trading actually occurs.

[Insert Figure 11]

5. Robustness Check: Maker-Taker Fees

[In process]

6. Conclusion

In light of pending proposals to widen the tick size to a nickel trading increment for certain issuers, we provide compelling evidence of the effect of a wider tick size on the competition for order flow that exists between stock exchanges and non-exchange trading venues. As with prior studies using smaller samples, our market-wide analysis of the trading environment surrounding the current $1.00 cutoff between subpenny and penny quoting rules highlights how a wider trading increment can enhance subpenny queue-jumping in dark pools, resulting in more trade executions within non-exchange venues. However, having replicated this result, we part company with prevailing academic and policy analyses that critique the effect of the current MPV rule given its potential cost for the provision of displayed liquidity. As noted above, such concerns typically focus on the questionable utility of facilitating trades that queue-jump exchanges given the potential for such trades to impair the incentive investors have to post displayed liquidity on public exchanges. The possibility that queue-jumping occurs primarily by means of subpenny "stepping ahead" only underscores these concerns given the de minimis benefit such trading provides to investors.

Notwithstanding these concerns, our empirical analysis of the trading environment surrounding the change in MPV at $1.00 per share provides several reasons to doubt whether an MPV rule that favors off-exchange venues necessarily harms the overall trading environment. First, our analysis of off-exchange trades during the sample period highlights the critical role of midpoint trading rather than stepping ahead in drawing order flow to dark venues rather than to conventional stock exchanges. As noted previously, this form of trading has historically been viewed as providing meaningful benefits to investors in the form of reduced trading costs. Second, while we show such trading may nevertheless impair the level of price competition on exchanges' displayed order books, we also show that changes in the institutional structure of liquidity provision on public exchanges complicates efforts to evaluate the overall welfare effects associated with this loss of liquidity. In particular, our empirical analysis highlights that the discontinuous drop in price competition above the $1.00 cut-off is most likely the result of diminished activity by firms engaged in HFT. Although empirical assessment of the benefits and costs of HFT remains mixed, our own analysis of the millisecond trading environment certainly highlights the possibility that HFT may in fact impair rather than enhance the price discovery process. In particular, while one can imagine fundamental market or issuer information occasionally increasing intra-millisecond volatility for particular issuers, there seems little reason why the two-decimal NBB should systematically be more volatile within the millisecond for securities priced at $0.99 than those priced at $1.00 but for the greater presence of HFT. In this fashion, the greater presence of HFT below the $1.00 cut-off may itself impair efficient price discovery and stock transactions which could potentially explain the overall greater rate of trading in the penny quoting environment notwithstanding the drop in displayed liquidity above $1.00 per share.

While our evaluation of how the MPV rule affects the overall trading environment must remain preliminary, these results provide intriguing evidence that pricing rules favoring non-exchange venues may have positive welfare benefits given the evolving nature of HFT liquidity on public exchanges. Disentangling the ways in which HFT on exchanges and queue-jumping in dark pools can have simultaneous and offsetting effects on market quality represents an exciting area of future research.

References

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell, 2011, Heaping-Induced Bias in Regression-Discontinuity Designs, NBER Working Paper No. 17408.

Bartlett, Robert P., III, and Justin McCrary, 2013, Shall We Haggle in Pennies at the Speed of Light or in Nickels in the Dark? How Minimum Price Variation Regulates High Frequency Trading and Dark Liquidity, Unpublished working paper, University of California, Berkeley.

Bessembinder, Hendrik, and Herbert Kaufman, 1997, A cross-exchange comparison on execution costs and information flow for NYSE-listed stocks, *Journal of Financial Economics* 46, 293-319.

Bessembinder, Hendrik, and Kumar Venkataraman, 2010, Bid-Ask Spreads: Measuring Trade Execution Costs in Financial Markets, in "Encyclopedia of Quantitative Finance" (Rama Cont, Ed.), John Wiley & Sons.

Bright Trading LLC, 2010, Comment Letter on Concept Release on Equity Market Structure, available at http://www.sec.gov/comments/s7-02-10/s70210-63.pdf.

Brogaard, Jonathan A., 2010, High-frequency trading and its impact on market quality, Unpublished working paper, Kellogg School of Management.

Buti, Sabrina and Barbara Rindi, 2012, Undisclosed Orders and Optimal Submission Strategies in a Dynamic Limit Order Market, IGIER Working paper 389.

Buti, Sabrina, Barbara Rindi, and Ingrid Werner, 2011, Dark pool trading strategies. Unpublished working paper. The Ohio State University.

Cartea, Alvara and Jose Penalva, 2011, Where is the value in high frequency trading?, Unpublished working paper, Bank of Spain.

Chakravarty, Sugato and Asani Sarkar, 2002, A model of broker's trading, with applications to order flow internalization, *Review of Financial Economics* 11, 19-36.

Delassus, Romain, and Stephane Tyc, 2011, Subpenny trading in US equity markets, in "Econophysics of Order-drive Markets" (F. Abergel et al., Eds.)., Springer, Milan.

Dick, Dennis, 2010, The Hidden Cost of "Sub-Pennying," *CFO Magazine*, Jan-Feb, 8-9.

Easley, David, Nicholas Kiefer, and Maureen O'Hara, 1996, Cream-skimming or profit-sharing? The curious role of purchased order flow, *Journal of Finance* 51, 811–834.

Egginton, Jared, Bonnie Van Ness, and Robert Van Ness, 2012, Quote stuffing, Unpublished working paper, Louisiana Tech University.

Ferrell, Allen, 2001, A proposal for solving the "payment for order flow" problem, *Southern California Law Review* 74, 1029-1088.

Gresse, C., 2012. Effects of lit and dark trading venue competition on liquidity: the MiFID experience. Unpublished working paper, Université Paris-Dauphine.

Groth, Sven, 2011, Does algorithmic trading increase volatility? Empirical evidence from the fully-electronic trading platform Xetra, proceedings of the 10th International Conference on Wirtschaftsinformatik, Zurich.

Hahn Jinyong, Petra Todd and Wilbert Van Der Klaauw, 2001, Identification and estimation of treatment effects with a regression-discontinuity design, *Econometrica* 69, 201–209.

Harris, Lawrence, 1995, Consolidation, fragmentation, segmentation, and regulation, in ‘‘Global Equity Markets: Technological, Competitive, and Regulatory Challenges’’ (Robert Schwarz, Ed.), Irwin Publishing, New York.

Harris, Lawrence, 1999, Trading in Pennies: A Survey of the Issues, Unpublished working paper, University of Southern California.

Hasbrouck, Joel, 2010, The Best Bid and Offer: A Short Note on Programs and Practices, Unpublished working paper, New York University.

