

Dark Trading at the Midpoint:
Pricing Rules, Order Flow and High Frequency Liquidity Provision

Robert P. Bartlett, III
University of California, Berkeley

Justin McCrary
University of California, Berkeley, NBER

June 10, 2015

JEL codes: G10, G15, G18, G23, G28, K22

Keywords: high-frequency trading; dark pools; tick sizes; market structure

Dark Trading at the Midpoint:
Pricing Rules, Order Flow, and High Frequency Liquidity Provision

Abstract:

Using over eight trillion observations of market data, we use a regression discontinuity design to analyze the effect of increasing the minimum price variation (MPV) for quoting equity securities in light of recent proposals to increase the MPV from \$0.01 to \$0.05. We show that a larger MPV encourages investors to trade in dark venues at the midpoint of the national best bid and offer. Enhanced order flow to dark venues reduces price competition by exchange liquidity providers, especially those using high frequency trading (HFT). Trading in dark venues due to a wider MPV reduces volatility and increases trading volume.

1. Introduction

A central question for the organization of equity markets is how to provide incentives for traders to provide liquidity on public stock exchanges while protecting the interests of traders wanting to access it. By standing ready to buy or sell a security at disclosed prices, liquidity providers allow liquidity takers to execute trades immediately, facilitate the continuous pricing of securities on which markets depend, and most fundamentally, encourage the holding of securities in the first place. Yet given well-known adverse selection and market risks for standing ready to trade at a fixed price, liquidity providers demand compensation for doing so in the form of the bid-ask spread (Glosten & Harris, 1988). By buying securities at a “bid” quote that is always lower than the “ask” quote demanded for selling, liquidity providers effectively force liquidity takers to compensate them for the public good of displayed liquidity. While in theory this distributional conflict could be resolved through the price mechanism, rules regulating the smallest allowable improvement to a bid or an ask—or a quote’s minimum price variation (MPV)—effectively place a floor on the size of the bid-ask spread.¹ The fact that the MPV has always been regulated thus immediately places securities regulators in the position of allocating gains from trade between liquidity providers and liquidity takers.

In this paper we investigate the effects of the current MPV rule on this allocation of gains in light of mounting concerns that it sub-optimally impairs the provision of displayed liquidity (see, e.g., Kwan, Masulis, and McInish, 2014; Weild, Kim and Newport, 2012; Buti, Rindi, and Werner, 2011). The rule, which is contained in Rule 612 of Regulation National Market System (NMS), requires all quotations submitted to exchanges and priced at or above \$1.00 per share to be priced in penny increments. As a result, the spread between the published national best bid and offer (NBBO) for an actively traded security commonly hovers at or just above a penny. More importantly, in promulgating Rule 612, the SEC explicitly permitted the ability to *trade* in sub-penny increments (the “Trade Exclusion”) so that liquidity takers (e.g., retail traders using market orders) might avoid paying even this penny spread by routing marketable orders to non-exchange “dark” venues such as dark pools and broker-dealer internalizers.² For instance, where

¹ The MPV is currently a penny, but up until 20 years ago was an eighth of a dollar, or 0.125.

² We adopt the conventional practice of classifying non-exchange trading venues as either dark pools or broker-dealer internalizers (see Securities Exchange Commission, 2010). As described in more detail below, dark pools are equity trading systems that do not display their best priced-orders for inclusion in the consolidated quotation feed that is used to calculate the NBBO. While the precise structure of dark pools varies (Zhu, 2014), they generally provide a means to trade securities for

the NBBO for a security is \$10.01 x \$10.02, a trader might seek to sell at a price that is better than the national best bid of \$10.01 by routing a marketable sell order to a dark pool in hopes of crossing with a buy order at the midpoint of the NBBO (e.g., \$10.015). In this fashion, Rule 612's Trade Exclusion induces liquidity takers and their brokers to route orders to dark trading venues instead of exchanges, but at the cost of potentially undermining the incentive of liquidity providers to display quotations on public exchanges. In light of these concerns, the SEC recently finalized a pilot study that will widen the MPV from a penny to a nickel, including a controversial "trade-at" rule that will prohibit any venue from trading a security unless it previously quoted the security at the NBBO.

Motivated by the controversy surrounding the trade-at rule, this paper empirically explores how Rule 612's Trade Exclusion allocates gains from trade between liquidity providers and liquidity takers and, consequently, how it affects the incentive to provide displayed liquidity. Notwithstanding evidence that the Trade Exclusion facilitates off-exchange trading (Kwan, Masulis, and McNish, 2014), the extent to which it actually redistributes gains from liquidity providers to liquidity takers and how it affects the provision of displayed liquidity remain open empirical questions. Because Rule 612 does not currently require a minimum amount of price improvement over the NBBO for off-exchange trades, prevailing studies of this phenomenon have generally emphasized that while traders might route orders to dark venues in hopes of securing price improvement, liquidity providers in dark venues are under no obligation to provide it. As such, liquidity providers in dark venues who use the Trade Exclusion to fill inbound marketable orders instead of competing for those orders on exchanges at the NBBO—a practice commonly described as "queue-jumping"—are presumed to do so by "providing little or no price improvement" (Kwan, Masulis, and McNish, 2014; p. 4). If accurate, such a practice would represent a wealth transfer from providers of displayed liquidity to providers of non-displayed liquidity in dark trading venues, thereby undermining the SEC's purported goal of reducing trading costs for liquidity takers. Yet as we show, the trading rules of the most active dark venues are commonly designed to facilitate trades occurring at the midpoint of the

investors who seek to execute large transactions without displaying their trading interest to minimize the movement of prices against them before a trade can occur. Broker-dealer internalizers represent broker-dealers that fill marketable orders received from customers (or marketable orders purchased from other broker-dealers) using their own proprietary capital or by matching customer orders against one another. Trades executed in either type of venue are, however, reported publicly on the consolidated tape as having been executed at a FINRA trade reporting facility. We use the term "dark venue" to refer collectively to dark pools and broker-dealer internalizers.

displayed NBBO, which effectively allows liquidity takers to avoid paying any spread at all. As such, whatever effect Rule 612's Trade Exclusion has on the incentive to display liquidity, it would at least appear to be allocating gains from trade to the intended beneficiaries.

Equally important, regardless of who benefits from this wealth transfer from providers of displayed liquidity, there are reasons to question whether a trading rule that discourages the display of liquidity on public exchanges necessarily harms the trading environment. Conventionally, U.S. policy in market structure has assumed that greater price competition among providers of displayed liquidity should result in faster, more efficient price discovery while enhancing overall trading liquidity (see, e.g., Securities and Exchange Commission, 2010). In recent years, however, the displacement of human providers of displayed liquidity by algorithmic, high-frequency trading (HFT) firms has given rise to concerns that the liquidity provision offered by HFT firms may have a number of undesirable consequences for the trading environment, such as enhanced volatility (Egginton, Van Ness and Van Ness, 2012; Kirilenko, Kyle, Samadi, and Tuzun, 2011; Zhang, 2010; Cartera and Penalva, 2011) and excessive message traffic (Ye, Yao and Gai, 2012). Given these concerns, if off-exchange trading deters the provision of displayed liquidity by HFT firms, such a result may actually represent a benefit rather than a cost of off-exchange trading to the extent that the market functions better with less price competition among HFT firms on public exchanges. Of course, champions of HFT argue precisely the opposite, that the public exchanges work better now than ever before, and emphasize a view that the prevalence of displayed liquidity makes markets more effective (Modern Markets Initiative, 2014).

To examine these issues empirically, we exploit a regulatory discontinuity in the quoting of U.S. equity securities. Rule 612 requires all quotations priced at or above \$1.00 per share to be priced in penny increments, yet quotations below \$1.00 per share may be priced in sub-penny increments as fine as a hundredth of a penny. As shown in Bartlett and McCrary (2013) and Kwan, Masulis, and McInish (2014), this discontinuous regulatory treatment of quotations priced at or above \$1.00 per share results in a sharp increase in the share of trades occurring in dark trading venues. The reason stems from the fact that the wider quoted spreads at \$1.00 incentivizes liquidity takers to use Rule 612's Trade Exclusion to search for sub-penny trades within the penny-constrained NBBO. The fact that the incentive to use the Trade Exclusion

changes discontinuously at \$1.00 raises the possibility of applying a regression discontinuity (RD) design.

In light of the policy rationale for permitting sub-penny trades in dark venues, we first examine whether Rule 612's Trade Exclusion successfully redistributes trading gains to liquidity takers through facilitating dark trading. We hypothesize that a primary channel through which the Trade Exclusion affects the incidence of dark trading is through its effect on trades done at the midpoint of the NBBO. Specifically, we conjecture that at least 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.³ As emphasized by Kwan, Masulis, and McInish (2014), 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.⁴ We hypothesize that in this environment dark venues can compete with exchanges by offering midpoint pricing that enables investors to post nondisplayed midpoint orders that execute against marketable orders sent to the venue.⁵ Although turning to dark venues raises execution risk due to uncertainty regarding the liquidity there, 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 venues 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

³ We ignore for the moment the effect of exchange access fees and rebates. Because exchanges often pay a rebate for orders that add liquidity to an order book while assessing a fee on orders that take liquidity from it, these so-called maker/taker pricing arrangements 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 venue 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.

⁴ 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.

⁵ As discussed below, the SEC views non-displayed midpoint orders as un-priced orders, and consequently, permissible under Rule 612 regardless of whether the midpoint of the NBBO sits at a subpenny price (e.g., \$10.015 in the previous example).

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 all TAQ data from 2011-2014 for high liquidity securities priced under \$4.50 per share, we provide evidence that a wider MPV 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 venue at the \$1.00 cut-off reveals an approximate 12% discontinuous increase as the MPV moves from sub-penny increments to penny increments. These results highlight that while a wider MPV might facilitate queue-jumping in dark venues, much of the wealth transfer from providers of displayed liquidity is likely to go to liquidity takers.

Turning to the effect of queue-jumping on liquidity provision, we similarly conduct an RD analysis of the \$1.00 cutoff to examine how the rise in queue-jumping at \$1.00 per share 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 traders' enhanced incentive to use Rule 612's Trade Exclusion to search for price-improving trades in dark venues.

While this evidence is consistent with queue-jumping harming the incentive to provide displayed liquidity, analysis of intra-millisecond quote activity confirms that the effect is driven largely by liquidity providers engaged in HFT. In particular, examination of the volatility of the national best bid (NBB) reveals a discontinuous decrease in the rate of within-millisecond changes to the NBB as it crosses above the \$1.00 cut-off, which we interpret as being consistent with a drop in price competition among liquidity providers using HFT algorithms. Remarkably, this discontinuity persists even when changes to the NBB are measured using an NBB truncated to two-decimal places, which effectively eliminates any price volatility below \$1.00 that might be induced mechanically by the ability to price in sub-penny increments. This latter finding is consistent with concerns that HFT might increase price volatility even when HFT firms engage in ordinary market-making strategies. As such, these results provide compelling evidence of the potential for HFT to enhance price volatility in the ordinary course of trading—a characteristic of HFT that has been hotly contested since the publication of Michael Lewis' *Flash Boys* in 2014 but one that has hitherto been empirically unproven.

While these negative effects of HFT complicate assessment of the ultimate welfare effects of an MPV rule that discourages the provision of displayed liquidity, our final empirical analysis provides an initial step in this direction by examining overall trading volume at the \$1.00 cut-off. Notwithstanding the diminished quote competition in the penny-priced trading environment, our RD analysis reveals a discontinuous increase in trading volume above the \$1.00 cut-off. This counterintuitive finding underscores the possibility that even if a wider MPV diminishes displayed liquidity provision because of greater queue-jumping in dark venues, overall market quality may nevertheless reflect a more liquid trading environment given the decrease in HFT activity. As we show and discuss below, all of our findings are robust to changes in maker-taker pricing on exchanges that might differ at the \$1.00 cut-off as well as delisting rules affecting securities that trade below \$1.00 per share.

Our study is most closely related to theoretical and empirical literature examining how the MPV rule can favor dark trading venues relative to public exchanges. Buti, Rindi and Werner (2011) and Buti, Rindi, Wen, and Werner (2013) model how the current MPV rule can cause trades to occur in dark venues. They highlight the ability of liquidity providers in dark venues to use smaller tick sizes to undercut the displayed price in exchanges' limit order books that must display prices using wider ticks. Kwan, Masulis, and McInish (2014) provide empirical support for such queue-jumping in dark venues due to the MPV rule. All of these papers speculate that queue-jumping results in a wealth transfer from providers of displayed liquidity to providers of non-displayed liquidity in dark venues, calling into question whether the MPV rule is advancing its underlying policy goal of reducing trading costs for liquidity takers. Our empirical results are inconsistent with those speculations, however, and show that queue-jumping occurs primarily by means of traders providing midpoint liquidity. This indicates that the MPV rule may in fact be redistributing gains from trade from liquidity providers to liquidity takers as intended by the SEC.

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. To appreciate the gravity of these concerns, note that it is precisely the

public exchanges' limit order books that determine the NBBO and therefore the transaction prices for both displayed and non-displayed venues. In other words, these books dictate trading prices across all venues. A specific, long-standing concern with broker-dealer internalization is the potential for dealers to engage in "cream skimming" whereby 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 queue-jumping in dark venues does indeed diminish price competition among liquidity providers on public exchanges. However, the evidence we marshal also paints a more complex picture than might first be expected. We demonstrate that the nature of the diminished liquidity cited above reflects in fact a diminished amount of high-speed, algorithmic liquidity provision, which is associated with both *greater* price volatility and *less* market activity. While priors may play an important role in interpretation, for many this finding calls into question whether some types of liquidity provision are more valuable than others. In this regard, our analysis contributes to an increasingly heated policy debate concerning whether HFT liquidity provision enhances or harms overall market quality.

While prevailing studies of HFT and market quality have tended to focus on the behavior of HFT firms in stressed markets such as during the Flash Crash (e.g., Kirilenko, Kyle, Samadi, and Tuzun, 2011) or on strategies designed to exploit longer-term investors (e.g., Zhang, 2010), our findings suggest that even ordinary market-making by HFT firms can contribute to enhanced price volatility. These results highlight a potentially negative effect of what is commonly cited as the most beneficial form of HFT trading insofar as it directly reduces quoted and effective spreads (Modern Markets Initiative, 2014).

Finally, our findings have immediate policy implications for the ongoing debate over the optimal tick size for emerging growth companies. As noted previously, the SEC recently finalized a two-year pilot tick size program that will widen the MPV from a penny to a nickel for select companies with the goal of encouraging market-making in the securities of these firms. Our findings suggest that a wider tick size will result in a substantial increase in the use of the Trade Exclusion to secure subpenny pricing in dark venues, which is likely to impair significantly the goal of encouraging the provision of displayed liquidity on public exchanges. To be sure, a component of the pilot also includes a “trade-at” rule to address this issue by prohibiting a venue filling an incoming order unless the venue was displaying the NBBO. Critically, however, the rule exempts from its application any trades executed at the midpoint of the NBBO. In light of our findings that queue-jumping occurs primarily by means of midpoint trading, the proposed trade-at rule is therefore unlikely to stem the flow of trading from exchanges to dark venues.

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.⁶ 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).

These rules benefit providers of displayed liquidity by making it more likely that market orders are routed to liquidity providers quoting on exchanges at the NBBO. However, in balancing the interests of passive liquidity providers and active liquidity traders, the SEC also allows for off-exchange trading. 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.⁷ Although the Order Protection Rule requires a broker-dealer to execute an incoming market order at a price that is no worse than the NBBO, the fact that the 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). To see how this works, suppose the market for a security stands at \$10.00 x \$10.05. A broker receiving from a customer a market order to buy the security could do two things. First, she could route the order to the venue offering to sell at \$10.05. Second, the broker could comply with the Order Protection Rule by simply selling the

⁶ 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 (subject 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.

⁷ 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, because in that case the broker would be competing with her customer. 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 account at a price that would satisfy the customer order." To avoid this conflict of interest, brokers often sell market orders (but not limit orders) to dedicated broker-dealer internalizers under payment-for-order flow arrangements.

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 her best execution obligations, providing a common justification for the practice.⁸

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 rather than principal 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 nondisplayed orders to sell or buy a security at the midpoint of the NBBO.⁹ 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.¹⁰ According to Tabb Group (2015), prominent dark pools such as Barclays DirectEx,

⁸ 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.

⁹ 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.

¹⁰ 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

IEX, and BIDs report more than seventy percent of their 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).¹¹ As the Table reveals, exchanges principally attract limit orders to interact with incoming market orders while non-exchange venues focus on attracting market orders that will interact with one another.

[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 exchange, 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.¹²

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.

¹¹ 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.

¹² 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, Rindi, Wen, and Werner 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 14) have highlighted the SEC's willingness to enforce the rule against both exchange and non-exchange venues alike.

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 sub-penny increments. While this rule also applies to any limit orders posted to a dark pool, the fact that dark pools seek to facilitate the interaction of marketable orders within the NBBO has allowed these venues to use Rule 612's Trade Exclusion. Specifically, these venues avoid Rule 612's restriction on order pricing by using marketable order types that are technically unpriced but still tied to the price of the prevailing NBBO. For instance, to facilitate midpoint trades when the midpoint of the NBBO is a fraction of a penny, the SEC has explicitly endorsed the use of "Midpoint Peg Orders."¹³ Even if the NBBO spread is a penny, these order types allow a trader to submit an immediately executable order that will execute at the midpoint of the NBBO against any incoming marketable orders. For stocks trading with narrow spreads, the 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.¹⁴

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. 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.¹⁵

¹³ 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.").

¹⁴ 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 smart 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)."

¹⁵ 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

2.3. Order Routing of Marketable Orders

The differentiated application of Rule 612 across different trading venues can ultimately have significant implications for order flow given the widespread use of smart-order routers to manage active liquidity. This is 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 broker-dealers to use smart-order routers that check these venues for a trade within the NBBO before routing an order to the exchange holding the best displayed price.¹⁶ In light of these order routing practices, the presence of orders in dark venues that improve on the NBBO in subpenny increments allows these venues to queue-jump exchanges' displayed limit order books.

3. Data and Empirical Design

3.1 Sample Construction

To analyze the effect of the MPV on queue-jumping and displayed liquidity, we use 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 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 the SIAC for dissemination on the Consolidated Tape last sale data in securities listed on the NYSE,

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.

¹⁶ 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.

Nasdaq, the Amex, and all regional exchanges. These reported transactions are then 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 a 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 the SIAC as having been executed at a FINRA TRF.

In addition to the Consolidated Trade file, TAQ also includes a daily Consolidated Quote File that records historical quotation data reported to the SIAC. As with their trade reporting obligations, all exchanges and FINRA are required to report to the 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.

In light of the research questions posed in this paper, the specific TAQ files we use often vary. 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 to do so.¹⁷

Because we are interested in how the change in MPV rule at \$1.00 affects dark trading, we limit our sample to trades and quotes that are priced at less than \$4.50 per share during the four-

¹⁷ 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).

year period spanning 2011-2014.¹⁸ To ensure that all quotes and trades occur during the trading day after the opening cross and before the closing auction, we also filter the TAQ data to exclude quotes and trades occurring before 9:45:00.000 and after 15:35:00.000. 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.¹⁹

Finally, given that many of these securities trade at spreads wider than the penny MPV, we focus our analysis on those securities where the penny MPV is most likely to be a binding constraint. Visual inspection of the data indicated that roughly the third most liquid securities traded at penny spreads; therefore, we restrict the sample to those securities that fell within the top tercile of traded securities based on average daily trading volume.²⁰

With these restrictions imposed on the TAQ data, the final sample contains 793 securities that, over 2011-2014, were associated with roughly 776 million trades and 8 trillion updates to venues' Consolidated BBO.²¹ Average quoted (midpoint) spreads were approximately \$0.011 (\$0.006), indicating that the penny MPV generally represents a binding constraint for traders seeking to display liquidity for these securities.²²

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 2011-2014, 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. In addition to being computationally efficient, this ensures that all of our securities are true securities, as opposed for example to the test securities that are present in the TAQ data and about which documentation is uneven.

¹⁹ When calculating the NBBO, we additionally restrict our analysis to those quotations that are eligible to establish an exchanges' consolidated BBO (i.e., quotation updates having a condition of A, B, H, O, R, or W). When quotation activity represents an outcome measure of interest, however, we use all quotation updates recorded in TAQ. Our results are robust to whether we exclude quotations that are marked as cancelled or corrected or if we strict our attention to only those that are eligible to establish an exchanges' consolidated BBO.

²⁰ Securities having an average daily trading volume of 336,000 qualified to be in the sample.

²¹ The 8 trillion BBO quote updates correspond to just over 6 trillion NBBO updates.

²² 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 lock the market, causing the order to be rejected or converted into a marketable buy (sell) order.

3.2 Regression Discontinuity: Overview

To assess how the MPV rule affects dark trading and liquidity provision on exchanges, we follow Kwan, Masulis, and McInish (2014) 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 to 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

$E[\text{Market Measure}_i | \text{Price}_i]$, 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, as discussed below, one of our analyses uses options data from OptionMetrics, where data on closing prices are available, but intra-day prices are not available; for that analysis, our notion of “security-time” is a given day for one of our sample securities and “price” is the closing price. As another example, in analyses involving the Consolidated Quote File, “security-time” is one of our sample securities during a given second (and in some tests, a given millisecond) and “price” is the security’s NBB as of the beginning of the second (millisecond).²³ We specify below in our empirical results the particular notion being adopted for each analysis.

An important point to note in analyses using the Consolidated Quote Data is that reported data include only updates to exchanges’ Consolidated BBO. As such, randomly selecting reported data from this file would not yield a randomly drawn quotation for a security at a moment in time, and rather would yield an oversample of the most liquid securities. To ensure

²³ We use the NBB in our empirical analyses for ease of exposition. Our results, however, are robust to using the NBO.

that our analyses of daily data and data relying on the intra-day quotation data (e.g., the NBBO) use the same sampling scheme, we utilize all intra-day quotation records, but generate a weighting scheme such that our intra-day analyses of the quoting environment can be thought of as reflecting randomly sampled seconds or milliseconds for our sample securities.²⁴

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 assess the magnitude of the discontinuity in outcomes given price as well as that of pre-determined characteristics 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 candidate for an RD research design, two issues relating to the trading environment around the \$1.00 cut-off could potentially bias our results and we address these issues first. 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 per share 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). Indeed, we observe in the data many securities engaging in reverse stock splits when prices remain below \$1.00 per share for many days.

While these institutional rules make it unlikely that delisting could operate in a discontinuous manner, the best way to assess the possibility is to examine it empirically. We examined each of our securities for each millisecond during our time window from 2011 to 2013. For each such security millisecond, we generate an indicator for whether that security was trading 180 calendar

²⁴ This would appear to be a methodological innovation, but is straightforward to accomplish with standard SAS code.

days in the future. We use that data to compute the probability that the security would still be trading in 180 days, and Figure 1 plots this probability as a function of a security’s two-decimal NBB.

This figure contains a number of distinctive features that are specific to the RD context and deserve explanation.²⁵ First, the probabilities for each two-decimal NBB 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 in Table 2 our formal RD estimates of the discontinuity in the probability of a security trading 180 days later. Because this is our first reported result, and because the structure of Table 2 parallels our remaining tables, we pause here to describe in some detail our methodological approach and choices before returning to a substantive discussion of these results.

The first column in Table 2 presents a basic local linear regression estimate of the discontinuous change in this probability at the \$1.00 cut-off. We use the triangle kernel since that is known to be boundary optimal (McCrary 2008) and a bandwidth of fifty 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 robustness to alternative bandwidths.²⁶ For all four models, standard errors are reported in parentheses below point estimates.²⁷

²⁵ The figure also reflects a number of computational features that are worth mentioning. The TAQ data are stored by trading date. We take repeated advantage of the fact that the mean is the weighted mean of the weighted means; we loop (in parallel where possible) over days, storing up the daily sufficient statistics, and then stack all of the daily sufficient statistics up before computing the mean of the means.

²⁶ Column 3 uses a narrow bandwidth of twenty-five 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 a 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 our data, fifty 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 fifty cents is probably too wide and for others it is probably too narrow. Column 4 seeks to replace transparency

Standard errors are notoriously difficult to estimate in the RD context. Many such designs, and ours is no exception, leverage great numbers of observations, leading to an impression that standard errors would be easy to estimate. However, in the end the focus of the econometric modeling is fitting a curve to perhaps a few dozen precisely estimated averages that do not necessarily follow a linear trend—model misspecification dominates as the sample size grows, rendering the sample effectively small despite the abundance of microdata.

Angrist and Pischke (2009, Chapters 3 and 8), summarizing several decades of experience with heteroskedasticity robust standard errors in the context of small samples, warn that in small samples estimated standard errors are too small. Fundamentally, in a nonparametric context it is hard to avoid problems with spurious rejection using standard t-ratio testing (Wald) approaches, as these involve either asymptotic approximations or assumptions of normality, neither of which are well supported. We provide empirical support for the sensitivity of Wald approaches to inference below (fn. 28).