Hasbrouck, Joel, and Gideon Saar, 2012, Low-latency trading, Working paper, New York University Stern School of Business.

Hausman, J., and C. Palmer, 2011, Heteroskedasticity-Robust Inference in Finite Samples, NBER Working Paper 17698.

Hendershott, Terrence and Haim Mendelson, 2000, Crossing networks and dealer markets: competition and performance, *Journal of Finance,* 55, 2071-2115.

Hendershott, Terrence, and Ryan Riordan, 2011, High-frequency trading and price discovery, Working paper, University of California at Berkeley.

Holden, Craig W., and Stacey Jacobsen, 2011, The Breakdown of Standard Microstructure Techniques and What to Do About It, Unpublished working paper, Indiana University.

Ikenberry, David L. and James P. Weston, 2008, Clustering in US Stock Prices after Decimalisation. *European Financial Management*, 14, 30–54.

Imbens, Guido W., and Karthik Kalyanaraman, 2012, Optimal Bandwidth Choice for the Regression Discontinuity Estimator, *Review of Economic Studies*, xx, 1–27.

Imbens, Guido W., and Thomas Lemieux, 2008, Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics*, 142, 615–35.

Jovanovic, Boyan, and Albert J. Menkveld, 2010, Middlemen in limit-order markets, Unpublished working paper, New York University.

Kirilenko, Andrei A., Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2011, The flash crash: The impact of high frequency trading on an electronic market, Unpublished working Paper, MIT Sloan School of Management.

Kwan, Amy, Ronald Masulis, and Thomas McInish, 2014. Trading rules, competition for order flows and market fragmentation. *Journal of Financial Economics* (forthcoming).

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

Lee, David S., 2008, Randomized Experiments from Non-Random Selection in U.S. House Elections, *Journal of Econometrics*, 142, 675-697.

Lee, David S., and Thomas Lemiuex, Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, 48, 281-355.

Macey, Jonathan, and Maureen O'Hara, 1997, The law and economics of best execution, *Journal of Financial Intermediation* 6, 188-223.

MacKinnon, J., and H. White, 1985, Some Heteroskedasticity-consistent Covariance Matrix Estimators with Improved Finite Sample Properties, *Journal of Econometrics*, 29, 305-325.

Madhavan, Ananth and Minder, Cheng, 1997, In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets, *Review of Financial Studies*, 10, 175-203.

McCrary, Justin, 2008, Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test, *Journal of Econometrics*, 142, 698-714.

Menkveld, Albert J., 2012, High frequency trading the new-market makers, Unpublished working paper, University of Amsterdam.

Nimalendran, Mahendrarajah, and Sugata Ray, 2014. Informational linkages between dark and lit trading venues, *Journal of Financial Markets* (forthcoming).

O'Hara, Maureen and Mao Ye, 2011, Is market fragmentation harming market quality? *Journal of Financial Economics* 100, 459-474.

O'Hara, Maureen, Gideon Saar, and Zhuo Zhong, 2014, Relative Rick Size and the Trading Environment, Unpublished working paper, Cornell University.

Securities and Exchange Commission, 2010, Concept release on equity market structure, Release No. 34-61358.

Securities and Exchange Commission, 2013, Hidden Volume Ratios, available at http://www.sec.gov/marketstructure/research/highlight-2013-02.html#.VNkcI_mjMeV.

Tabb Group, 2015, Equities LiquidityMatrix: December 2014, available at http://tabbforum.com/liquidity-matrix/equities

TD Ameritrade, 2010, Comment Letter on Concept Release on Equity Market Structure, available at http://www.sec.gov/comments/s7-02-10/s70210-124.pdf.

Weaver, Daniel, 2011, Internalization and market quality in a fragmented market structure, Unpublished working paper, Rutgers Business School.

Ye, Mao, 2011. A glimpse into the dark: price formation, transaction costs, and market share in the crossing network. Unpublished working paper. University of Illinois.

Zhang, Frank, 2010, High-frequency trading, stock volatility, and price discovery, Unpublished working paper, Yale School of Management.

Zhu, H Haoxiang, 2014, Do dark pools harm price discovery? *Review of Financial Studies* (forthcoming).

Table 1:
Market and Limit Orders Exchanges, Dark Pools, and Broker-Dealer Internalizers

Venue	Limit Orders As % of All Reported Orders	Market Orders As % of All Reported Orders
Nasadaq	80.9%	19.1%
G1 Execution Services	4.1%	95.9%
Credit Suisse CrossFinder	4.4%	95.6%

Table 2:
High, Medium, and Low Liquidity Securities

Liquidity Bin	No of Securities	No. of Trades	No. of BBO Updates	Average Quoted Spread	Average Midpoint Spread
Low†	584	8,527,693	482,543,681	0.346	0.055
Medium††	593	55,922,979	1,027,244,161	0.280	0.02
High†††	683	612,229,488	6,386,985,527	0.013	0.008

†Securities having less than 50,000 in average daily trading volume.
††Securities having more than 50,000 and less than 250,000 in average daily trading volume.
†††Securities having greater than 250,000 in average daily trading volume.

Table 3:
Delisting and Smoothness Tests

Outcome Measure	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Probability of Trading in 180 Days	0.0085 (0.0083)	0.0097 (0.0101)	0.0062 (0.0118)	0.0035 (0.0104)
Implied Volatility	-0.041 (0.0433)	-0.0398 (0.0466)	-0.0659 (0.0536)	-0.0815 (0.0461)
Monthly Returns	0.0121 (0.006)	0.0115 (0.0068)	0.0117 (0.0076)	0.0006 (0.0088)
Quarterly Returns	-0.0011 (0.0128)	-0.0063 (0.0155)	-0.0056 (0.018)	-0.0143 (0.018)

Table 4:
Queue-Jumping at the $1.00 Cut-off

Outcome Measure	Model 1	Model 2	Model 3	Model 4
Quoted bid depth	2.6631 (0.0624)	2.6013 (0.0454)	2.6240 (0.0469)	2.6975 (0.0669)
Rate of FINRA Trades	0.0105 (0.0006)	0.0103 (0.0007)	0.0100 (0.0008)	0.0096 (0.0009)
Probability of Midpoint Trade	0.1177 (0.0042)	0.1148 (0.0053)	0.1124 (0.0052)	0.1153 (0.0045)
Probability of Stepping Ahead	-0.0169 (0.0037)	-0.0155 (0.0042)	-0.0133 (0.0047)	-0.0119 (0.0046)
Probability of Trading at the NBBO	-0.0374 (0.0152)	-0.0379 (0.0172)	-0.0432 (0.0198)	-0.0622 (0.0232)