However, a pair of important recent papers, Cattaneo, Frandsen, and Titiunik (2015, “CFT”) and Sales and Hansen (2015, “SH”), place inference in the RD context on stronger footing, drawing on randomization inference approaches rather than standard Wald approaches. Randomization inference, which R.A. Fisher famously referred to as the “reasoned basis for inference” in the context of the randomized control trial, avoids asymptotic approximations by contemplating a distribution of the test statistic of interest under the null hypothesis obtained in a particular way: the outcomes are held fixed at their observed values, treatment indicators for the sample are chosen at random holding fixed the number of treated, and the test statistic of interest is calculated for each such shuffling of the treatment indicators. CFT and SH both extend these notions to the RD context, but differ somewhat in their implementation. CFT assumes that the indicator for being to the right of the cutoff can be treated as if assigned randomly in a neighborhood of the cutoff, leading to an assumption that that indicator is independent of counterfactual outcomes. The approach taken in SH weakens the assumption of unconditional

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.

²⁷ The standard errors we use in this context are the so-called “HC3” standard errors that adjust for heteroskedasticity. See MacKinnon and White (1985), Angrist and Pischke (2009), and Hausman and Palmer (2011).

independence to conditional independence given the running variable. We follow the approach of SH.

The randomization inference p-value, reported in brackets beneath the estimated standard error, is the fraction of F-statistics so obtained that are more extreme than the F-statistic for the actual data. A p-value below 0.05 is interpreted as evidence of a discontinuous function, whereas a p-value above 0.05 is interpreted as evidence suggesting the function is more likely smooth.

Returning to our substantive discussion of delisting risk, all four specifications confirm the visual impression from Figure 1. Delisting risk is substantial at price points below, say, \$0.25, but does not appear to vary discontinuously at \$1.00 per share. While the estimated t-ratios range from 2.4 to 3, the more reliable randomization inference p-values range from 0.116 to 0.293, with three of the four values larger than 0.27.²⁸

[Insert Table 2]

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

²⁸ To understand the discrepancy between the t-ratios, which suggest smoothness should be rejected, and the random inference p-values, which suggest smoothness should be accepted, consider the distribution of placebo estimates of the discontinuity in delisting risk, presented in Appendix Figure 2. We generated this distribution by estimating a discontinuity at each price point from \$0.50 to \$4.00 using the approach of Column 1 of Table 2. The estimated discontinuities range widely, from -0.076 to 0.036 and the associated t-ratios range similarly widely, from -6.82 to 6.57. Indeed, in the placebo distribution, a t-ratio at least as extreme as the 2.4 observed at the actual cutoff of \$1 occurs 117 times out of the 300 placebo estimates. That can be understood as a p-value of 0.39, where the p-value arises from the placebo distribution. It is the evident fragility of the t-ratio approach in this context that originally motivated our consideration of the randomization inference approach.

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 is 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 is 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.²⁹ For purposes of this analysis, 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%.³⁰ The plot reveals a smooth, downward sloping curve, reflecting a decline in implied volatility as stock prices increase. Table 2 provides point estimates, standard errors, and randomization inference p-values 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 5-day and 100-day returns based on data from CRSP. For each security on trading day t , we compute 5-day (100-day) returns as the conditional expectation, given the security’s closing price on day t , of the cumulative returns over the past 5 (100) trading days, respectively, which we think of as prior

²⁹ 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.

³⁰ 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).

weekly returns and prior long-run returns. Cumulative returns are calculated using CRSP daily closing prices as $-1 + \exp(\sum_{t=1}^T \ln(1 + R_{r-t}))$, where T is either 5 or 100 and R_r is the daily return. Panel A plots weekly returns as a function of the intra-day NBB truncated to two decimal places, and Panel B plots long-run returns. Neither shows evidence of any discontinuity at the \$1.00 cut-off. Point estimates, standard errors, and randomization inference p-values are provided in Table 2 and confirm this conclusion.

Finally, following the recommendation of SH, we additionally compute for all bandwidths from 0.05 to 1.50 randomization inference p-values using Hotelling's T^2 statistic, testing jointly for zero discontinuities in delisting, the implied volatility of call options, prior weekly returns, and prior long-run returns. These results, presented in Appendix Figure 1, show that bandwidths larger than 0.65 are squarely rejected by the data, but that those in the range of 0.05 to 0.55 are likely reliable in this context.

4. Empirical Results

In this section we first re-examine how increasing the MPV to a penny at the \$1.00 cut-off enhances the incentive for traders to engage in so-called “queue-jumping” in dark venues—that is, using Rule 612's Trade Exclusion to post orders to dark venues that fill in-bound marketable orders instead of competing for marketable orders on exchanges at the NBBO. As we demonstrate, traders engaging in queue-jumping typically do so by offering midpoint trades within dark venues. We then document the adverse effect of queue-jumping on price competition among liquidity providers on exchanges, particularly those using HFT algorithms. Finally, given these findings, we examine the welfare effects of a tick size rule that facilitates queue-jumping relative to one that facilitates trading on public exchanges with HFT liquidity providers.

4.1 Tick Size and Dark Trading

As noted previously, a wider tick size can affect queue jumping by forcing traders who post limit orders to join long queue-lines at the NBBO. For example, a trader looking to post a bid at the displayed NBB when a stock trades at \$0.994 x \$0.9999 will face a shorter queue of limit orders than if the same stock had a penny MPV. The reason arises from the simple fact that a penny MPV would require the NBBO to be priced at \$0.99 x \$1.00, pushing any traders willing

to trade at \$0.99 to \$0.994 to post orders at \$0.99. For an investor looking to avoid paying the spread, this heightened risk of non-execution when posting a limit order at the NBBO can increase the attractiveness of midpoint trading within a dark venue.

To examine whether this in fact occurs, we first analyze whether a wider tick size results in longer queue lines at the NBBO. Figure 4 plots for the sample securities the average log quoted depth at the end of a millisecond as a function of the two-decimal NBB for that millisecond.³¹ As the figure shows, quoted depth at the NBB increases discontinuously above the \$1.00 cut-off. Table 3 provides point estimates of this discontinuity using the four different specifications discussed previously. Considering that average quoted bid depth just below \$1.00 was approximately 7 round lots when expressed in logs (i.e., e^2), our preferred estimate indicates that quoted depth at the NBB increases by 87 lots—a nearly twelve-fold increase—when priced at \$1.00 relative to when the two-decimal NBB is priced at \$0.99.

[Insert Figure 4]

[Insert Table 3]

Given these longer queue lines in the penny quoting regime, we next analyze whether the penny MPV is also associated with a greater incidence of non-exchange trading as measured by trade executions reported to a FINRA TRF (exchange code of “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.³²

Using this approach, we calculate 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 5A provides a scatterplot of the results. As the figure shows, TRF-reported trades occurring when the NBB was at or above \$1.00 per share reveal a sharp increase relative to TRF-

³¹ We utilize log quoted depth because of the long right-hand tail of the quoted depth distribution.

³² To examine the reasonableness of this proxy, we regress the NBB for a security-millisecond against the NBB as of the beginning of the associated security-second. The regression results (unreported) yield a constant of 0.0010, a slope coefficient of 0.9993, and an R-squared of 0.9957. These results confirm that the NBB as of the beginning of a security-second was virtually a perfect predictor of any randomly drawn NBB. 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.

reported trades below this price per share. To examine how this phenomenon affects the market share of trading between exchanges and non-exchange venues, Figure 5B plots the fraction of all trades in a second that are reported to a FINRA TRF as a function of the two-decimal NBB prevailing at the beginning of that security-second. Similar to Figure 5A, the fraction of trades that occur off-exchange reveals a large, discontinuous increase at the \$1.00 cut-off. Table 3 provides point estimates of both discontinuities. Overall, these results are consistent with the findings of Kwan, Masulis, and McNish (2014) concerning the effect of a wider tick size on the incidence of dark trading.

[Insert Figure 5]

While Figure 4 and Figure 5 are consistent with claims that wider tick sizes enhance the incentive to use Rule 612's Trade Exclusion to trade in dark venues, neither speaks to how the spread that would otherwise be paid to a liquidity provider on an exchange is distributed among participants in a dark venue. To examine this question, we analyze 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 the execution of midpoint peg orders that are designed to be filled at the midpoint of the NBBO and benefit liquidity takers by eliminating the spread paid on their trade. Second, for non-exchange trades we estimate the incidence of pricing with exactly \$0.0001 of price improvement over the NBBO, using the Lee and Ready (1991) algorithm to classify orders as buy or sell initiated.³³ We refer to this form of trade as "stepping ahead" given that the trade was filled by a liquidity provider in a dark venue offering de minimis price improvement over the NBBO. We focus 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, commentators that are critical of queue-jumping in dark venues have commonly focused on the use of off-exchange trades that offer just \$0.0001 of improvement over the NBBO as evidence that queue-jumping provides few pricing benefits to liquidity takers (see, e.g., Buti, Rindi, and Werner (2011); Dick (2010)).³⁴ For similar reasons, we also estimate for non-exchange trades the incidence of trades priced at exactly the

³³ Where quoted spreads were \$0.0002, we classified a trade with \$0.0001 of price improvement as a midpoint trade.

³⁴ 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).

NBBO (again, based on the direction of the trade) given that such trades simply redistribute trading gains from providers of displayed liquidity on exchanges to providers of non-displayed liquidity in dark venues. Given well known limitations of the Lee and Ready (1991) algorithm with regard to mismatching trades to the appropriate NBBO (Holden and Jacobsen, 2011), we limit 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 6 plots the rate of each form of off-exchange trading as a function of the two-decimal NBB that applied to each trade. Figure 6A 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 between the penny and subpenny quoting environments. Table 3 reports point estimates of this change at the \$1.00 cut-off; these indicate a discontinuous increase of approximately 12 percentage points in the frequency of midpoint trading. In combination with the sharp change in quoted depth presented in Figure 4, Figure 6A is consistent with traders in the penny quoting environment opting to seek midpoint executions in dark venues rather than posting orders to exchanges given the longer queue lines at the NBBO for orders priced at or above \$1.00 per share.

[Insert Figure 6]

In contrast, Figures 6B and 6C reveal the opposite result with respect to the frequency of non-exchange trades that provide little or no price improvement over the NBBO, respectively. Point estimates in Table 3 reveal a discontinuous drop just above the \$1.00 cut-off in both forms of trades, particularly with respect to trades offering just \$0.0001 of price improvement. This latter finding is in direct conflict with the conventional wisdom that the penny MPV favors dark trading venues because of the ability of traders in such venues to queue-jump exchanges through offering *de minimis* price improvement over the NBBO.

We attribute the persistence of trades offering little or no price improvement below the \$1.00 cut-off to the activity of broker-dealer internalizers. As noted previously, broker-dealer internalizers commonly acquire marketable orders from retail brokerage firms that they then fill in accordance with their 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, below the \$1.00 cut-off 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. 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 from other broker-dealers.³⁵ Because these venues typically observe price-time priority 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 prevents us from testing this empirically, the increase just below the \$1.00 cut-off in dark trades with *de minimis* price improvement (Figure 6B) and in dark trades priced at exactly the NBBO (Figure 6C) is consistent with this institutional structure.

³⁵ 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 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.

4.2 Liquidity provision

To examine how the MPV affects the supply of displayed liquidity, we analyze price competition among providers of displayed liquidity across and within trading venues. We leverage 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 the SIAC for publication in the CQS of 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.³⁶

For any given stock the result is a steady stream of changing BBOs across exchanges throughout the trading day. Table 4, 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 the 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 number of BBO updates in a second as our initial measure for price competition among displayed liquidity. Our reasoning is that a larger supply of displayed liquidity should, all other things being equal, yield a greater number of quote updates of exchanges' Consolidated BBOs as brokers compete to optimize pricing and depth at the top of a venue's order book.

[Insert Table 4]

³⁶ 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."

Figure 7 plots the rate of BBO updates in a second as a function of the two-decimal price of the NBB at the beginning of each second, revealing a sharp discontinuous drop in BBO updates at exactly \$1.00 per share. Table 5 presents point estimates. As shown in the figure, BBO 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 7]

[Insert Table 5]

Accordingly, to examine more precisely the effect of the change in tick size on the provision of displayed liquidity we also analyze whether there were within-second changes to the NBB across the \$1.00 cut-off; this can be thought of as measuring whether there are within-second fluctuations of the NBB. To eliminate fluctuations arising simply from the ability to quote in subpenny prices, we estimate this outcome measure by calculating the price change of the NBB as actually reported (e.g., \$1.00, \$0.9999, \$0.9998, etc.) as well as if truncated to two-decimal places (e.g., \$1.00, \$0.99, \$0.98, etc.). Figure 8A plots the within-second fluctuation rate 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 5.

Remarkably, as Table 5 and Figure 8B reveal, this discontinuity persists *even when changes to the NBB are measured using the NBB truncated 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 truncating 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 8]

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).³⁷ 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 temporarily 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-examine 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 displayed liquidity. Although the sharp change in liquidity provision in Figures 7 and 8 (where we used security-seconds) is

³⁷ 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.

highly suggestive of this result, the persistence of these results at the level of the security-millisecond would indicate the influence of HFT liquidity providers given that only computer algorithms could respond to prices that change by the millisecond. Moreover, we also examine 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 BBO updates per millisecond, we also examine the incidence of security-milliseconds that experienced at least five BBO updates over the course of the millisecond. As with our use of security-seconds, we assume that the two-decimal NBB at the beginning of a security-millisecond represents a good proxy for the NBB prevailing over the duration of the millisecond.

Figure 9 presents the results of this approach. As in Figure 8A and 8B, Figures 9A and 9B examine the within-millisecond fluctuation rate of the NBB 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 NBB volatility (truncated to 2-decimal places) persists even when the unit of analysis is the security-millisecond. These results are confirmed in Table 6, which provides point estimates. With regard to the presence of strategic runs, Figure 9C 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 9C, 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 9C indicates that security-milliseconds having a 2-decimal NBB priced at \$0.99 are more than twice as likely to experience five BBO updates as those priced at \$1.00.

[Insert Figure 9]

[Insert Table 6]

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 incentive to trade in a dark venue, 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 immediate 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) as prices move from the subpenny to penny 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 6B reveals 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 below the cut-off, more orders will have to "walk up the book" and execute against limit orders with less attractive prices.³⁸

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 volume of trading. Our general intuition is 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 average number of trades reported to the consolidated tape, the average size of each trade, and the overall volume of shares traded. As above, we further assume that the NBB as of the

³⁸ Consistent with this concern, examination of Intermarket Sweep Orders (ISO) at the \$1.00 cut-off reveals a discontinuous decrease from approximately 44% to 30% in the fraction of trades reported as ISOs as trades moves from the subpenny to penny quoting environment (unreported). In general, under Reg NMS, an order marked as an ISO is exempt from the Order Protection Rule; therefore, 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.

beginning of a security-second provides a reasonable estimate of the NBB that prevailed for trades occurring within that security-second to estimate how total trading activity changes across the \$1.00 cut-off.

Figure 10 presents the results. In Figure 10A, we plot the number of trades in a second as a function of the two-decimal NBB. Notwithstanding the greater amount of price competition below the \$1.00 cut-off, the figure reveals no evidence of a discontinuous change in the rate of trading. This continuity in trades across the \$1.00 cut-off is particularly surprising in light of the sharp decline in quoted depth in the subpenny quoting environment (Figure 4). As shown in Figure 10B, the drop in quoted depth below \$1.00 per share has the unsurprising effect of reducing average trade size. Maintaining the same level of trading volume across the \$1.00 cut-off would accordingly require significantly more trades to occur below this threshold. As shown in Figure 10C, the overall result is a discontinuous drop in total trading volume just below the \$1.00 cut-off. Table 6 presents point estimates for the \$1.00 cut-off for all three plots. We interpret these findings as indicating that, while wider tick sizes can harm displayed liquidity, they may improve the amount of realized trading volume.

[Insert Figure 10]

5. Robustness Check: Maker-Taker Fees

Our results are robust to rebates paid by exchanges for providing displayed liquidity and to fees charged to traders for taking this liquidity. To draw liquidity providers to a venue, exchanges commonly pay liquidity providers a small rebate for every share they sell on an exchange and assess liquidity takers a slightly larger “access fee” for taking it. A small number of exchanges (e.g., BATS Y, Nasdaq BX) also utilize so-called “inverted pricing” whereby the reverse payment structure applies. In these cases, liquidity takers receive a rebate for every share purchased on the venue, and exchanges charge a slightly larger fee to the liquidity provider who provided the liquidity. While the size of rebates paid to liquidity providers is unregulated, Reg. NMS limits access fees for transactions priced above \$1.00 per share to 30 cents per hundred shares while limiting access fees for all other transactions to 0.3% of the transaction value. As such, exchanges during the sample period routinely used different pricing regimes depending on whether a transaction was priced above or below \$1.00 per share.

To examine whether this differential treatment of fees and rebates might affect our results, we collected data on all exchanges' fee and rebate programs from 2011-2012. During this time period, exchanges competed vigorously for trading volume on the basis of their maker-taker pricing, commonly changing their fees and rebates on a weekly basis. This competition was also characterized by considerable experimentation, with some exchanges flipping from standard to inverted pricing structures either entirely or for select securities. We exploit these time-varying changes in maker-taker fees to explore the sensitivity of our results to these pricing programs.

Table 7 illustrates the typical structure of these fees for transactions above and below \$1.00 per share by summarizing the fees and rebates across exchanges for June 3, 2011, a randomly selected trading day. Because exchanges often differ in whether a fee or rebate is payable on a per share basis or as a percentage of the trade, we convert all fees and rebates to reflect their percentage of a \$10,000 trade.³⁹ Moreover, because exchanges often calculate fees and rebates based on the volume of trading conducted by a trader, Table 7 provides the maximum and minimum fee or rebate for each exchange.

[Insert Table 7]

Table 7 highlights the difficulty of assessing the effect of maker-taker pricing on trading across the \$1.00 cut-off. Although the average maximum fee assessed on liquidity takers for a trade priced below \$1.00 per share was generally less than the average maximum fee imposed on a trade priced above the \$1.00 cut-off, several trading venues (including Nasdaq) assessed the same maximum taker-fee regardless of transaction price. At the same time, two exchanges (BATS Y and Nasdaq BX) applied inverted pricing only to trades above \$1.00 per share such that liquidity takers on these venues received a rebate for accessing displayed liquidity for trades priced above this price. Rebates paid to liquidity providers were generally lower for trades priced below \$1.00 per share, but situations could similarly arise where the reverse applied. For instance, a liquidity provider for shares priced below \$1.00 per share on the National Stock Exchange could receive a rebate as high as .25% of the transaction value, while a liquidity provider for shares priced at \$1.00 per share on the CBOE would be required to pay a fee of 0.02% of transaction value.

³⁹ Converting fees and rebates to a percentage of a \$10,000 trade also allows us to analyze the fee and rebate owing on a similarly valued trade regardless of the price paid per share.

In light of this heterogeneity, we re-examine each of our core results controlling for the expected fee and rebate payable on a trade during each security-millisecond. Given the range of fees and rebates payable on a venue, we run the analysis separately using the maximum, the minimum, and the midpoint of the fee and rebate payable on each security on each exchange for each security-millisecond from 2011-2012.

Our approach to controlling for fees and rebates takes advantage of the natural variation in fees and rebates. The fee and rebate structure changes on a millisecond by millisecond basis, as the NBBO moves from one exchange to another, changes over time as exchanges alter the fee and rebate structure, and differs across securities as exchanges sometimes offer specialized programs for some securities. We merge the fee and rebate structure onto our TAQ data so that each security-millisecond is characterized by the NBB, by fees, and by rebates and then compute averages for each unique configuration of the NBB, fees, and rebates, as well as the total number of security-milliseconds corresponding to each configuration. We then estimate the effect of fees and rebates on each outcome by running a fixed effect regression of each outcome on a set of indicators for the NBB, as well as fees and rebates. To maintain a focus on behavior local to the cutoff, the regression uses a triangle kernel that weights observations more if they are closer to the cutoff, and we use the same bandwidths as in our main analyses. The fixed effects for the NBB from this regression, up to scale, are a fee-rebate adjusted version of the local average displayed in the main figures for the paper. We then perform the same analysis as in the main part of the paper, but using the NBB fixed effects in place of the raw means.

Table 8 presents point estimates, standard errors, and randomization inference p-values for the analyses, analogous to the results in the main part of the paper. Regardless of the method for calculating maker-taker fees, our results remain consistent with those obtained without regard to maker-taker pricing. Intuitively, the similarity of results holds because while fees and rebates are sometimes predictive of the outcomes under study, they do not exert a large enough influence to change the results qualitatively.

[Insert Table 8]

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 will enhance the incentive to use Rule 612's Trade Exclusion to engage in subpenny trading in non-exchange venues. However, having replicated this result, we part company with prevailing academic and policy analyses that critique such queue-jumping. As noted above, these criticisms typically focus on the questionable utility of allowing dark venues to queue-jump exchanges which must quote using the MPV given the potential for such trades to impair the incentive investors have to post displayed liquidity on public exchanges. The possibility that queue-jumping might occur with little or no price improvement only underscores these concerns given the questionable 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 queue-jumping necessarily harms the overall trading environment. First, our analysis of off-exchange trades during the sample period highlights the critical role of midpoint trading in drawing order flow to dark venues rather than to conventional stock exchanges. As noted previously, midpoint trading has historically been viewed as providing meaningful benefits to investors in the form of reduced trading costs. Indeed, the SEC expressly adopted this viewpoint in exempting midpoint trade executions from the trade-at rule incorporated in its 2015 pilot study. Given that the trade-at rule was designed to deter off-exchange trading unless a venue quotes at the NBBO, this finding alone has critical implications for the SEC's ability to generate data in the pilot study. With a nickel trading increment, traders will have enhanced incentives to seek midpoint liquidity in off-exchange venues, which will likely result in a significant share of trades being exempt from the trade-at rule.

Second, while we show midpoint trading may 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 analysis of the millisecond trading environment 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 for 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 volume 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 venues can have simultaneous and offsetting effects on market quality represents an exciting area of future research.

References

- Angrist, Joshua D. and Jörn-Steffen Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- 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.
- Buti, Sabrina, Barbara Rindi, Yuanji Wen and Ingrid Werner, 2013. Tick Size Regulation and Sub-Penny Trading. Working paper. Fisher College of Business.
- Cartea, Alvara and Jose Penalva, 2011, Where is the value in high frequency trading?, Unpublished working paper, Bank of Spain.
- Cattaneo, Matias D, Brigham R. Frandsen, and Rocío Titiunik, 2015, Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate, *Journal of Causal Inference*, 3(1), 1—24
- 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.
- Dick, Dennis, 2010, The Hidden Cost of "Sub-Pennyng," *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.

Glosten, Lawrence R., and Lawrence E. Harris, 1988, Estimating the Components of the Bid-Ask Spread, *Journal of Financial Economics* 21, 123-142.

Gresse, Carole, 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 McNish, 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 Lemieux, 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.

Modern Markets Initiative, 2014, When Improving Markets, First Do No Harm, <http://modernmarketsinitiative.org/improving-markets-do-no-harm/>.

- 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.
- Sales, Adam and Ben B. Hansen, 2015, Limitless Regression Discontinuity, Unpublished working paper, University of Michigan.
- 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.
- Weild, David, Edward Kim, and Lisa Newport, 2012, The Trouble with Small Tick Sizes, Grant Thornton Capital Market Series.
- 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.
- Ye, Mao, Chen Yao, and Jiading Gai, 2012, The externalities of high frequency trading, Working paper, University of Illinois at Urbana-Champaign.
- 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 Received at Exchanges, Dark Pools, and
Broker-Dealer Internalizers

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

Table 2:
Delisting and Smoothness Tests

| <i>Outcome Measure</i> | Model 1 | Model 2 | Model 3 | Model 4 |
|------------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Probability of Trading in 180 Days | 0.0111 (0.0046) [0.2790] | 0.0127 (0.0051) [0.2720] | 0.0179 (0.0053) [0.1160] | 0.0158 (0.0052) [0.2930] |
| Implied Volatility | -0.0044 (0.0213) [0.7970] | 0.0002 (0.0237) [0.6480] | -0.0360 (0.0310) [0.6800] | -0.0179 (0.0249) [0.7800] |
| 5 Day Prior Returns | -0.0019 (0.0012) [0.2120] | -0.0025 (0.0011) [0.1460] | -0.0036 (0.0012) [0.3220] | -0.0032 (0.0012) [0.4900] |
| 100 Day Prior Returns | 0.0087 (0.0066) [0.9940] | 0.0092 (0.0071) [0.9610] | 0.0046 (0.0081) [0.4050] | 0.0094 (0.0072) [0.1630] |