Table 5:
Quotations for StemCells, Inc. Reflected in the NYSE Consolidate Quote File

DATE	TIME_M	SYM_ROOT	BID	BIDSIZ	ASK	ASKSIZ	BIDEX	ASKEX
20110601	10:56:41.157	STEM	0.6541	5	0.6551	25	P	P
20110601	10:56:42.867	STEM	0.6541	5	0.6551	21	P	P
20110601	10:56:42.873	STEM	0.655	196	0.6551	26	T	T
20110601	10:56:44.360	STEM	0.655	196	0.6551	24	T	T
20110601	10:56:45.540	STEM	0.655	196	0.6551	14	T	T
20110601	10:56:45.540	STEM	0.65	21	0.6551	15	K	K
20110601	10:56:45.540	STEM	0.65	2	0.6551	15	Z	Z
20110601	10:56:45.540	STEM	0.655	158	0.6551	14	T	T
20110601	10:56:45.680	STEM	0.6541	5	0.6551	11	P	P
20110601	10:56:50.323	STEM	0.6541	5	0.659	2	C	C
20110601	10:56:50.540	STEM	0.655	158	0.6551	1	T	T
20110601	10:56:50.543	STEM	0.65	21	0.6593	1	K	K
20110601	10:56:50.543	STEM	0.65	2	0.66	3	Z	Z
20110601	10:56:50.543	STEM	0.655	102	0.6551	1	T	T
20110601	10:56:50.543	STEM	0.65	2	0.67	3	Z	Z
20110601	10:56:50.547	STEM	0.655	2	0.67	3	Z	Z
20110601	10:56:50.610	STEM	0.6541	5	0.66	11	P	P
20110601	10:56:50.610	STEM	0.6541	5	0.66	8	P	P

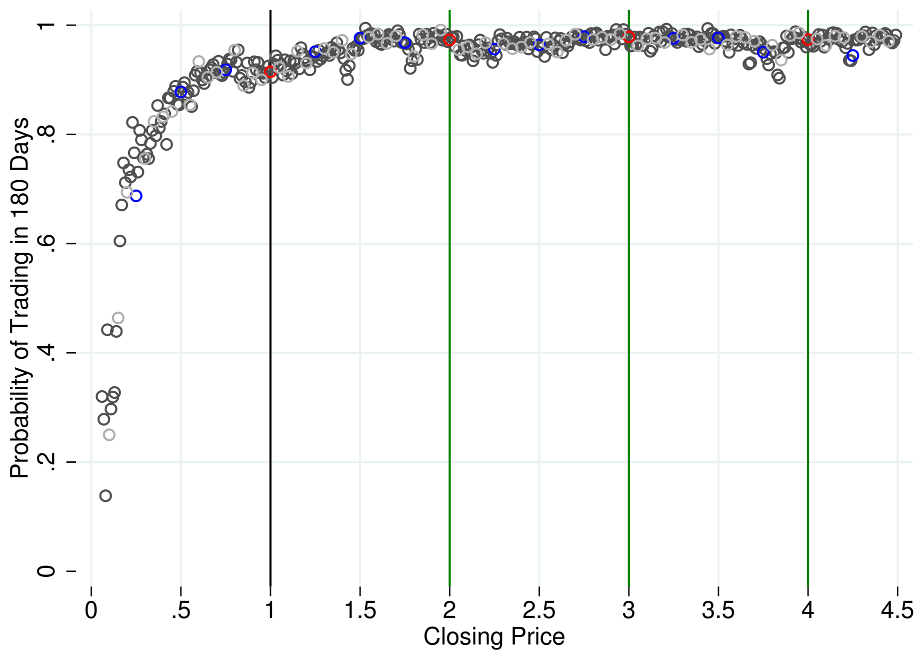
Table 6:
Liquidity Provision at the $1.00 Cut-Off

Outcome Measure	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Quotes per second	-0.5132 (0.0381)	-0.5279 (0.0382)	-0.5506 (0.0404)	-0.6312 (0.0294)
Intra-second NBB Volatility (using unadjusted NBB prices) x 1000	-4.5515 (0.1346)	-4.5385 (0.1485)	-4.5279 (0.1789)	-4.6672 (0.2335)
Intra-second NBB Volatility (using two-decimal NBB prices) x 1000	-0.4740 (0.0739)	-0.4805 (0.0843)	-0.4723 (0.0982)	-0.4720 (0.0744)

Table 7:
HFT and Frequency of Trades at the $1.00 Cut-Off

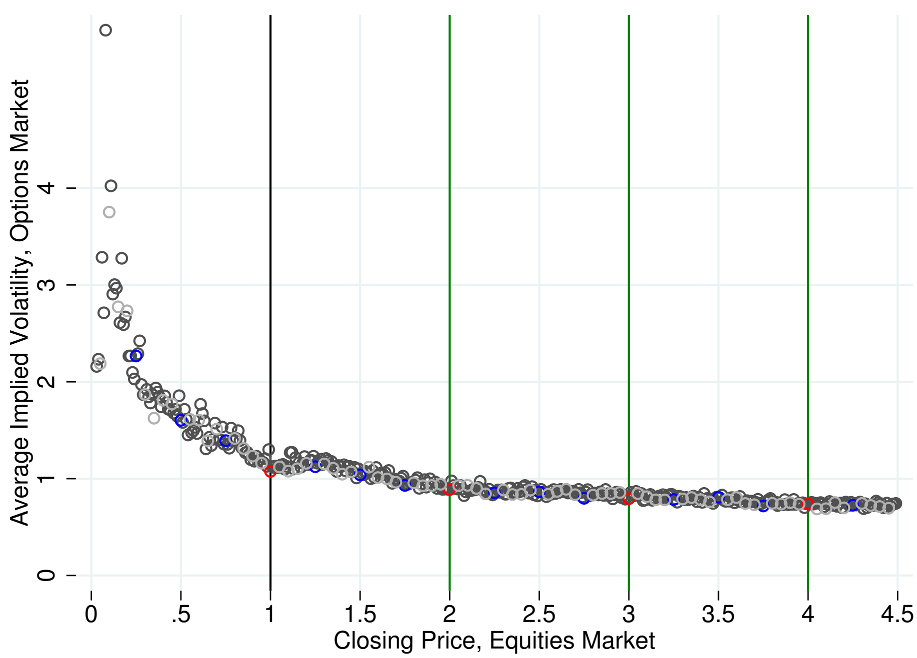
Outcome Measure	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Quotes per millisecond	-1.7858 (0.1375)	-1.8417 (0.1546)	-1.8924 (0.1626)	-1.8993 (0.1426)
Intra-millisecond NBB Volatility (using unadjusted NBB prices) x 100,000	-0.6832 (0.1316)	-0.7141 (0.1453)	-0.7237 (0.1742)	-0.7441 (0.1992)
Intra-millisecond NBB Volatility (using two-decimal NBB prices) x 100,000	-0.0396 (0.0101)	-0.0476 (0.0095)	-0.0495 (0.0109)	-0.0399 (0.0108)
Rate of Trades	0.0096 (0.0028)	0.0092 (0.0032)	0.0082 (0.0037)	0.0021 (0.0045)