Table 3:
Queue-Jumping at the \$1.00 Cut-off

| <i>Outcome Measure</i> | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Quoted Bid Depth | 2.5486 (0.0351) [0.0000] | 2.5284 (0.0342) [0.0000] | 2.4867 (0.0440) [0.0000] | 2.5329 (0.0363) [0.0000] |
| FINRA Trades in a Second | 0.0093 (0.0006) [0.0000] | 0.0091 (0.0007) [0.0000] | 0.0078 (0.0009) [0.0000] | 0.0094 (0.0006) [0.0000] |
| FINRA Trades as a % of All Trades in a Second | 0.1209 (0.0022) [0.0000] | 0.1208 (0.0026) [0.0000] | 0.1210 (0.0027) [0.0000] | 0.1232 (0.0025) [0.0000] |
| Probability of Midpoint Trade | 0.1218 (0.0028) [0.0000] | 0.1217 (0.0032) [0.0000] | 0.1182 (0.0036) [0.0000] | 0.1218 (0.0027) [0.0000] |
| Probability of De Minimis Price Improvement | -0.0173 (0.0023) [0.0000] | -0.0164 (0.0025) [0.0000] | -0.0128 (0.0036) [0.0060] | -0.0169 (0.0024) [0.0000] |
| Probability of Trading at the NBBO | -0.0401 (0.0085) [0.0000] | -0.0418 (0.0096) [0.0010] | -0.0496 (0.0154) [0.0020] | -0.0445 (0.0107) [0.0000] |

Table 4:
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 5:
Liquidity Provision at the \$1.00 Cut-Off

| <i>Outcome Measure</i> | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| BBO Updates per Second | -0.4606 (0.0221) [0.0000] | -0.4696 (0.0231) [0.0000] | -0.5021 (0.0285) [0.0000] | -0.4107 (0.0220) [0.0000] |
| Intra-second NBB Volatility (using unadjusted NBB prices) x 1000 | -4.1733 (0.0696) [0.0000] | -4.1797 (0.0730) [0.0000] | -4.0977 (0.0729) [0.0000] | -4.2074 (0.0640) [0.0000] |
| Intra-second NBB Volatility (using two-decimal NBB prices) x 1000 | -0.3590 (0.0459) [0.0000] | -0.3629 (0.0505) [0.0000] | -0.3353 (0.0635) [0.0030] | -0.3085 (0.0311) [0.0000] |

Table 6:
HFT and Frequency of Trades at the \$1.00 Cut-Off

| <i>Outcome Measure</i> | Model 1 | Model 2 | Model 3 | Model 4 |
|---|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Intra-millisecond NBB Fluctuation Rate (using unadjusted NBB prices) x 100,000 | -0.6225 (0.0304) [0.0000] | -0.6305 (0.0330) [0.0000] | -0.5499 (0.0288) [0.0000] | -0.6348 (0.0266) [0.0000] |
| Intra-millisecond NBB Fluctuation Rate (using two-decimal NBB prices) x 100,000 | -0.0173 (0.0042) [0.0240] | -0.0194 (0.0044) [0.0130] | -0.0166 (0.0068) [0.0880] | -0.0080 (0.0031) [0.6440] |
| Security-Milliseconds with > 5 quote updates x 100,000 | -1.5160 (0.1044) [0.0000] | -1.5344 (0.1171) [0.0000] | -1.7593 (0.1354) [0.0000] | -1.4594 (0.1013) [0.0000] |
| Rate of Trades | 0.0015 (0.0022) [0.3900] | 0.0007 (0.0024) [0.4920] | -0.0034 (0.0034) [0.4930] | 0.0014 (0.0022) [0.3800] |
| Trade Size | 114.2932 (13.9032) [0.0000] | 113.3570 (13.8426) [0.0000] | 111.4452 (16.6577) [0.0000] | 117.4661 (14.7848) [0.0000] |
| Total Shares Traded | 8.3249 (1.4493) [0.0000] | 7.8191 (1.5746) [0.0020] | 6.1148 (2.5248) [0.0050] | 6.3114 (2.4375) [0.0000] |

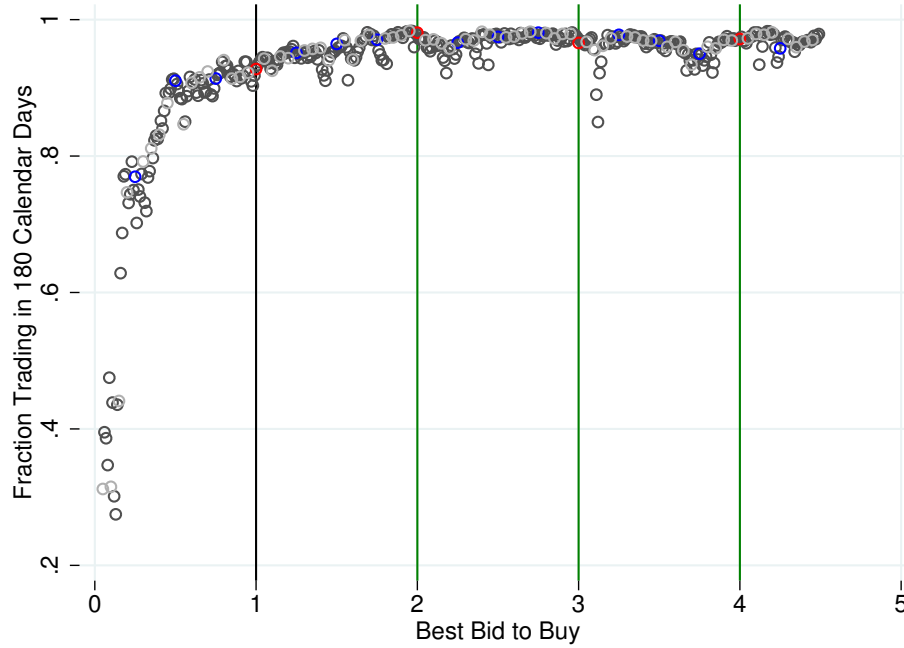
Table 7:
Maker-Taker Fees as of June 3, 2011

| Exchange: | Fee/Rebate for Accessing Liquidity | | | | Rebate/Fee for Providing Liquidity | | | |
|--------------------------------|------------------------------------|--------|----------------------|--------|------------------------------------|--------|----------------------|--------|
| | Share Price = \$1.00 | | Share Price = \$0.99 | | Share Price = \$1.00 | | Share Price = \$0.99 | |
| | Max. | Min. | Max. | Min. | Max. | Min. | Max. | Min. |
| BATS Z | -0.28% | -0.28% | -0.10% | -0.10% | 0.27% | 0.27% | 0.00% | 0.00% |
| BATS Y | 0.03% | 0.03% | -0.10% | -0.10% | 0.00% | -0.02% | 0.00% | 0.00% |
| DirectEdge A | -0.30% | 0.02% | -0.20% | 0.00% | -0.01% | -0.03% | 0.00% | 0.00% |
| DirectEdge X | -0.30% | -0.30% | -0.10% | -0.10% | 0.34% | 0.23% | 0.00% | 0.00% |
| Chicago Board Options Exchange | -0.02% | -0.02% | -0.02% | -0.02% | -0.02% | -0.02% | 0.00% | -0.02% |
| National Stock Exchange | -0.30% | -0.30% | -0.30% | -0.30% | 0.30% | 0.26% | 0.25% | 0.00% |
| Chicago Stock Exchange | -0.30% | -0.30% | -0.30% | -0.30% | 0.31% | 0.25% | 0.01% | 0.00% |
| Nasdaq | -0.30% | -0.30% | -0.30% | -0.30% | 0.30% | 0.20% | 0.01% | 0.00% |
| Nasdaq BX | 0.05% | 0.14% | -0.10% | -0.10% | -0.15% | -0.18% | 0.00% | 0.00% |
| Nasdaq PSX | -0.25% | -0.25% | -0.20% | -0.20% | 0.24% | 0.22% | 0.00% | 0.00% |
| NYSE | -0.23% | -0.23% | -0.23% | -0.23% | 0.35% | 0.15% | 0.05% | 0.00% |
| NYSE Arca | -0.30% | -0.28% | -0.10% | -0.10% | 0.32% | 0.21% | 0.00% | 0.00% |
| NYSE Amex | -0.28% | -0.19% | -0.25% | -0.20% | 0.45% | 0.16% | 0.25% | 0.00% |
| Mean | -0.21% | -0.17% | -0.18% | -0.16% | 0.21% | 0.13% | 0.04% | 0.00% |
| Maximum | -0.30% | -0.30% | -0.30% | -0.30% | -0.15% | -0.18% | 0.00% | -0.02% |
| Minimum | 0.05% | 0.14% | -0.02% | 0.00% | 0.45% | 0.27% | 0.25% | 0.00% |

**Table 8:
Discontinuities at the \$1.00 Cut-Off Controlling for Maker-Taker Fees**

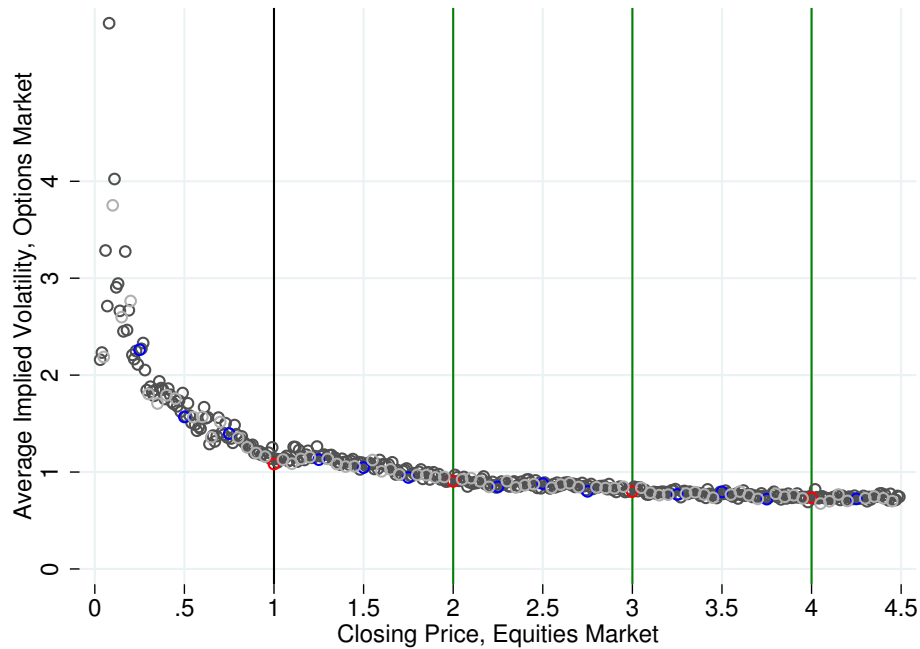
| <i>Outcome Measure</i> | Model 1 | Model 2 | Model 3 | Model 4 |
|---|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Quoted bid depth | 2.5096 (0.0356) [0.0000] | 2.4882 (0.0343) [0.0000] | 2.4497 (0.0453) [0.0000] | 2.5279 (0.0354) [0.0000] |
| FINRA Trades in a Second | 0.0097 (0.0006) [0.0000] | 0.0095 (0.0007) [0.0000] | 0.0082 (0.0009) [0.0120] | 0.0096 (0.0006) [0.0000] |
| FINRA Trades as a % of All Trades in a Second | 0.1181 (0.0020) [0.0000] | 0.1187 (0.0021) [0.0000] | 0.1165 (0.0026) [0.0000] | 0.1203 (0.0026) [0.0000] |
| Probability of Midpoint Trade | 0.1249 (0.0028) [0.0000] | 0.1248 (0.0032) [0.0000] | 0.1213 (0.0036) [0.0000] | 0.1245 (0.0025) [0.0000] |
| Probability of De Minimis Price Improvement | -0.0200 (0.0023) [0.0000] | -0.0192 (0.0025) [0.0000] | -0.0155 (0.0036) [0.0000] | -0.0216 (0.0026) [0.0000] |
| Probability of Trading at the NBBO | -0.0342 (0.0085) [0.0010] | -0.0359 (0.0096) [0.0020] | -0.0437 (0.0154) [0.8580] | -0.0360 (0.0108) [0.0010] |
| BBO Updates per Second | -0.4641 (0.0220) [0.0000] | -0.4733 (0.0231) [0.0000] | -0.5053 (0.0285) [0.0010] | -0.3605 (0.0254) [0.0000] |
| Intra-second NBB Fluctuation Rate (using unadjusted NBB prices) x 1000 | -4.2419 (0.0688) [0.0000] | -4.2501 (0.0728) [0.0000] | -4.1630 (0.0713) [0.0000] | -4.2648 (0.0605) [0.0000] |
| Intra-second NBB Fluctuation Rate (using two-decimal NBB prices) x 1000 | -0.3640 (0.0453) [0.0000] | -0.3692 (0.0502) [0.0000] | -0.3364 (0.0634) [0.3950] | -0.3009 (0.0334) [0.0000] |
| Intra-millisecond NBB Fluctuation Rate (using unadjusted NBB prices) x 100,000 | -0.5846 (0.0309) [0.0000] | -0.5904 (0.0334) [0.0000] | -0.5177 (0.0292) [0.0000] | -0.5772 (0.0258) [0.0000] |
| Intra-millisecond NBB Fluctuation Rate (using two-decimal NBB prices) x 100,000 | -0.0184 (0.0042) [0.0170] | -0.0204 (0.0045) [0.0100] | -0.0181 (0.0068) [0.6180] | -0.0035 (0.0031) [0.6490] |
| Security-Milliseconds with > 5 quote updates x 100,000 | -1.4783 (0.1041) [0.0000] | -1.4975 (0.1167) [0.0000] | -1.7173 (0.1356) [0.0060] | -1.2609 (0.0969) [0.0000] |
| Rate of Trades | 0.0033 (0.0022) [0.1740] | 0.0026 (0.0024) [0.2690] | -0.0016 (0.0034) [0.2480] | 0.0047 (0.0021) [0.1920] |
| Trade Size | 106.6080 (13.3088) [0.0000] | 106.9286 (12.2659) [0.0000] | 102.6688 (15.8390) [0.0000] | 116.8140 (13.8474) [0.0000] |
| Total Shares Traded | 9.7549 (1.4399) [0.0000] | 9.2844 (1.5726) [0.0000] | 7.4864 (2.5010) [0.0010] | 9.8910 (2.1071) [0.0000] |

FIGURE 1. PROBABILITY OF TRADING 180 DAYS LATER



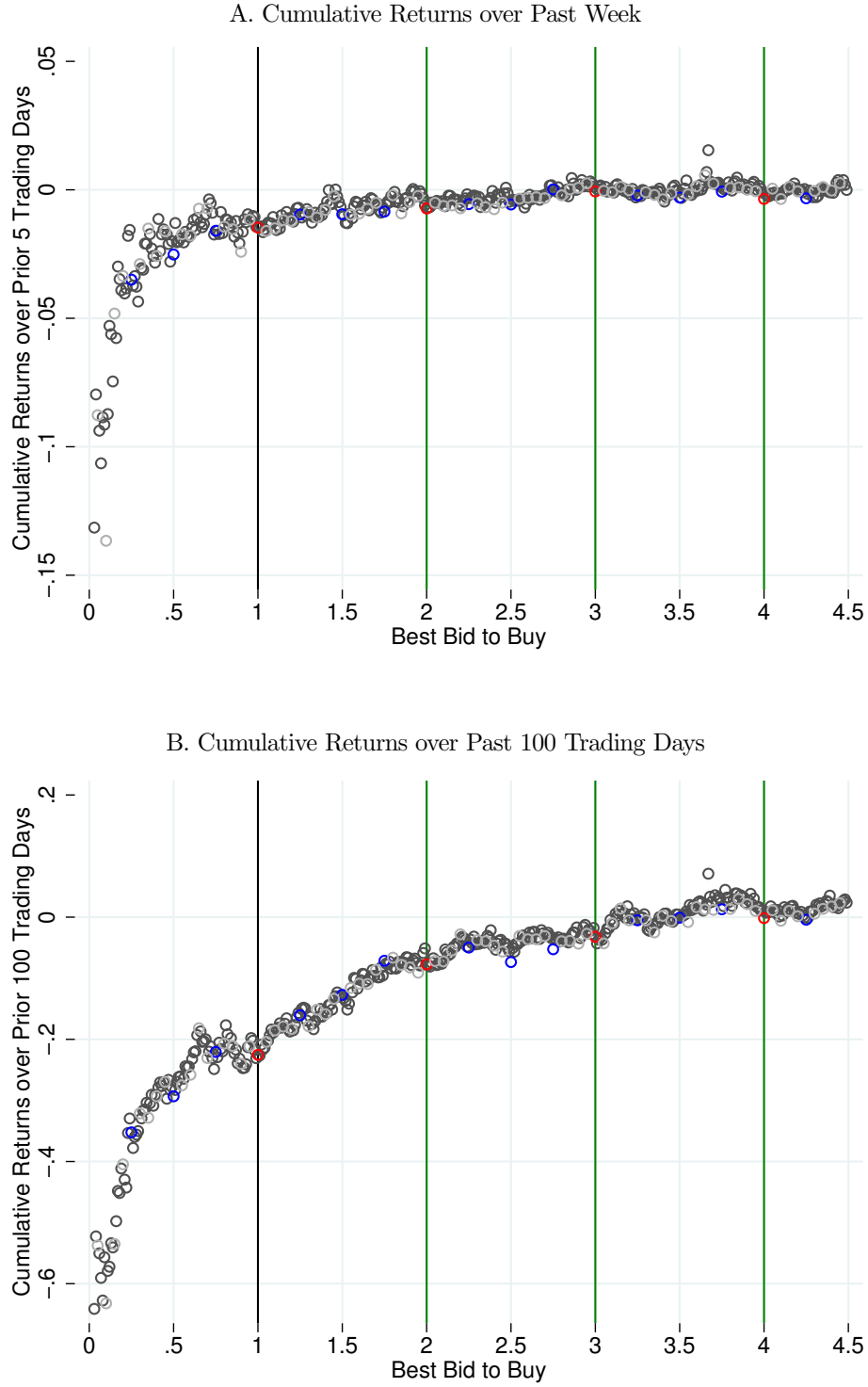
Note: Figure shows the probability, conditional on the NBB truncated to two digits in a given millisecond of date t , of the security being traded at date $t+180$. Estimates based on CRSP and TAQ data. See text for details.

FIGURE 2. IMPLIED VOLATILITY OF OPTIONS MARKET CALL CONTRACTS



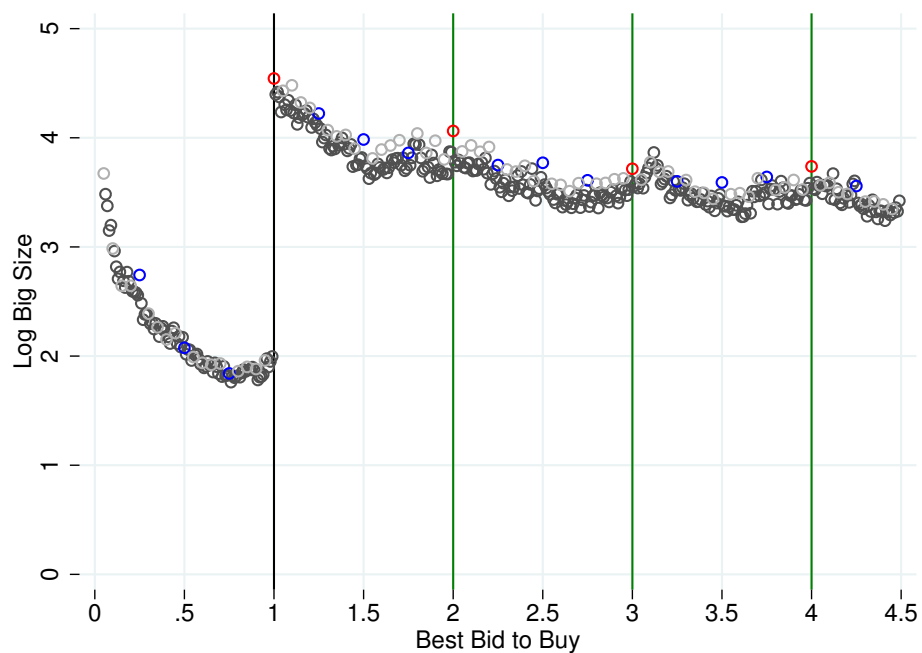
Note: Figure shows the conditional expectation of the implied volatility of options market call contracts given the corresponding closing price in the equities market. Implied volatility based on Black-Scholes using a risk-free rate of 0.5%. Estimates based on OptionMetrics data. See text for details.

FIGURE 3. RELATIONSHIP BETWEEN CURRENT PRICES AND PAST CUMULATIVE RETURNS



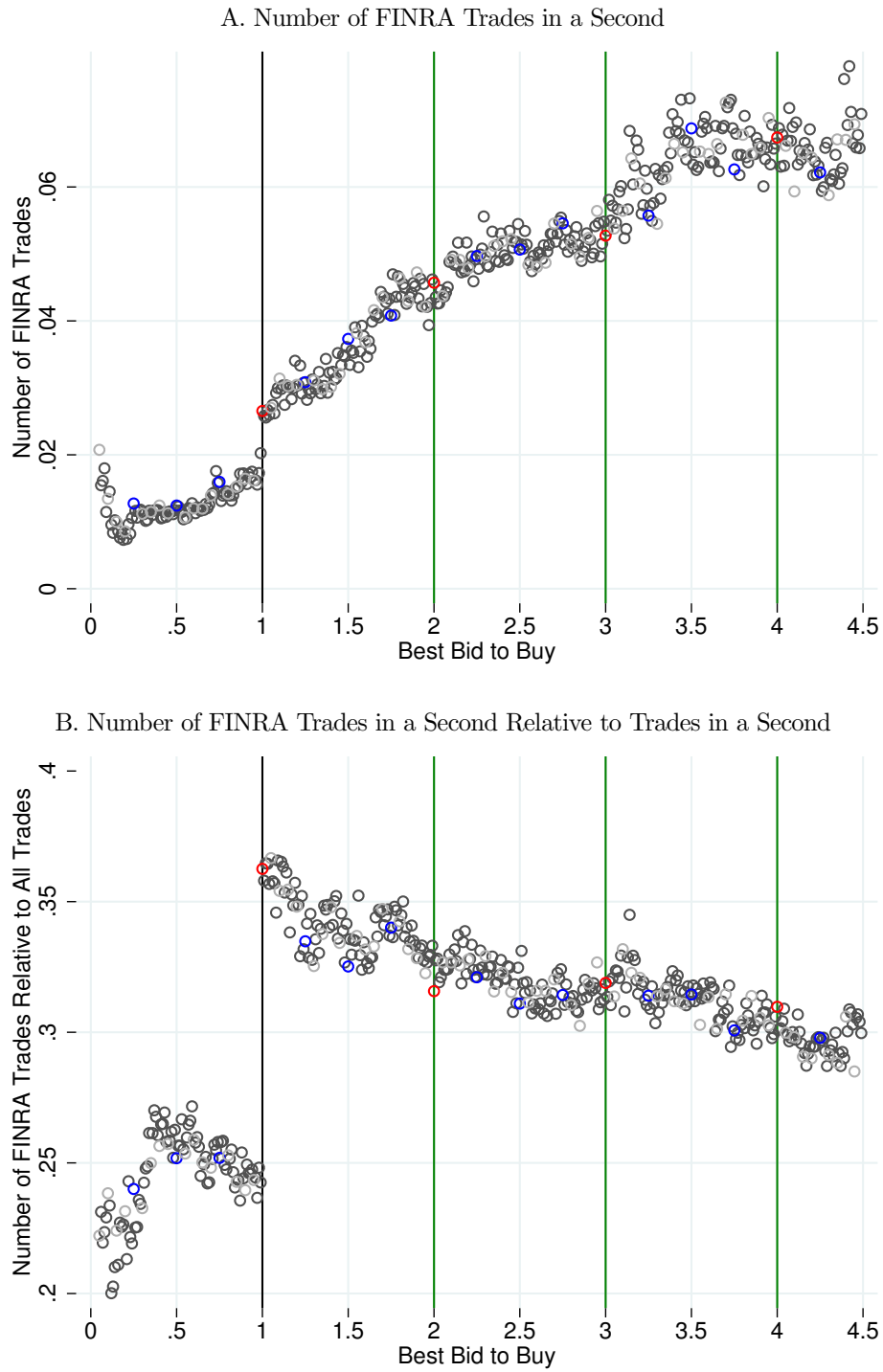
Note: Figure shows conditional expectations of prior returns given the NBB. Panel A (B) shows the conditional expectation of the cumulative returns over the past 5 (100) trading days, calculated as $-1 + \exp(\sum_{t=1}^T \ln(1 + R_{\tau-t}))$, where T is either 5 or 100 and R_{τ} is the daily return. Estimates based on CRSP and TAQ data. See text for details.