Figure 1. Probability of Trading 180 Days Later



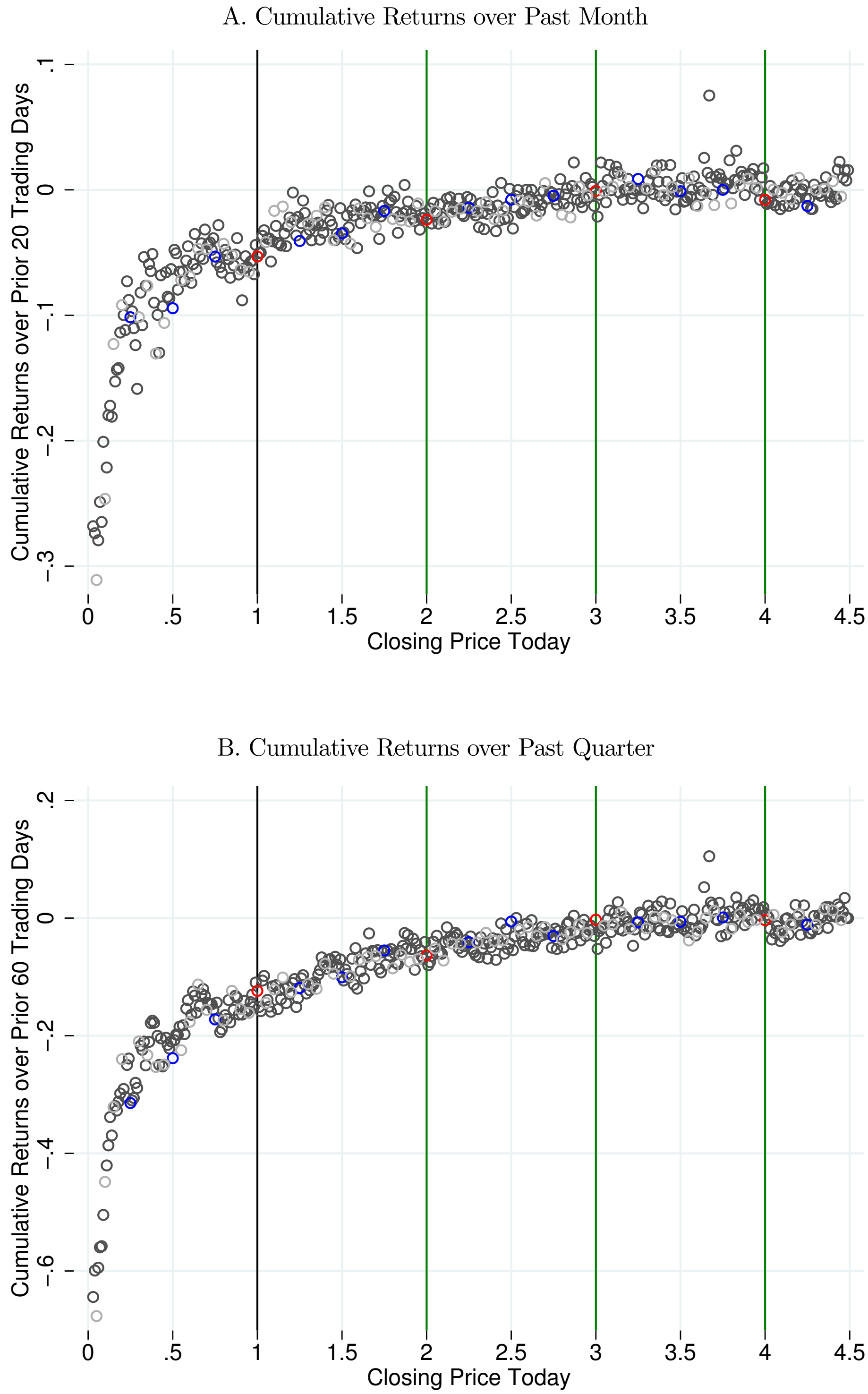
Note: Figure shows the probability, conditional on a two-decimal closing price on day t, of a stock being traded at date $t+180$. Estimates based on CRSP data. See text for details.

Figure 2. Implied Volatility of Options Market Call Contracts



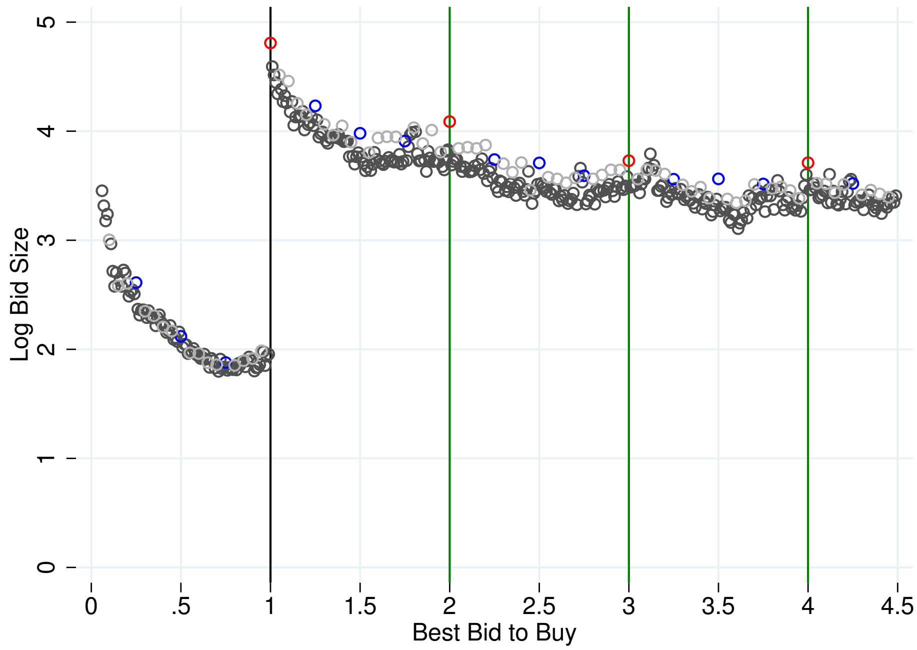
Note: Figure shows the conditional expectation of the implied volatility of options market call contracts given the corresponding closing price in the equities market. Implied volatility based on Black-Scholes using a risk-free rate of 0.5%. Estimates based on OptionMetrics data. See text for details.

Figure 3. Relationship Between Current Prices and Past Cumulative Returns



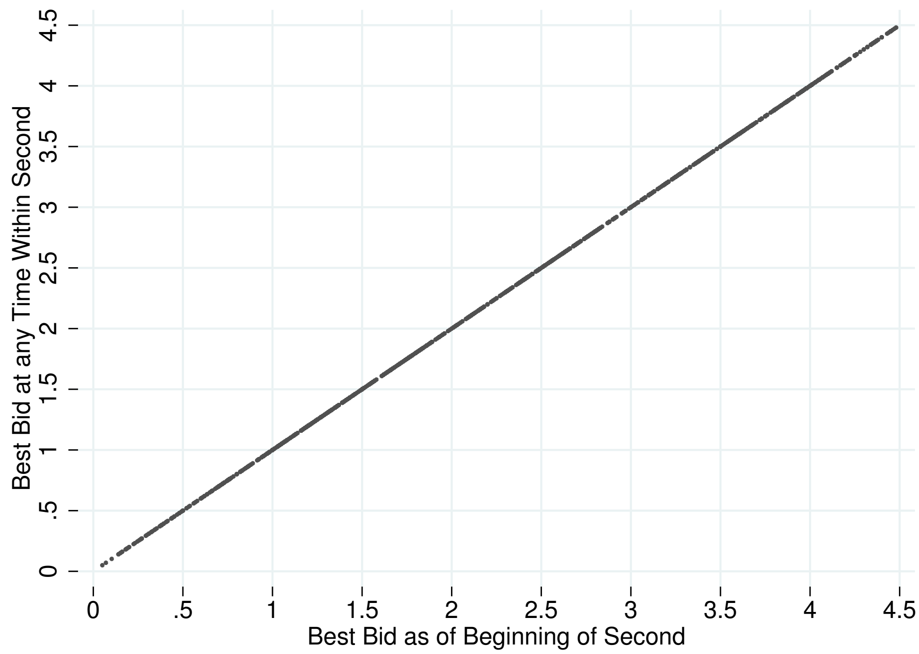
Note: Figure shows conditional expectations of prior returns given current closing price. Panel A (B) shows the conditional expectation of the cumulative returns over the past 20 (60) trading days, calculated as $-1+\exp(\sum_{t=1}^{T}\ln(1+R_{\tau-t}))$, where T is either 20 or 60 and R_τ is the daily return. Estimates based on CRSP data. See text for details.

Figure 4. Inside Depth



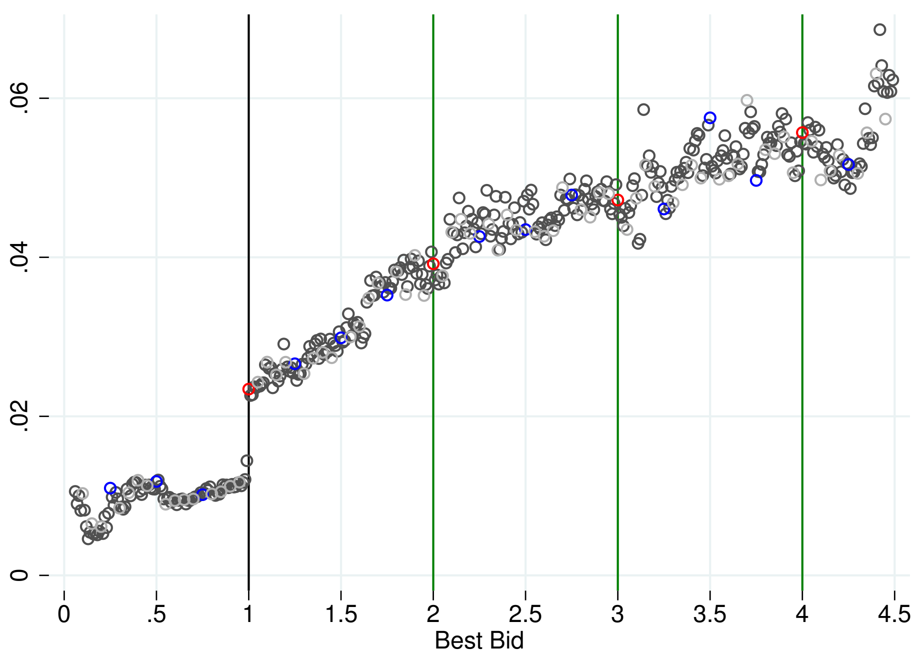
Note: Figure shows the conditional expectation of log bid size given the NBB. Estimates based on TAQ data. See text for details.

Figure 5. Relationship Between the NBB in a Given Millisecond and the NBB at the Beginning of the Second



Note: Figure shows the scatterplot of the NBB at any time against the NBB as of the beginning of the second. Displaying all NBB updates leads to an unworkably large image file, so we display 1,000 updates drawn at randomly selected security-milliseconds. Estimates based on TAQ data. See text for details.

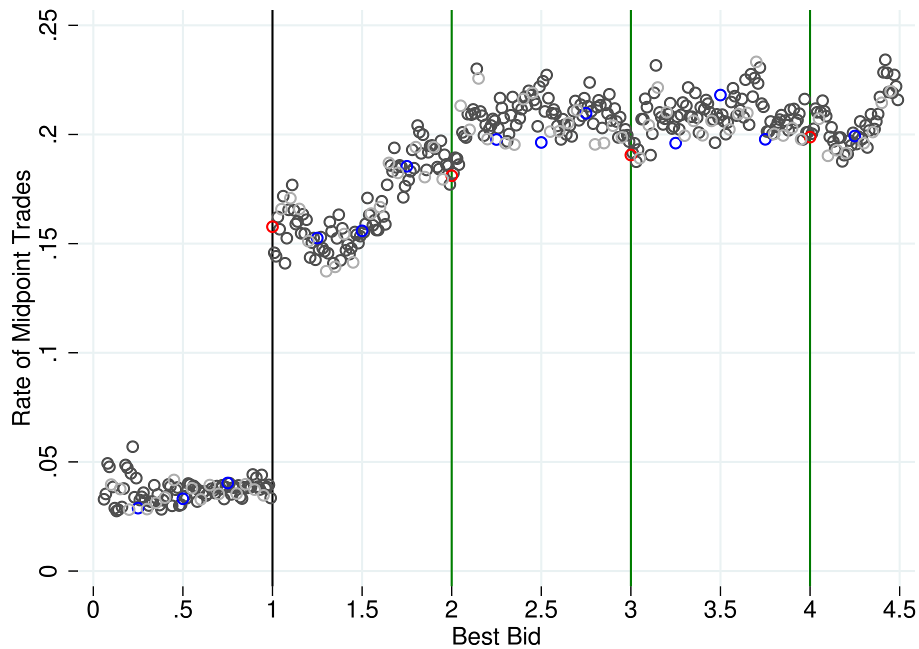
Figure 6. Number of FINRA Trades per Second