FIGURE 4. INSIDE DEPTH



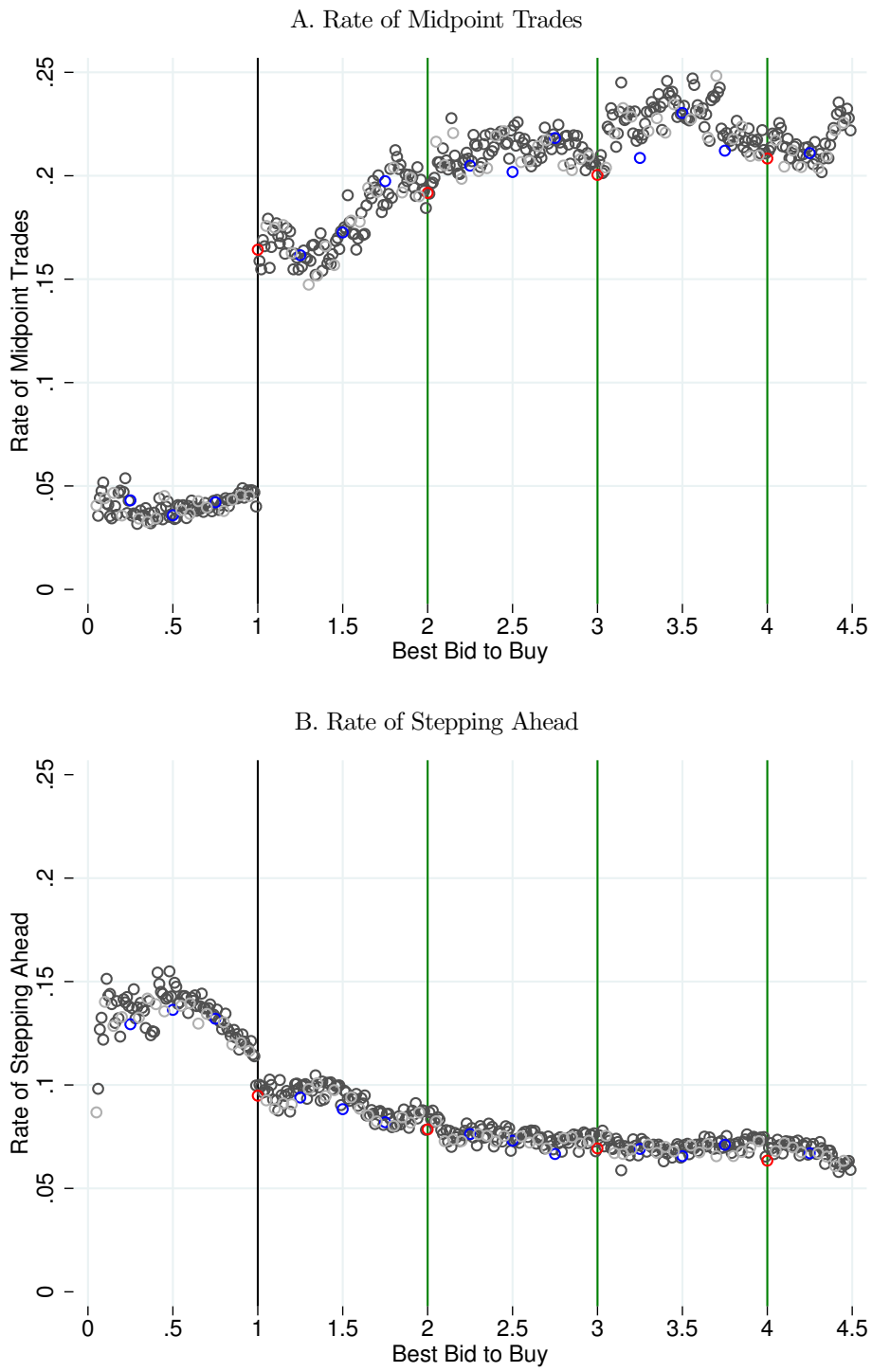
Note: Figure shows the conditional expectation of log bid size given the NBB. Estimates based on TAQ data. See text for details.

FIGURE 5. PREVALENCE OF FINRA TRADES BY PRICE



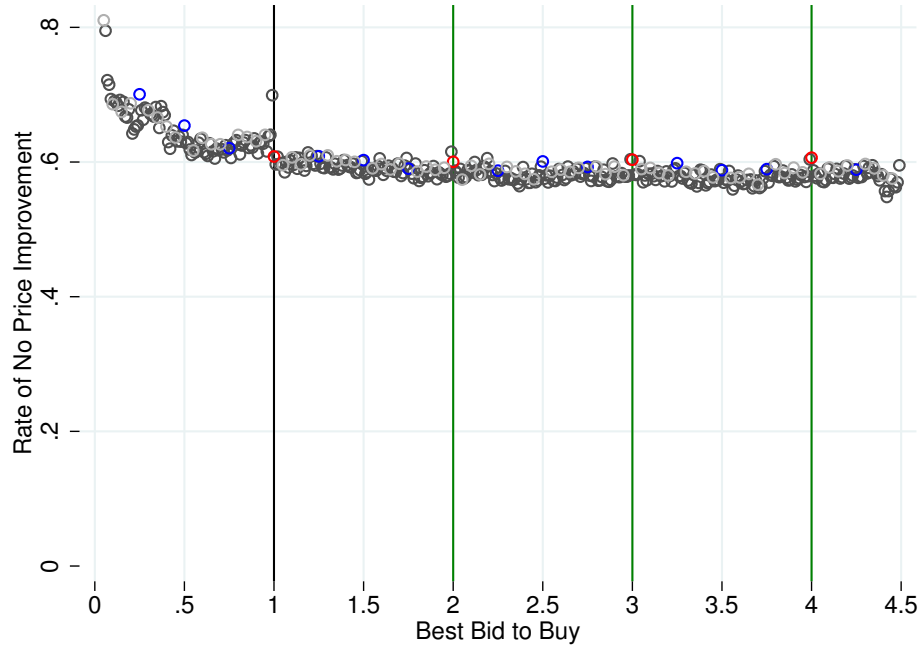
Note: Figure shows the conditional expectation of the number of FINRA trades in a second given the NBB as of the beginning of the second (A) and the fraction of trades in a second that are FINRA trades (B). Estimates based on TAQ data. See text for details.

FIGURE 6. PRICE IMPROVEMENT FOR FINRA TRADES



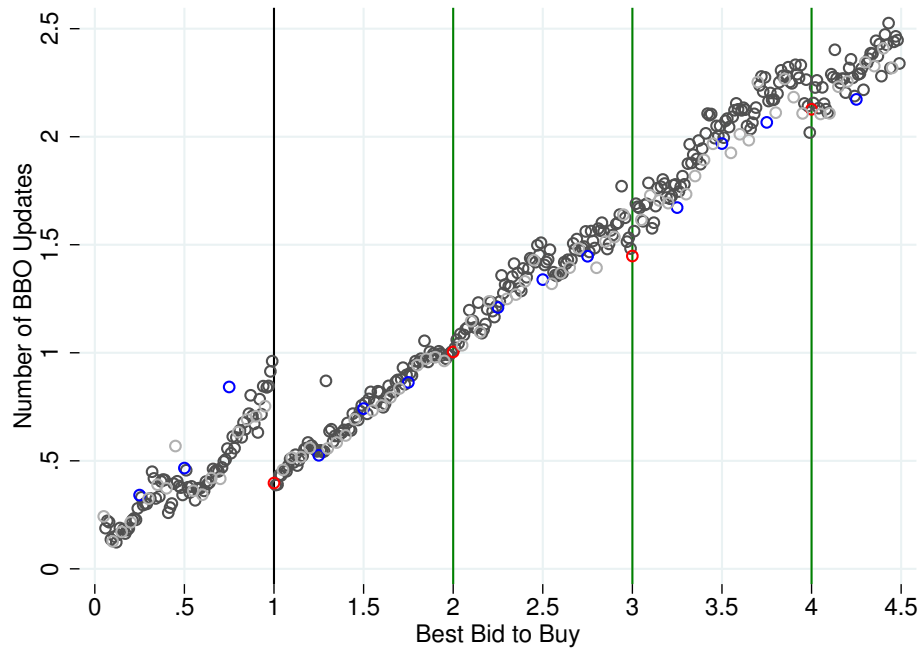
Notes at end of table.

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 price improvement of just 0.0001 exclusive of midpoint pricing; and C shows rate of no price improvement. Estimates based on TAQ data. See text for details.

FIGURE 7. NUMBER OF BBO UPDATES IN A SECOND



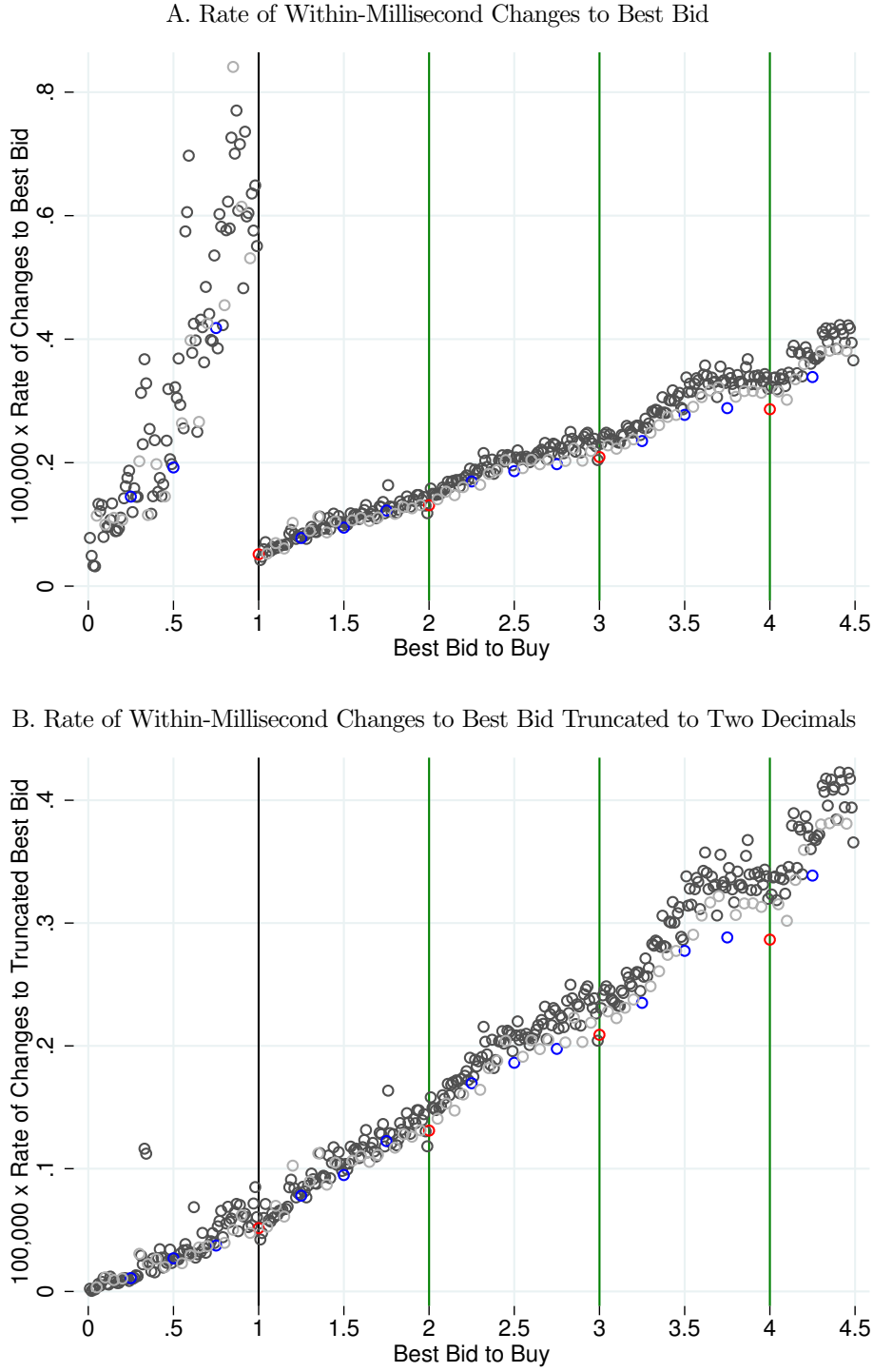
Note: Figure shows the conditional expectation of the number of BBO updates in a second given the NBB at the beginning of the second. Estimates based on TAQ data. See text for details.

FIGURE 8. FLUCTUATIONS IN NBB UPDATES WITHIN ONE SECOND



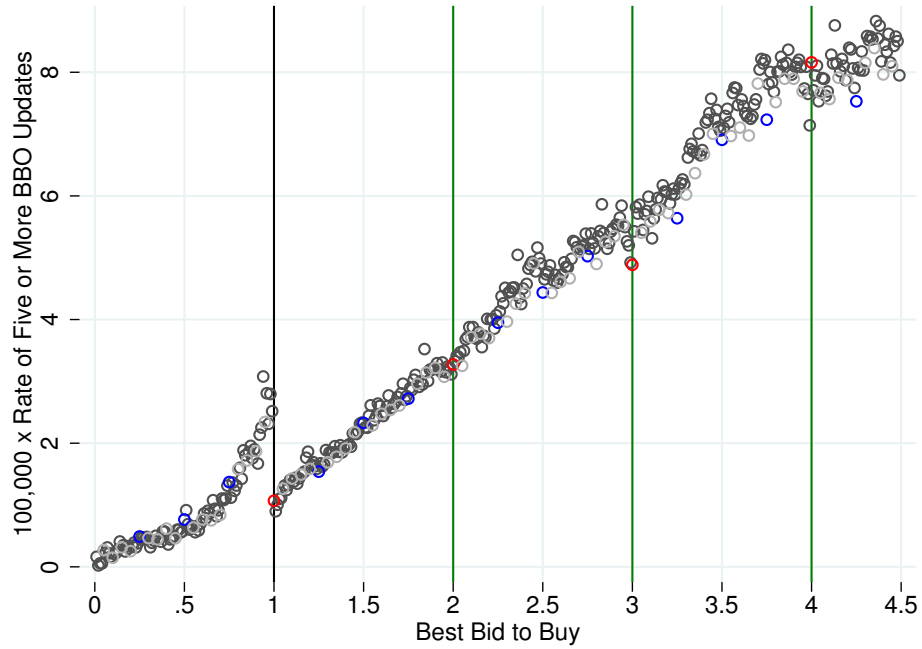
Note: Figure shows rate at which the NBB varies within the second (Panel A) and the rate at which the NBB, truncated to two decimals, varies within the second (Panel B). Estimates based on TAQ data. See text for details.

FIGURE 9. PRESENCE AND EFFECTS OF HIGH FREQUENCY TRADING: WITHIN MILLISECOND EVIDENCE



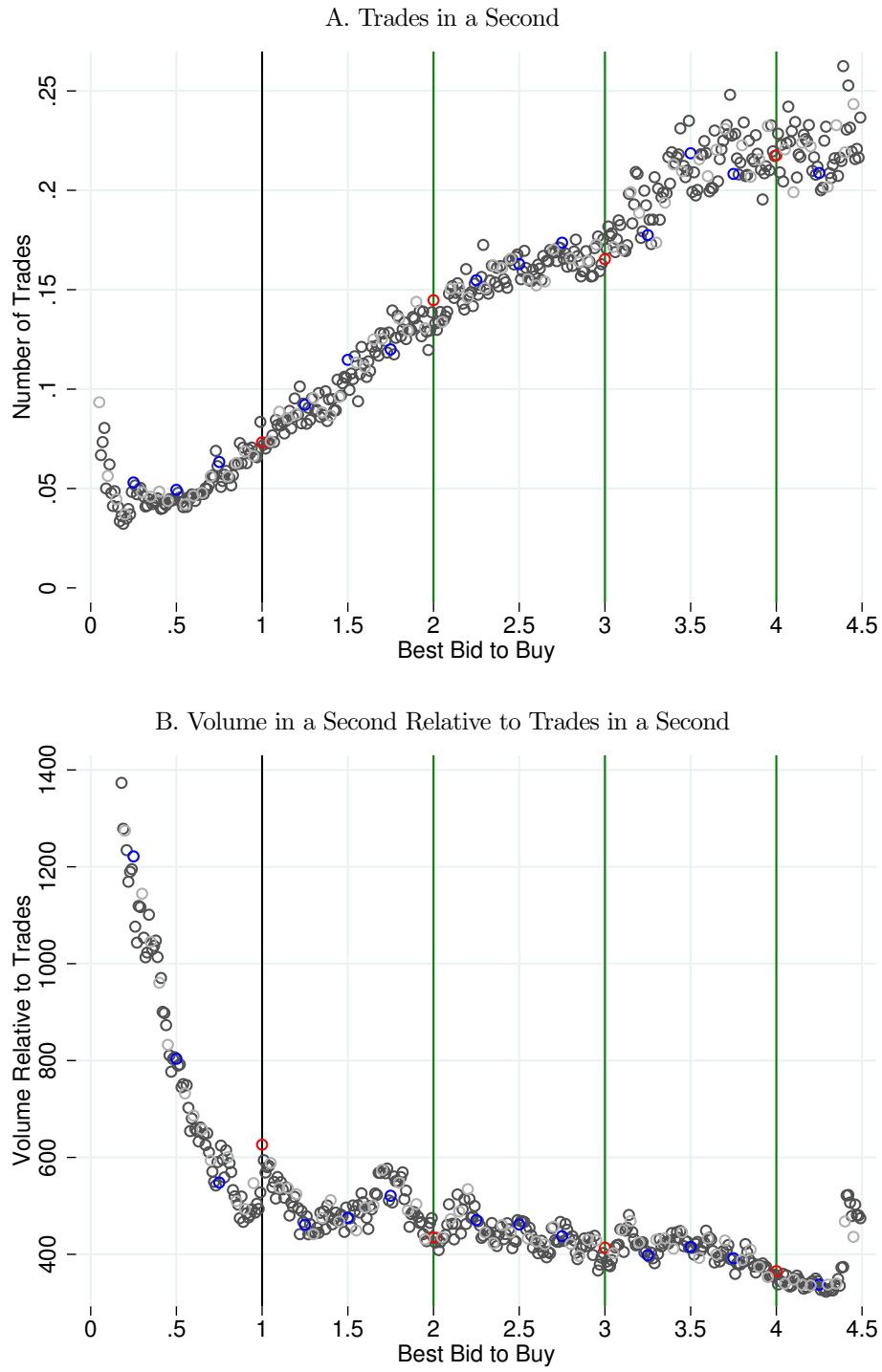
Notes at end of table.

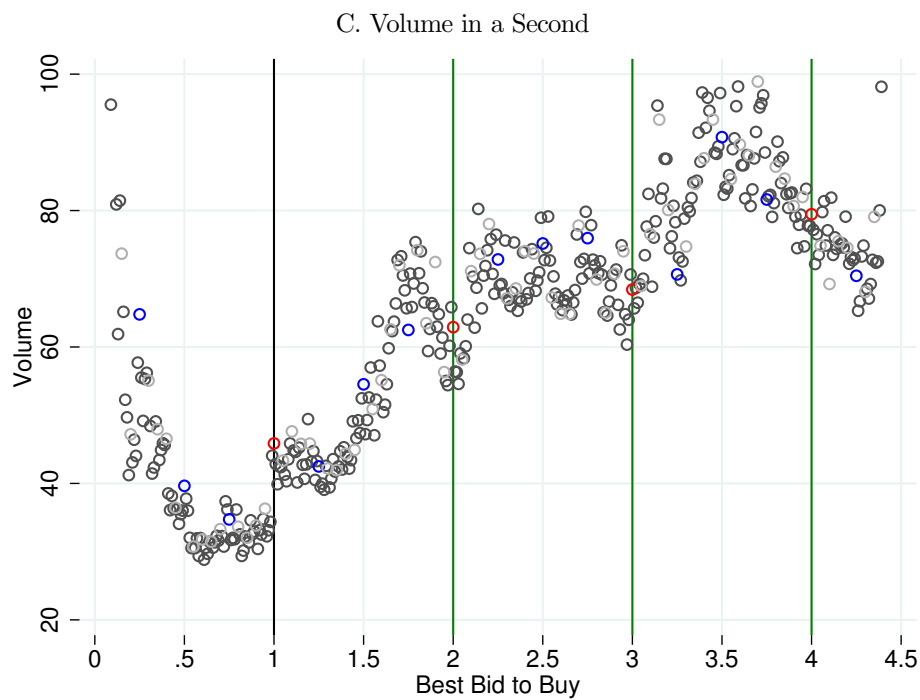
C. Rate of Five or More BBO Updates in a Millisecond



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

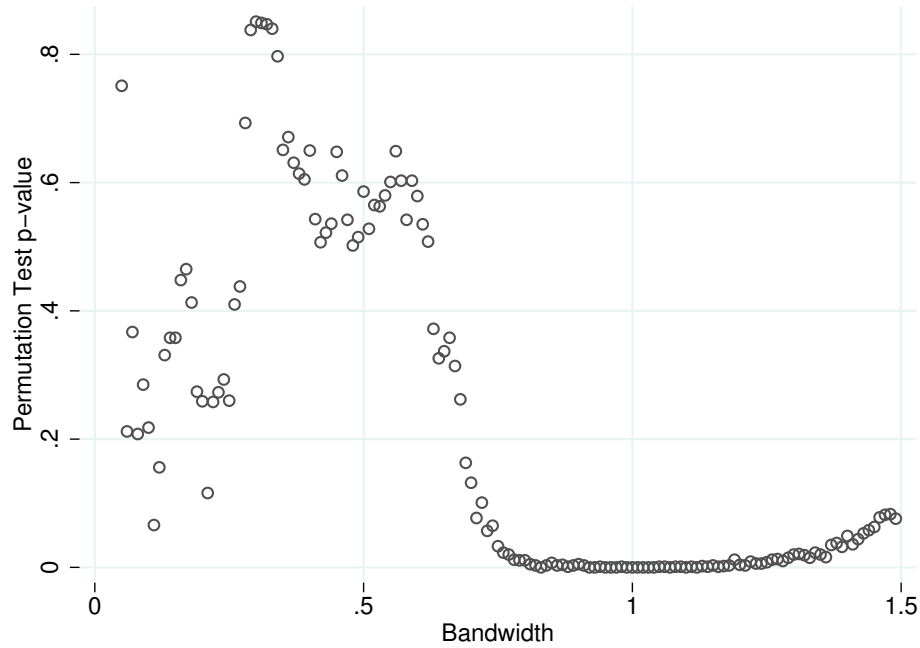
FIGURE 10. MARKET ACTIVITY





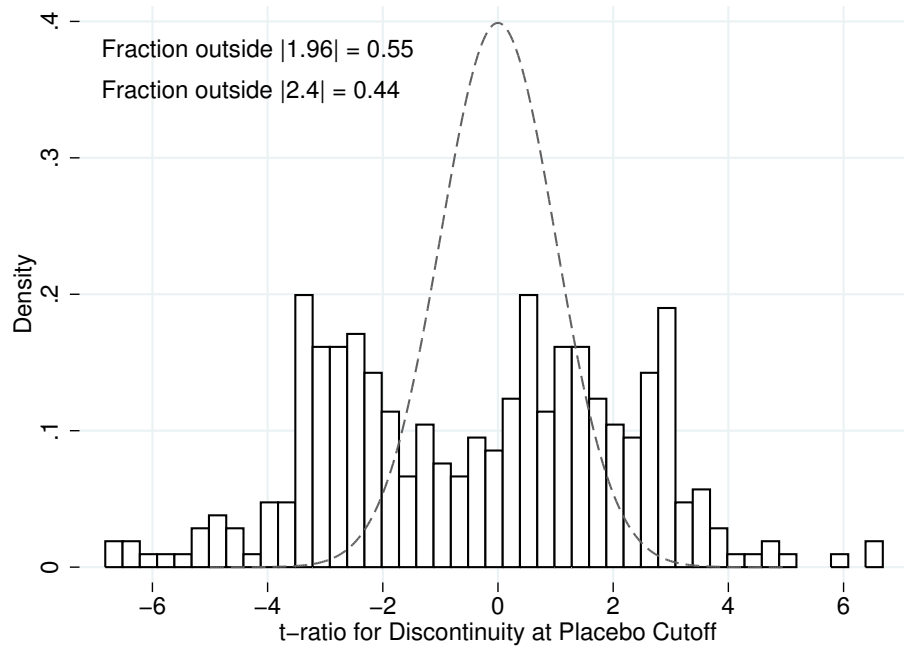
Note: Panels A, B, and C plot the conditional expectation of trades in a second, the conditional expectation of average trade size, and the conditional expectation of volume in a second given the NBB at the beginning of the second, respectively. Estimates based on the TAQ data. For panel B, observations above 1500 excluded for visual purposes. For panel C, observations above 100 excluded for visual purposes.

APPENDIX FIGURE 1. RANDOMIZATION INFERENCE P-VALUE AS A FUNCTION OF BANDWIDTH



Note: For bandwidths over a grid of $\{0.05, 0.06, \dots, 1.5\}$, figure shows randomization inference p-value associated with Hotelling's T^2 and the four covariates from Figures 1, 2, 3A, and 3B.

APPENDIX FIGURE 2. PLACEBO DISTRIBUTION OF t -RATIOS FOR DELISTING RISK



Note: Figure shows the histogram of t -ratios obtained by estimating placebo discontinuities in delisting risk from \$0.50 to \$4.00. The superimposed grey curve is the standard normal density function.