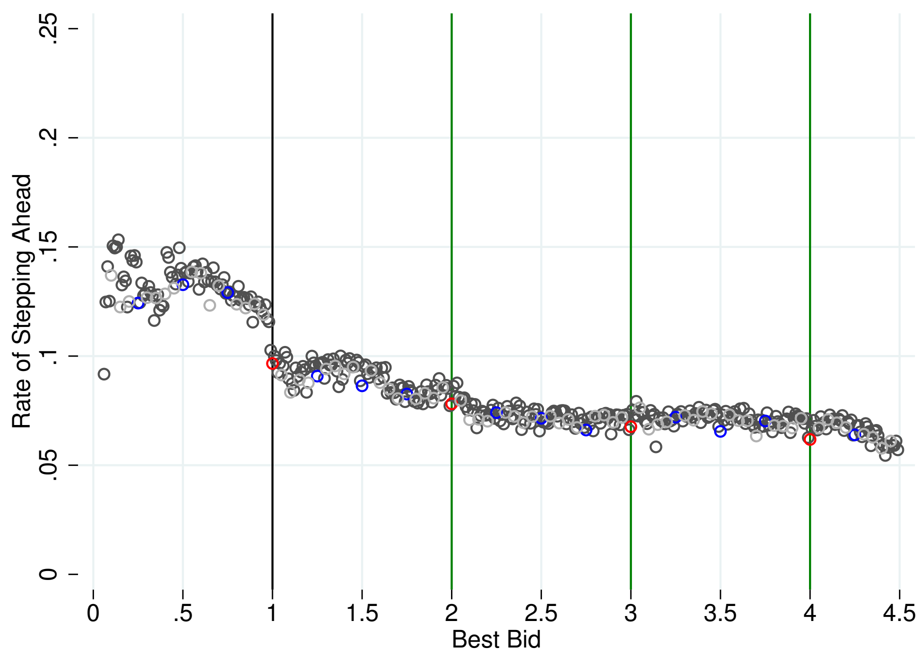
Note: Figure shows the conditional expectation of the number of FINRA trades per second given the NBB as of the beginning of the second. Estimates based on TAQ data. See text for details.

Figure 7. Price Improvement for FINRA Trades

A. Rate of Midpoint Trades

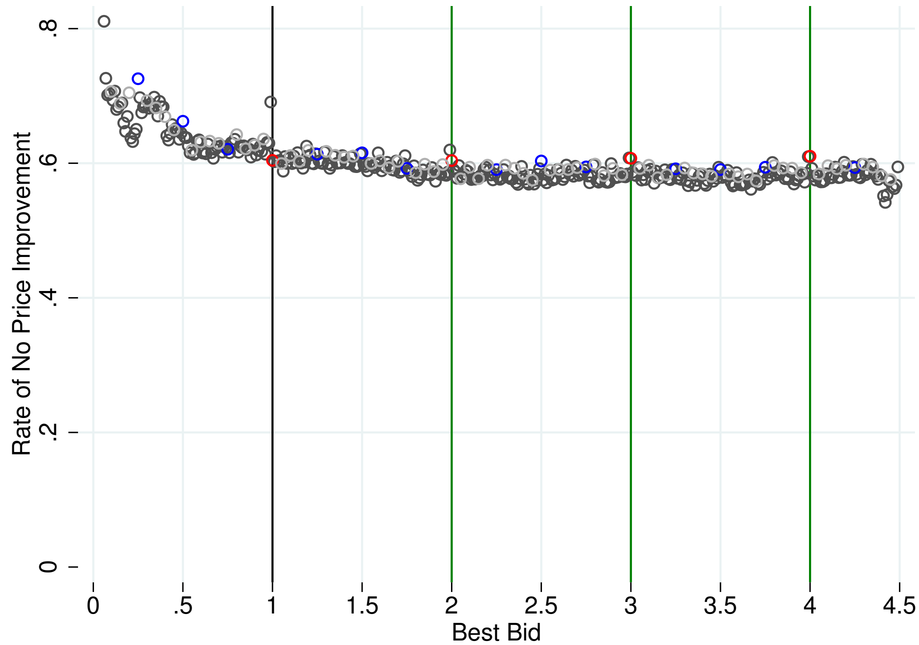


B. Rate of Stepping Ahead



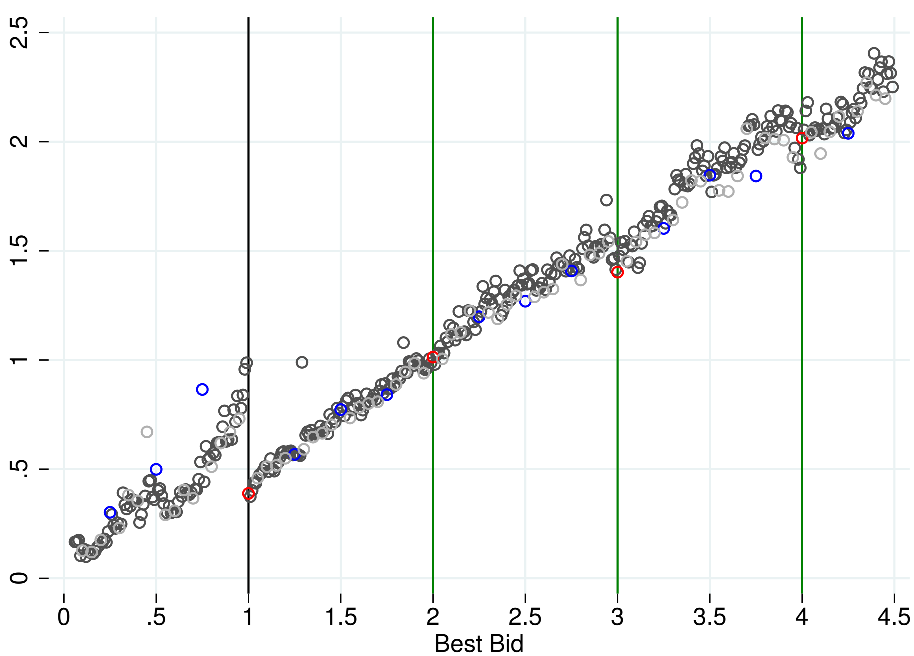
Notes at end of table.

C. Rate of Zero Price Improvement



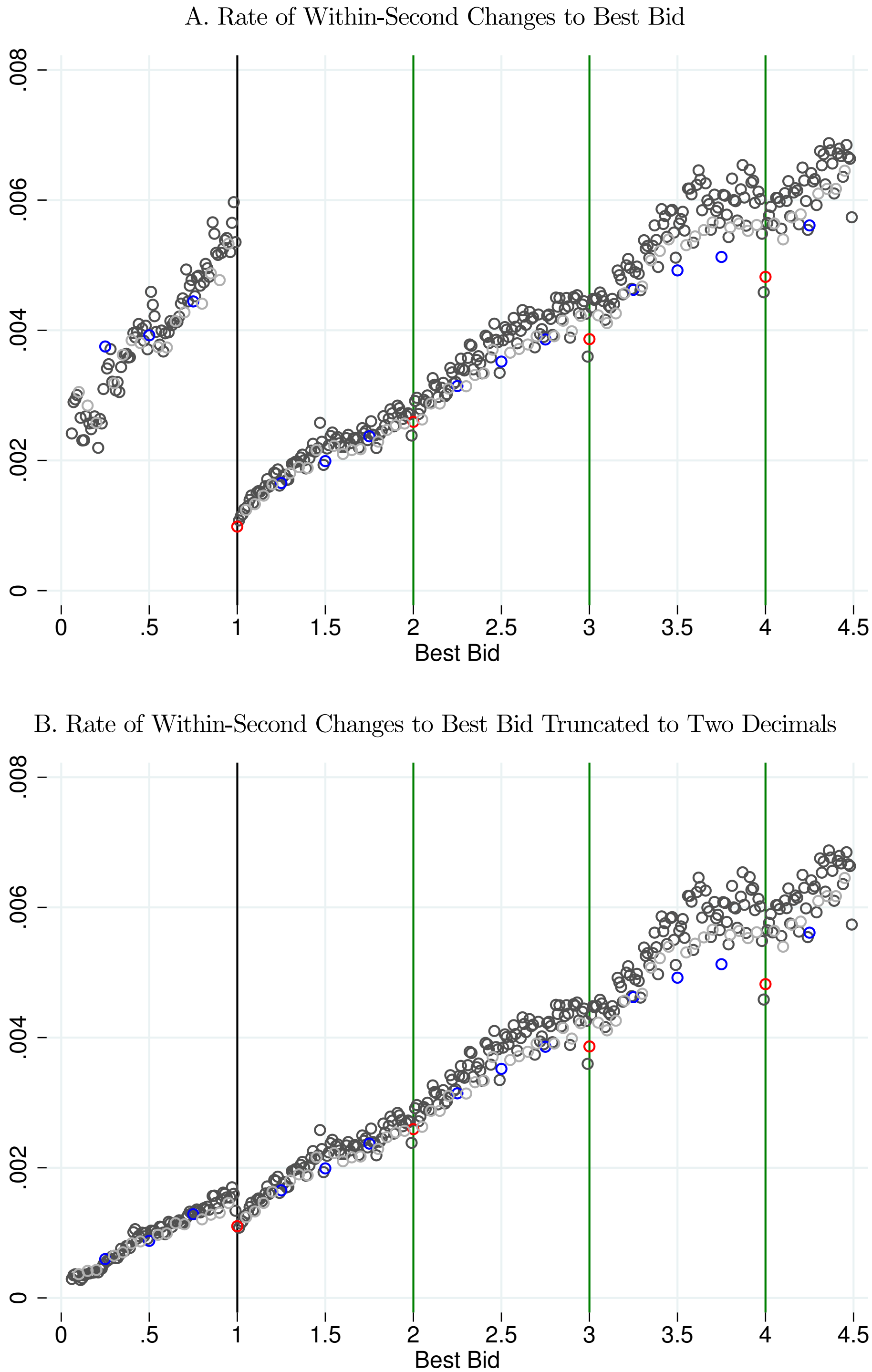
Note: Figure shows conditional probabilities of functions of price improvement, or the difference between trade price and NBO (for buy orders) or NBB (for sell orders), for off-exchange trades. Orders classified using Lee-Ready. Panel A shows rate of midpoint pricing; B shows rate of no price improvement; and C shows rate of price improvement of just 0.0001 exclusive of midpoint pricing. Estimates based on TAQ data. See text for details.

Figure 8. Number of Quote Updates per Second



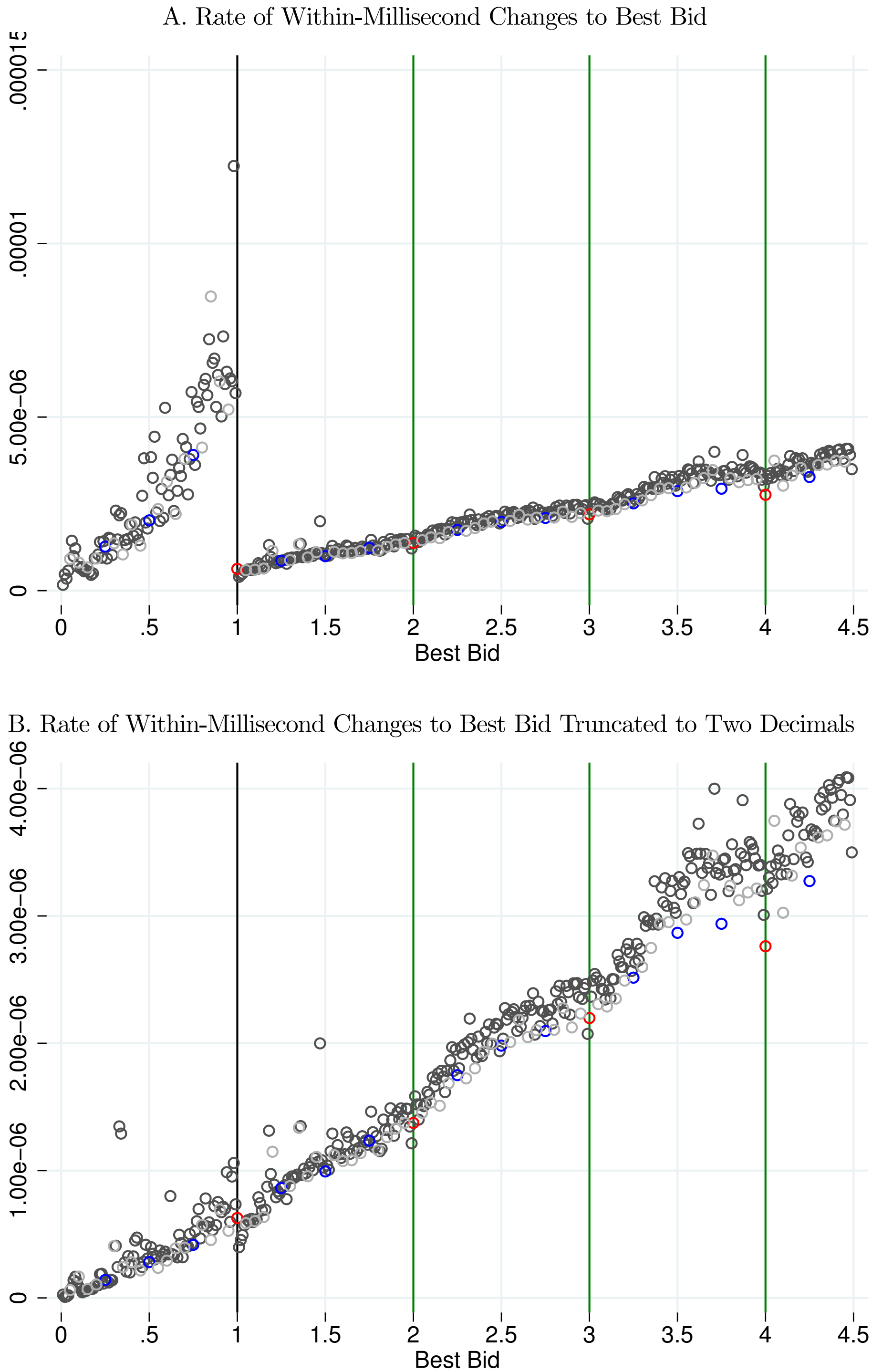
Note: Figure shows the conditional expectation of the number of quotes per second given the NBB at the beginning of the second. Estimates based on TAQ data. See text for details.

Figure 9. Fluctuations in NBB Updates Within One Second



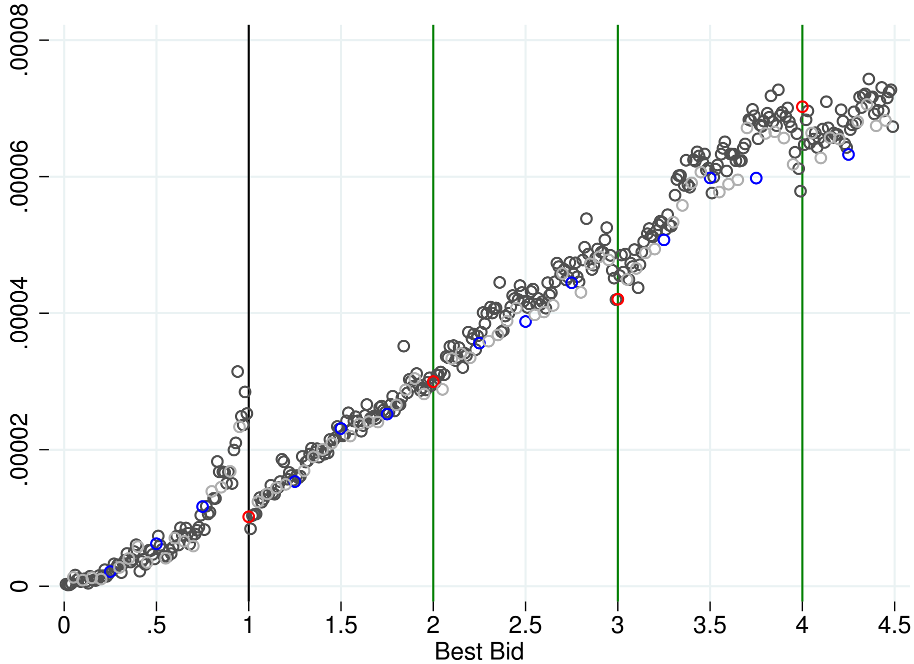
Note: Figure shows rate at which the NBB varies within the second (Panel A) and the rate at which the NBB, truncated to two decimals, varies within the second (Panel B). Estimates based on TAQ data. See text for details.

Figure 10. Presence and Effects of High Frequency Trading: Within Millisecond Evidence



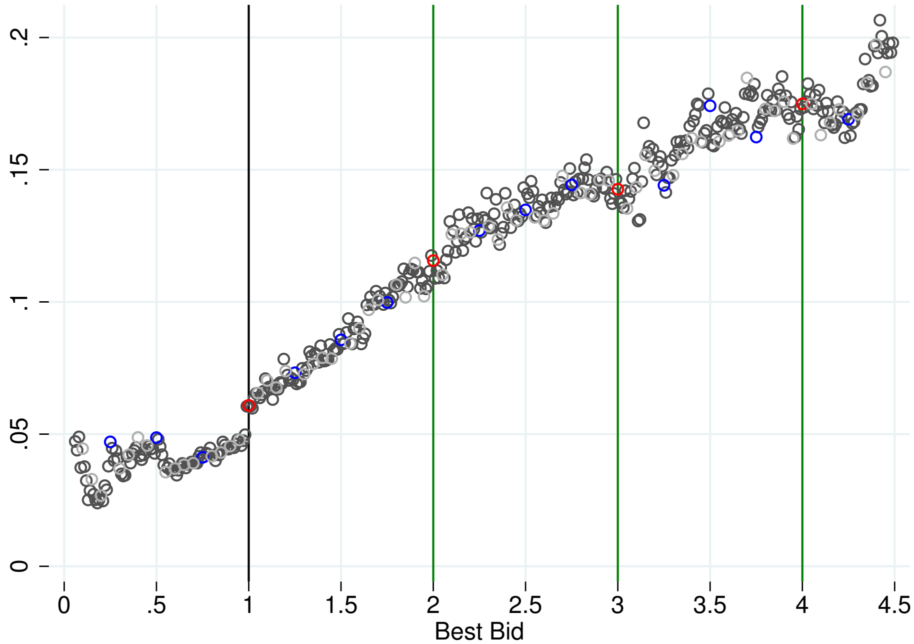
Notes at end of table.

C. Rate of Five or More Quote Updates per Millisecond



Note: Panel A shows the rate at which the NBB differs within the millisecond, as a function of the NBB as of the end of the millisecond. Panel B is the same as Panel A, but truncates the NBB to two decimals. Panel A shows the rate of five or more quote updates per millisecond as a function of the NBB as of the end of the millisecond. Estimates based on TAQ data. See text for details.

Figure 11. Trades per Second



Note: Figure plots the conditional expectation of trades per second given the NBB at the beginning of the second. Estimates based on the TAQ data. See text for details.